DOCUMENT RESUME

ED 118 482 SO 008 879

TITLE	Child Development and the Housing Environment. Volume 1: Statistical Design and Analysis.
INSTITUTION	Urban Systems Research and Engineering, Inc., Cambridge, Mass.
SPONS AGENCY	Office of Economic Opportunity, Washington, D.C.
PUB DATE	Aug 72
NOTE	214p.; For related documents, see SO 008 880-881
EDRS PRICE	MF-$0.83 HC-$11.37 Plus Postage
DESCRIPTORS	*Child Development; Elementary Secondary Education; *Family Characteristics; Family Environment; Interaction Process Analysis; Measurement Instruments; Parent Child Relationship; *Public Housing; *Research Design; Research Tools; Social Environment; Social Science Research; *Statistical Analysis; Statistical Studies

ABSTRACT

The first part of a three-volume study, this report presents and justifies a research design for investigation of the relationship between the housing environment and the range of child development and family measures. The recommended design is nonexperimental in nature and focuses on comparisons of residents in selected housing programs (publicly supported residential housing complexes). In chapter 1 the results of a computer simulation of the entire investigation are used to calculate optimal decision rules for the conduct of the study. Basic problems of inference that are associated with investigation of the interaction of human subjects with their environment are reviewed in the context of independent variables in chapter 2. Chapters 3 and 4 deal with nonexperimental and experimental approaches respectively, with major emphasis given to the description and documentation of the recommended nonexperimental approach. The final chapter presents recommended instruments for the measurement of the cognitive, socioemotional, and physical development in subject children, plus supplementary measures of family well-being. Three appendices provide descriptions of the instruments and documentation of the simulation study. (Author/ND)

CHILD DEVELOPMENT AND THE HOUSING ENVIRONMENT

VOLUME 1: STATISTICAL DESIGN AND ANALYSIS

2

Urban Systems Research & Engineering, Inc.
1218 Massachusetts Avenue, Cambridge, Massachusetts 02138

CHILD DEVELOPMENT AND THE HOUSING ENVIRONMENT

Volume 1. Statistical Design and Analysis

August 1972

Prepared for
Office of Economic Opportunity
1200 19th Street, N.W.
Washington, D.C. 20506

Prepared by
Urban Systems Research & Engineering, Inc.
1218 Massachusetts Avenue
Cambridge, Massachusetts 02138

This first volume of the Design Study on the relationship between Child Development and the Housing Environment reflects many beneficial influences. I would like to mention in particular the contributions of Mr. Kenneth Alper, Dr. David Cross, Dr. Patricia Greenfield, Dr. David Haynor, Dr. Richard Light, Dr. Don Olivier, and Miss Diana Hobson. Their efforts would have been far less productive had the project not been subject to the exacting standards and sense of purpose of its technical monitors, Dr. Iris Rotberg and Dr. Joy Frechtling; to an unusual degree, they are responsible for what is of merit in this report.

Anthony J. Blackburn
Project Director

Cambridge, Massachusetts

TABLE OF CONTENTS

CHAPTER THREE (continued)

CHAPTER FIVE (continued)

APPENDICES

LIST OF TABLES

CHAPTER ONE

INTRODUCTION

Child growth and development can be influenced not only by educational programs which focus explicitly on the child, but also by broader economic and social efforts which impact on the larger family unit. Programs such as income maintenance, employment, and housing which can directly alter the environment or affect the resources available to parents and families can also, through these changes, significantly affect the child. Little attention has been given, however, to assessing the implications of these broader programs for child development. Typically, only "child development" efforts have been evaluated in terms of their impact on children. Different poverty programs are, however, far from independent, and there is a need for assessing the ways in which differently focused programs serve to meet similar goals.

This study is designed to investigate the effects on child development of housing location change. There are many reasons to expect that housing programs which permit location change will have important and beneficial implications for child development. Location change may be expected to be beneficial to the family and thus the child by allowing escape from oppressive environmental conditions, and by permitting access to conditions that are more favorable. Many aspects of the inner city environment--poverty, crime, drugs, poor municipal services and public facilities, lack of educational opportunity, etc.--make it difficult for individuals forced to live there to achieve desired goals for themselves and their children. Fragmentary support programs which aim at altering one aspect of such families' lives have generally not been found to yield significant long term changes in child development, possibly because they leave the major portion of a debilitating environment unmodified. It is reasonable to hypothesize, however, that programs which lead to more fundamental environmental changes, programs which make it possible for interested families to move to areas with more favorable characteristics--improved physical, social, and educational environments--will have important and long lasting effects on the development of the family and the child.

The purpose of this report is to present and justify a research design for the investigation of the relationship between the housing environment on the one hand, and a range of child development and family measures on the other. The recommended design is non-experimental in nature and focuses upon comparisons of residents in selected housing programs; housing programs are here defined as publicly supported residential housing complexes. Because of the non-experimental nature of the proposed design, the danger of self-selection bias in inter-program comparisons is quite severe, and considerable emphasis has been given to the resolution of this problem in the development of the proposed research design. This emphasis finds expression in the sequential nature of the design proposed and in the introduction of tests for self-selection at each phase. A computer simulation of the entire investigation has been conducted using a variety of assumptions with respect to the existence and nature of self-selection, and the results of this simulation have been used to calculate optimal decision rules for the conduct of the study.

The choice of a non-experimental approach to this investigation was not made without a thorough consideration of its relative advantages and disadvantages. The basic problems of inference which are associated with investigation of the interaction of human subjects with their environment are reviewed in the context of the independent variables of interest in Chapter Two of this report. Chapters Three and Four deal with non-experimental and experimental approaches respectively, with major emphasis being given to the description and documentation of the recommended non-experimental approach. The final chapter presents recommended instruments for the measurement of the cognitive, socio-emotional, and physical development in subject children, together with supplementary measures of family well-being. Lastly, three appendices provide descriptions of the recommended measurement instruments and documentation of the simulation study.

CHAPTER TWO

PROBLEMS OF INFERENCE IN THE INVESTIGATION OF ENVIRONMENTAL CHANGE

The basic objectives of this study relate to a conjectured causal relationship between, on the one hand, the environment of home and neighborhood in which a child grows up and, on the other, his ability to learn and to stay healthy. Realization of these objectives is unlikely if the associated problems of inference are not carefully and systematically examined.

Statement of the hypotheses under investigation requires identification of the dependent variables which characterize child development and of the independent variables which characterize the child's environment. Statement of the hypotheses also requires the identification of the subject population for whom the hypotheses are expected to hold. Other things being equal, it is desirable that the subject population should be as inclusive as possible, thus investigation of the impact of the housing environment on the development of the child from families of all income groups, racial backgrounds, and current housing types will, in general, be more useful than an investigation of only a subset of the population. Unfortunately, there is frequently a conflict between a desire to make the investigation as inclusive as possible and the necessity to isolate causal, rather than statistical, associations between variables.

In order to identify causal relationships between variables, it is necessary to design experiments so that the levels of selected factors can be varied while all other factors are held constant. In the ideal experiment, all the factors likely to affect child development variables can be controlled and independently manipulated. Subjects can be independently sampled and randomly assigned to planned combinations of conditions. In such an experiment the effects or impact of a program can be partitioned into additive component effects attributable to the separate factors or combinations of factors. Hypotheses of the form "if A then B" can then be specifically tested. Unfortunately, it is never possible to know all the pertinent variables that influence an outcome, nor can they always be controlled even if they are known. Randomization helps here by counter-balancing the biases that the unknown factors introduce. Since each member in the population has equal opportunity of being selected by random sampling, members with certain distinguishing characteristics will be balanced, in the long run, by members with opposite characteristics.

To meet the requirements of such an ideal experiment, it must be possible both to control the variables which characterize the housing treatments and also to assign families at random to alternative treatments. In the context of this design, and indeed of most social science designs, neither of these conditions can be assumed. In the first place, control over the variables which characterize the housing treatments is limited by what is available in the real world--and the real world does not provide planned combinations of conditions. Secondly, it is generally necessary to work with subjects who have assigned themselves to treatments, and no matter how skillfully "matching" or covariance adjustments are carried out, the basic problem of post factum research remains: groups can be matched on a hundred variables and still differ on one which is relevant to the level of child development.

In developing the designs presented in this section, consideration has been given to genuine experimental settings in which families are randomly assigned to alternative treatments and in which planned combinations of housing characteristics are developed. The methodological advantages of these alternatives are only achieved at great expense, however, and ordinary prudence suggests that they should not be undertaken without first conducting a careful and well-designed analysis of the data which the world already provides. The emphasis of this work has, therefore, been on the development of sampling designs which overcome, to the greatest extent possible, the limitations of the post factum experiment.

2.1 Problems of Self-Selection in the Design

In reviewing the general problems of post factum experiments, emphasis has been placed on the problems of inference caused by the self-assignment of individuals to different housing treatments. In contrasting the independent (child development) variables across different treatments, it is difficult to be sure whether the observed variation is caused by the variation in treatments or by differences in the populations which have assigned themselves to these treatments. It would, for example, be clearly absurd to contrast the achievement levels of different housing groups in a random sample of all families in the U.S. and thereby infer a causal relationship between housing and child development, although a highly significant association would undoubtedly exist.

Two kinds of self-selection bias may be present in the investigation of the relationship between housing and child development. The first of these is the bias introduced when the entering populations of different housing programs are not comparable with respect to the variables which influence child development. Less obvious is the bias which results from differences in the characteristics of families exiting the program; if mobility is correlated, either positively or negatively with dependent variables, comparisons between resident and control groups will tend either to under- or over-estimate effects; this is still further complicated when mobility itself varies between treatments.

Attempts to minimize the self-assignment or self-selection problem must necessarily focus on insuring that the populations assigned to different treatments are, as far as is possible, comparable in all dimensions relevant to child development. At a minimum--and most designs stop at a minimum--this requires that the treatment groups be matched on socio-economic and family variables; thus, for example, it may be necessary to quota sample within the treatment populations to insure comparability of groups with respect to income, ethnic mix, family size, family structure, and sex of child. This, by itself, does not provide adequate protection against problems of self-selection, for one critical difference between the treatment groups remains: their choice of treatment. Nor is this simply a theoretical objection. To the extent that certain housing and school environments are conducive to healthy child development, parental attitudes and methods of child-rearing may well be associated with particular housing preference patterns. We cannot, therefore, infer that observed differences between the children of matched families living in suburban neighborhoods and in dilapidated inner-city housing are attributable to the housing/neighborhood change alone, for if the inner-city children were to be reassigned to the suburban homes, and the suburban children were to be reassigned to the inner-city homes, they would take with them parents who had made a fundamentally different choice about the way they should lead their lives.

There is an obvious conflict between the logic of this argument and the basic requirements of statistical inference. To achieve comparability amongst treatment groups, it is necessary to insure that their housing choices are similar; to estimate the impact of different treatments it is necessary to insure that housing choices are different. The conflict is similar where mobility leads to inter-temporal self-selection; to avoid bias, comparison groups should have similar lengths of residence; to permit contrasts, their lengths of residence should be different.

It is clearly necessary to effect a compromise between the need for variation in treatments and the need for homogeneity amongst treatment groups. Efforts to achieve such a compromise are reflected in the designs proposed. There are two complementary ways to approach the problem: in the first place the choices of treatments should be carried out in such a way as to provide reasonable assurance that the entering populations will be comparable. In the designs proposed this is achieved through an exclusive emphasis on subsidized housing programs. This provides some assurance that treatment groups are comparable in the sense that they are all families for whom subsidized housing is an acceptable alternative.

The second approach to the problem of self-selection bias is to test for its presence. In the designs developed, tests for the homogeneity of both entering and exiting populations are included. If these tests are passed, the findings of the investigation would be extremely hard to refute.

In reviewing the treatment of self-selection problems in the designs proposed, it is fair to say that the attempt to solve these problems has exercised a profound influence on the choices made. The early thinking about these designs included work on methods of analysis which would allocate developmental gains between successive housing environments and which would require the tracing of families moving through both private and public housing markets. Apart from other questions of feasibility, this approach could not be reconciled with the theoretical objections reviewed above. It is fair to acknowledge that by concentrating on residents of publicly supported housing programs, a number of interesting hypotheses can no longer be investigated. Nevertheless, the hypotheses which remain are of vital interest to those concerned with the role of the government in the sponsorship of good housing.

2.2 Selection of Independent Variables

2.2.1 Independent Environmental Variables

In the non-experimental setting of these designs, the variables that characterize a housing program cannot be directly controlled or manipulated. The selection of independent variables in this design is therefore determined by the decision to include or to exclude existing or planned public housing programs, and these decisions are, in turn, constrained by what is available. Within these constraints, it is possible to emphasize selected dimensions of variation in program characteristics and the choice of emphasis must necessarily reflect the nature of the hypotheses under investigation. The principal emphasis of the designs proposed is upon the socio-economic dimensions of relocation rather than upon its physical and spatial aspects. This emphasis reflects the view that social and economic change is a potentially more powerful influence on the development of children than is improvement in the physical and spatial aspects of home and neighborhood.

Alternative approaches to the measurement of socio-economic change can be contrasted in terms of the decision to use a composite or non-composite measure and in terms of the choice of variables. Composite measures lump together several variables, such as household income, occupational distribution, educational achievement, and so on, to form an index of the socio-economic status of a selected population. The difficulties associated with comprehensive measures of this kind are generally the availability of data, the non-operational nature of the measure and the arbitrary element associated with choice of the index weighting scheme. These difficulties all point in the direction of a simple measure of socio-economic change if there is one available which is highly correlated with omitted measures in the populations of interest. Household income is the obvious candidate in this application, being highly correlated with other measures of socio-economic status, such as occupational breakdown and educational achievement. Since household income data is available from the 1970 Census at a relatively fine level of spatial disaggregation, its use as the characterizing variable seems highly desirable.

In characterizing environmental change in terms of its socio-economic dimensions, a distinction can be made between the program itself, the neighborhood in which it is located and the school in which its children are enrolled. All three can, in principle, exhibit independent variation with respect to the socio-economic mix of their members, but the real world with its limited range of combinations may not provide an adequate base for investigation of the separate contributions of each. In particular there will tend to be a high degree of collinearity between the socio-economic status of the neighborhood and of the school populations; indeed if the neighborhood is defined as coterminous with the school district, there is an effective problem of singularity. These considerations--as well as considerations of cost and feasibility--lead to the elimination of either school or neighborhood as an independent source of variation.

Two dimensions of environmental change remain: the neighborhood and the program itself. These together determine the income mix of the program child's peer group. The difficulty here is not to define variables which characterize neighborhood and program SES, but to find programs which provide adequate independent variation along both of these dimensions. In the real world, the income balance of program and neighborhood are heavily confounded with very few low-income programs in high-income neighborhoods, and still more elusive, moderate-income programs in low-income areas. Only through a comprehensive search of programs has it been possible to complete the designs presented here.

2.2.2 Independent Family and Child Variables

The primary hypotheses of this investigation relate to the effects of socio-economic integration on the development of children in subsidized housing programs; implicit in this hypothesis is the assumption that differences between the socio-economic status of the subject and that of his neighbors may affect the subject's rate of development. It follows, therefore, that those family and child variables which are associated with cultural or economic deprivation in urban communities should be treated as independent variables in this analysis.

There are, in addition, characteristics of the family and the child which, although they may be influential in terms of child development, do not interact directly with the socio-economic variables used to characterize environmental change. Variables such as family size, family structure, and sex of child must therefore be experimentally controlled in the sense that the samples must be balanced across treatment combinations--but they need not be subject to special investigation. These, together with those variables which must be statistically controlled, such as age of respondent, constitute the set of "nuisance variables" in the designs which follow.

The requirement that samples be balanced across programs with respect to both the independent variables and the experimentally controlled nuisance variables constitutes quite a severe constraint on the selection of housing programs, since it must be possible to find appropriately balanced samples in all the programs.

CHAPTER THREE

NON-EXPERIMENTAL DESIGNS FOR HOUSING PROGRAM IMPACT

The principal emphasis of the design effort has been upon the development of non-experimental approaches to the investigation of child-development/housing interaction. The difficulties of the non-experimental approach, and in particular those relating to self-selection bias, have already been discussed; the emphasis on non-experimental designs does not, therefore, reflect any lack of appreciation of the extent and nature of the associated problems of inference.

The emphasis on non-experimental approaches does, however, represent a judgment with respect to the comparative cost-effectiveness of experimental and non-experimental approaches. Experimental approaches are characteristically costly and difficult to administer--and housing experiments are no exception to this rule. Commitments to families participating in any housing experiment cannot be transient, and the costs of data collection and analysis, which represent the complete expense for the non-experimental approach, characteristically represent only a small part of the total costs of a truly experimental design. If there were no conceptually adequate non-experimental approach available, the cost difference would not, in itself, be sufficient reason for rejecting the experimental design. If there are grounds for believing that a conceptually adequate non-experimental approach exists, however, it is extremely hard to justify a more elaborate expensive recommendation.

The choice between non-experimental and experimental approaches is not the only question of strategy. Within the non-experimental approach itself, there are a set of alternative methods for collecting and analyzing data; and the choice among these strategies depends upon prior judgments as to the validity of the hypotheses themselves and the reliability of the methods by which they are to be investigated. These issues and their relationship to the manner in which the non-experimental approach is to be implemented, are now reviewed in detail.

3.1 Strategies for the Non-Experimental Investigation

In reviewing the strategies for a major non-experimental investigation of housing/child development interaction, it is useful to distinguish between all the logically possible outcomes. These outcomes can be summarized in terms of change in the dependent variables, the influence of the independent variables, and the validity of the assumptions underlying the design.

It is useful to begin by considering a very general type of design in which there are several child development--or outcome--variables on which measurements are made. These measurements are made on children in programs which differ with respect to the independent variables of interest. In analyzing the variation in the dependent measures between programs there are two mutually exclusive and collectively exhaustive possibilities:

(1) Significant variation between programs on at least some outcome measures

(2) No significant variation between programs on any outcome measure.

In the last instance, since there are no detectable differences between programs along any dimension of child development, efforts to explain it through contrived variation in program variables are clearly doomed. In the second case, hindsight might suggest that some of the outcome measures should not be investigated; to the extent that it is possible to extend the investigation of those measures which show significant variation between programs at the expense of those which do not, this is clearly desirable.

It has already been pointed out that if no significant variation between programs is found on any outcome measure, investigation of the independent variables is not a useful exercise. If significant variation in some or all outcome measures does exist, however, it does not follow that this variation will necessarily be associated with the planned variation in the selected program variables. The programs included in the design will vary amongst each other, not only with respect to the independent variables selected for investigation, but also with respect to others which cannot be completely controlled experimentally. It is, therefore, conceivable that the observed variation in outcome measures cannot be effectively explained by reference to the independent variables selected for analysis. Even when the included variables do appear to "explain" much of the observed variation, this may simply reflect correlation between included variables and the excluded variables (supportive services, tenant management, etc.) which are the principal sources of variation. This aspect will be discussed in the context of the implicit assumptions of the design. The set of logical possibilities relating to investigation of the influence of selected independent variables are:

(1) Selected independent variables "explain" variation

(2) Selected independent variables fail to "explain" variation

As already discussed, the validity of results obtained through a non-experimental study depends upon the elimination of self-selection and upon the experimental matching of programs upon those excluded program variables which may affect child development. There can never be any complete assurance that all the relevant program variables are adequately treated, but there are methods of testing for the presence of self-selection. In either event, the detection of significant variation in outcome measures between programs and the "explanation" of a significant part of this variation by means of the selected independent variables will be valueless and misleading if the premises on which causal inference is based cannot be justified.

These considerations have led to the recommendation that the study be conducted in three distinguishable phases; the scope and objectives of each of these phases can be summarized as follows:

Phase I. The Phase I study is essentially a pilot study involving a comparison of measured child development in two program classifications, differing with respect to the level of both environmental variables. The objectives of the pilot include: test of significant differences between programs, test of homogeneity of subjects between programs and screening of outcome measures.

Phase II. The Phase II study involves data collection and analysis in additional program classifications. The data will be collected in one wave and analyzed jointly with the Phase I data. The objectives of this phase are to test the explanatory power of the independent program variables, to test for homogeneity of subjects between programs and to analyze possibly extended outcome measures.

Phase III. The Phase III study uses the data from Phases I and II together with data from the same programs collected in a single subsequent wave. The data from these two waves are then analyzed longitudinally. The Phase III effort is similar in its objectives to Phase II, but, being longitudinal, it endows the tests with much greater power of discrimination. It also provides an opportunity to test for inter-temporal self-selection resulting from mobility in the subject population.

It is useful to regard these three phases as steps in a sequential sampling and decision-making exercise. The data collected at each phase are used to determine whether or not to continue the investigation, and, if the answer is affirmative, these data are then reanalyzed in the framework of the next stage. Since nothing is thrown away, the sequential conduct of the investigation does not lead to any significant increase in cost even if all three phases are implemented.

The sequential conduct of the investigation is of no value unless rules are developed, in advance, which tell the investigator under what conditions the experiment should be continued. These rules cannot be based on conventional 5% or 10% levels of significance because such rules are arbitrary and bear no relationship to the objectives of the study. Thus in the first phase, for example, several of the programs are quite new and the maximum length of residence is in some instances less than two years; under these circumstances it is not reasonable to expect the data to reflect very large gains in measured child development; yet use of a significance test for effects at the 5% level might lead to rejection of the (true) hypothesis that effects exist over 50% of the time.

To deal with this problem, a computer simulation of the investigation through all three phases has been developed. This simulation, which is described in section 3.6, has permitted the development of decision rules (i.e., critical values of test statistics) by attaching payoffs to outcomes such as finding an effect when one exists and finding an effect when no effect exists. These rules form an integral part of the three phase design.

3.2 Selection of Sample Populations

The three phases of the non-experimental study can be distinguished from each other in terms of their inclusion of programs (Phase I vs. Phases II and III). All phases share a common requirement that the sample populations be comparable across programs, in the sense that observed variations in child development are properly attributable to the differences between programs rather than to the differences between the characteristics of the children and families sampled. The methods by which sample populations are to be selected to meet--or attempt to meet--this requirement are now reviewed.

The first consideration relates to the need to control for age amongst the population of subject children. Because it is reasonable to hypothesize that the environmental effects of relocation will differ according to the developmental stage of the child, it was considered desirable to separate subjects into three groups based on the age of the child. These groups are defined in terms of age range at the time of first testing; they are preschool (0-4), elementary school (5-12), and high school (13-17). These age ranges were selected with a view to reducing variability in scores attributable to the emotional transitions of latency and puberty.

It is useful to distinguish, for the remainder of this section, between independent and nuisance variables. With respect to the independent family variables--SES, race and program residence--the design itself will require that the populations sampled in each program can be broken down into specified proportions of families belonging to each of these groups. Subject to the constraints which the treatment of the independent variables imposes, the subjects must then be chosen to ensure comparability along other objective dimensions, such as family structure, sex, and size of family; these are the so-called nuisance variables. The importance of the distinction between independent and nuisance variables for this discussion is that the composition of the sample with respect to the independent variables is fixed by the design, while its composition with respect to the nuisance variables may be anything--as long as it is the same for each configuration of the independent (program and family) variables.[1] With this distinction in mind, the procedures for selecting the sample populations are now reviewed.

3.2.1 Sample Selection--Independent Variables

There are three basic independent classifications of the sample population within each program. Those classifications are: family SES, race of family and program residence. Each is now reviewed in turn.

Family Income

The conjectured benefits to children which result from movement into an enriched environment are, at least in part, assumed to result from increased contact with peers, both in and out of school, who have had the benefits of higher social and economic status conferred on them. In order to examine this effect, the sample population must be made to include housing program families of both low and moderate income. It is desirable and it also appears to be feasible, to sample low and moderate income families in equal proportions within

---

[1]While it is true that from a statistical viewpoint, it is only necessary to control for the nuisance variables somehow, the manner in which it is done affects the interpretation of the results. Thus, if only large, one-parent families are included, the results are only generalizable to other large, one-parent families. These considerations suggest the need to insure that the families included are reasonable representatives of the population of interest.

the selected programs. For the purposes of defining these categories, it is important to maintain a consistency across programs, not provided by local eligibility criteria. Accordingly, selection of low and moderate income definitions will reflect the need to insure adequacy of sample sizes in each program classification and the need to adjust categories to reflect regional cost-of-living variations.

Race of Family

In comparing programs with significantly different racial compositions, it is reasonable to suppose that the impact of these programs will vary according to the race of the family sampled. Thus, the advantages of the economic integration achieved through program participation may be inhibited if they are masked by racial discrimination in the neighborhood or in the schools. Other things being equal, it is clearly desirable to achieve as detailed an ethnic breakdown as possible, but the extent to which this can be done is limited by the availability of programs. For this reason, only white and black families will be included, and it is recommended that the subject populations be equally divided between the two.

Program Residence

On each of the selected sites, samples will be drawn from families who are program residents and from families who have applied for residence, but who are not yet residents. The applicant families serve two purposes in this design; firstly, by comparing residents with applicants it is possible to test for the effects of program residence on families and children; secondly, by comparing applicant groups across sites, it is possible to test for the existence of self-selection bias. Both these purposes require that the applicant group on each site should resemble the resident group on that site at its times of entry. Systematic variation in the eligibility requirements for a selected program is therefore undesirable.

The effects of environmental change, if they exist, no doubt take place gradually; other things being equal, therefore, the longer the "treatment," the greater will be the effects. To avoid confounding the influence of length of residence with the independent program and neighborhood variables, it is therefore necessary to control for length of residence in the resident samples. This can be achieved most effectively by attempting to insure experimental balance between samples drawn from different programs; it is also desirable to introduce a lower limit on length of residence in order to allow some opportunity for differentiating residents from non-residents. Since all of the selected programs are less than two years old at the time of this report, a lower limit of eighteen months appears to be a reasonable compromise between feasibility and longer residence.

3.2.2 Sample Selection--Nuisance Variables

Within each of the independent variable sub-classifications, it is necessary to insure that the populations are comparable with respect to the so-called nuisance variables; failure to insure balanced samples with respect to these variables may lead to spurious significance in both program and family variables.

This design calls for the experimental control of four nuisance variables. These are: family structure, family size, sex of child, and previous residence.

Family Structure

This is the euphemism for whether or not there are two parents in the family. Some of the programs under consideration have significantly more one-parent families than two-parent families, and the problem will be to select a ratio between one-parent and two-parent families which is feasible for all programs. Inspection of the housing program data suggests than up to 20% of those selected should come from one-parent families.

Family Size

It is necessary to control for family size in the selection of samples. This is most conveniently done by considering two categories: families with fewer than three children and families with three or more children. Even though there may be some birth-order effects within these categories, they should not be large.

Sex of Child

It is probably adequate to omit sex as a nuisance variable entirely and simply to trust randomization to insure balance between groups. As an additional caution, a check can be run on the proportions of male and female children in the population.

Previous Residence

If the hypotheses of environmental change are true, it is essential that families be experimentally controlled with respect to previous residence for the resident group and current residence for the applicant group. The way in which they are controlled must necessarily reflect the basic emphasis of this study, which is upon significant environmental change. For these reasons, only families from low income neighborhoods will be included; since previous residence will be highly correlated with family SES, elimination of families from the sample on this criterion will probably only affect the moderate income families to any great degree. If it proves impossible to find adequate samples of moderate income families satisfying the inner-city previous residence constraint, this constraint may have to be relaxed. This would have the effect of confounding family income and previous residence variables.

3.2.3 Sample Selection--Procedures

Selection of samples to meet the criteria stated above will be conducted on the basis of the parent interview data.[1] It would be disingenuous to assume that no difficulties will be experienced in meeting the fairly stringent requirements which the sample selection criteria impose, and it is recognized that some changes in the recommended ratios within each category may be necessary after inspection of parent interview data. It may be useful, however, to review a worked out, but hypothetical example of selection within a given program.

In this program it is assumed that there are 75 children, all similar with respect to previous residence. These are distributed between categories as shown in Table 1. By selecting samples from each cell of sizes shown in the bottom right-hand corner of the cell, it is possible to get a well-balanced sample of 32 children within this age group.

In this example, there is some limiting confounding of the independent and nuisance variables, but it is effectively negligible as can be seen from Table 2.

---

[1]See Parent Interview (Part I), Screening, Appendix A.

Family Income		Moderate				Low				
Race of Family		Black		White		Black		White		
Program Residents		Yes	No	Yes	No	Yes	No	Yes	No	
Two parents	3*	3/1	1/1	2/1	4/1	2/2	4/1	1/1	0/0	17/8
	2	1/1	2/1	2/2	1/1	2/0	1/0	4/2	2/1	15/8
One parent	3	2/1	0/0	1/1	4/1	8/1	10/1	1/1	3/2	29/8
	2	2/1	3/2	0/0	2/1	3/1	3/2	0/0	1/1	14/8
TOTAL		8/4	6/4	5/4	11/4	15/4	18/4	6/4	6/4	75/32

*Family size

Table 1

Hypothetical Example of Selection Within a Given Program

		Income		Race		Resident	
		Mod.	Low	Black	White	Yes	No
Family structure	1 par.	7	9	9	7	6	10
	2 par.	9	7	7	9	10	6
Family size	Small	9	7	8	8	7	9
	Large	7	9	8	8	9	7

Table 2

Distributions of Hypothetical Sample Observations by Selected Independent Variables

3.3 Non-Experimental Design--Phase I (Pilot)

3.3.1 Objectives of the Pilot Study

The pilot study is designed to achieve two principal objectives. These are:

(1) A test for the comparability of subject populations

(2) A test for the existence of effects.

The simulation results demonstrate that both the test for comparability (absence of self-selection bias) and the test for effects should be quite weak. During this phase, this corresponds with intuition; since subsequent phases of the design have greater power of discrimination, it is only desirable to stop the investigation at this stage if the probability of subsequent findings is very low. Nevertheless, even with relatively weak tests, there appears to be a significant chance of terminating at the end of the pilot study.[1]

In addition, several secondary objectives will be realized in the pilot study. These include analysis of the influence of independent family variables, screening of outcome measures, procedural improvements, estimates of test variances for use in later sample size selection, and the detection of problems not anticipated in this design.

The objectives of the pilot study are most effectively realized by contrasting programs which are polar extremes in terms of the independent program variables. Thus, it is desirable to contrast programs which are located in

---

[1]See Table 13. Simulation Results.

high income neighborhoods and which have a high proportion of moderate income residents with low-income programs in low-income neighborhoods. At the same time, it is desirable that these programs be as well matched as is possible on the excluded program variables--or there is no assurance that the variations are attributable to the kinds of environmental change on which the investigation is focussed. Specific recommendations for Phase I program selection are given in Volume II, Housing Program Survey.

3.3.3 Design and Analysis in Phase I

The pilot design permits the investigation of program group differences, the influence of the independent family variables and a check on the extent of self-selection between programs. It is useful to review these components of the investigation first of all in terms of the analysis of variance, although other techniques of analysis will also be employed.

Analysis of Variance

The pilot design is essentially a 2^4 factorial design. The four 2-way classifications are: program residence, family income, race of family, and program group; groups will be referred to as Group 1 (High/Moderate Neighborhood SES, Moderate Program SES) and Group 2 (Low Neighborhood SES, Low Program SES). Data from individual programs are pooled within groups. The cells in the pilot design shown in Table 2 are occupied by measured levels of development of children within that category. Table 3 can be interpreted as the design for a single measure on children of one age grouping.

Although the pilot study can be laid out as a 2^4 factorial design, it will be principally analyzed in three dimensions. The elimination of dimensions is dictated by the particular hypotheses of interest. They are now dealt with in turn.

Program Residence		No		Yes	
Program Group		Group I	Group II	Group I	Group II
Mod.	B				
	W				
Low	B				
	W				
Family Income	Race of Family				

Table 3

2^4 Full Factorial Design for Pilot Study

Hypothesis 1.1. For non-residents there are no significant differences between program groups.

This is the test of homogeneity or absence of self-selection amongst the entering populations. If there are significant differences between the populations at entry, then subsequent differences between children in different programs cannot be causally attributed to the treatments.

Hypothesis 1.2. Amongst resident children there are no significant differences between program groups.

If Hypothesis 1.1 is accepted (i.e., no self-selection), then rejection of this second hypothesis is equivalent to confirmation of the environmental change hypothesis; since the populations are homogenous at entry and non-homogenous later on, the treatments must have some effect.

Hypotheses 1.3 a&b. Group 1 programs have no significant impact on residents; Group 2 programs have no significant impact on their residents.

If Hypothesis 1.1 is accepted (i.e., homogeneity of entering populations) and Hypothesis 1.2 is rejected (i.e., there are significant differences between treatment groups after 18 months), then at least one of this third pair of hypotheses will be rejected. On the other hand, if Hypothesis 1.1 is rejected, 1.3 will still be of interest. It tells us that, even if treatment groups are not comparable because of self-selection, there are (are not) significant gains for the separate populations. This is done by comparing, within program groups, residents and non-residents. The difference, when age is controlled for, is the imputed developmental gain.

Hypothesis 1.4. Family income and race of family have no significant effects on child development.

Hypothesis 1.5. Race and income variables do not interact significantly with program residence.

This final hypothesis goes beyond a simple test for the main effects of family income and race to see whether program residence interacts significantly with the race and income variables.

Automatic Interaction Detection

The basic method of analysis applied in the pilot design is the analysis of variance. It is recommended that the data be further analyzed through automatic interaction detection applied to the independent and nuisance variables used in this design. The automatic interaction detector is basically an analysis of variance procedure, designed to analyze the importance of interactions among up to 36 variables in predicting a single dependent variable.[1] It is a computer program that employs a nonsymmetrical branching process to subdivide a given sample into a series of consecutive subgroups in such a way as to maximize, at each consecutive branching of the process, the ability to predict the dependent variable. The technique differs from conventional multiple regression techniques in that linearity and additivity assumptions are not required.

The overall strategy of the design is to find which of the several independent variables are most important in "accounting for" the variation in the dependent variable. For a given sample the total sum of squares is partitioned in various ways until that partition is found which maximizes the resulting between group sum of

---

[1]For a complete description of this technique, see The Detection of Interaction Effects: A Report on a Computer Program for the Selection of Optimal Combinations of Explanatory Variables, John Sonquist and James N. Morgan, Survey Research Center, Institute for Social Research, University of Michigan, Monograph #35, 1964.

squares. The partition is used to separate the sample into two subgroups, each of which can then be subjected to further treatment.

The proposed subdivision of the pilot design variables are as follows:

Program Residence
18-24 months
24-30 months
30-36 months

Program Size (units)
0-50
50-100
100 and over

Program Unit Type
apartments
townhouses

Program Income
(% moderate income)
75-100
50-74
25-49
0-24

Neighborhood Income
low
moderate
high

Family Income
very low
low
moderate
moderate-high

Family Race
white
black
other

Family Structure
1-parent
2-parent

Family Size
one child
two children
three children
four or more children

Sex of Child
male
female

Occupation of Family Head
(Parent Interview)

Education Level of Parents
(Parent Interview)

One of the interesting features of this technique is that the criterion for a split is one of importance or explanatory power rather than statistical significance. It is based on the notion that although correlation or association may not be sufficient to show causation, it _is_ necessary. The procedure is decidedly "post hoc," however, and as such is relatively useless as a means of testing _a priori_ hypotheses about the relationships obtained. The importance of a variable as measured by the proportion of variance it accounts for in the dependent variable is itself subject to random fluctuations. What turns out to be an important factor in one experiment may be shadowed in further experimentation by some other variable.

On the other hand, the technique can be a valuable heuristic in a research strategy that permits interaction between hypothesis formation and data collection. The initial conceptual scheme determines the variables that are included in the research in the first place, but their order of importance as suggested by the outcome of the analysis suggests hypotheses of a more precise model than the original conceptual scheme from which hypotheses may be deduced and then tested against additional data.

3.3.4 Sample Sizes in the Pilot Design

The size of the samples collected in the pilot investigation is constrained by the number of children in the programs selected; in particular, the relative scarcity of moderate income programs effectively restricts the number of children to be included in the sample. Since there are more applications than units, the number of families with residence over 18 months will be the binding constraint. By pooling programs in this category it is possible to find an experimentally balanced sample of 108 children. Of these, 24 should be of high-school age, 36 in the lowest age group, and 48 in the intermediate group.

These children constitute only one fourth of the total number of subjects included in the pilot study. Additional balanced samples will be drawn from the Group 1 non-resident and from Group 2 residents and non-residents giving a total sample size of 432; these will be divided between age groups as follows: pre-school - 144; elementary school - 192; high school - 96. It is necessary to consider the adequacy of each of these samples independently.

The adequacy of sample sizes can only be defined in terms of the ability of the analysis to discriminate between alternative hypotheses; in defining what is meant by this power of discrimination, it is necessary to distinguish between the null hypothesis (no effects) and the alternative hypothesis (effects). Unfortunately these two hypotheses are not, as stated here, strictly comparable. This null hypothesis that the effect is zero is a point hypothesis while the alternative hypothesis, that the effect is a non-zero, is composite in the sense that it only restricts the hypothesized level of the effect to a range of values. To avoid this difficulty, it is necessary to define the alternative as a point hypothesis specifying precisely the size of the effects obtained; when the alternative is specified in this way, it is then possible to define the power to discriminate between hypotheses in terms of the probability of accepting the alternative hypothesis when it is true. This is called the "power" of the test.

Within each age group, there is a set of observations to be made on each child; these observations are drawn from populations with different (and generally unknown) means and variances, and numerical specification of the alternative hypotheses in each case is clearly an arbitrary and time-consuming undertaking. To avoid carrying out separate calculations for each test and each age group, it is sensible to define scale-free alternatives which define the hypothesis in terms of percentiles of the standard deviation of the measure. Inevitably there is an arbitrary aspect to this, but it nevertheless constitutes a viable and useful means for reviewing the adequacy of the sample size.

For the purposes of this discussion, separate computations were made for two alternative hypotheses:

(1) The effect is greater than or equal to half a standard deviation of the measure in the subject population

(2) The effect is greater than or equal to one standard deviation of the measure in the subject population.[1]

Testing the null hypothesis at the 5% level, the power of the test for the levels of replication proposed for each age group in a 2^3 factorial design are as shown in Table 4.

It is not surprising to find that the test has only limited power when the alternative is less than one standard deviation away from the null hypothesis. This suggests that for purposes of making decisions about how to proceed after the data from the pilot study have been analyzed, use of a lower significance level on the null hypothesis may be indicated. This problem is addressed in the simulation discussion of Section 3.6.

---

[1]These alternatives represent non-trivial changes in measured levels of development. For a test such as the Wechsler which has a mean of 100 and a standard deviation of 20, the first alternative indicates a gain of at least ten points.

Size of treatment effect under alternative hypothesis	Number of replicates per cell		
	6	9	12
0.5σ	0.40	0.55	0.68
1.0σ	0.93	0.99	1.00

Table 4

Power of 95% level F Test in 2^3 Factorial Design[1]

[1] See Tang, P.C., "The power function of the analysis of variance tests with tables and illustrations of their use." Stat. Res. Mem., 2, pp. 126-149, 1938.

3.3.5 Costs and Scheduling

The pilot design calls for tests to be administered to 432 children at six locations. These children will be drawn from a larger population which must be surveyed for purpose of sample selection. The estimated costs of the pilot investigation, including survey and test work, data analysis, report preparation, and project administration are as follows:

Estimated Costs of Phase I (Pilot) Investigation

1. Direct Labor				
Project Direction & Data Analysis				
1 Senior Principal Scientist	6mm @	$2,500	$15,000	
1 Principal Scientist	6mm @	1,500	9,000	
1 Clerical	6mm @	750	4,500	
Survey and Testing Families 432 @ $80 (includes coordination, training, coding, but no payment to family)			34,560	
TOTAL DIRECT LABOR				$63,060
2. Overhead at 100% of Direct Labor[1]				63,060
3. Travel				
36 trips @ $150 average			5,400	
240 per diems per program @ $25			6,000	
TOTAL TRAVEL				11,400

[1] Overhead charges might be significantly reduced if interviewer salaries are not included in the direct labor base. Policy with respect to this varies between contractors.

4. <u>Computer Rental</u>

10 hours at $425	$4,250	
TOTAL DIRECT COST		$141,770

5. <u>General and Administrative Expense</u>

At 15% of Total Direct Cost		21,265
TOTAL COST		$163,035

The pilot study can be conducted on a relatively tight schedule if the project is well administered. The average number of children per program is less than 90, and with 4 interviewers and 2 testers, this can be completed in less than 25 working dyas. Allowing for start-up, on-the-site training, and unforeseen circumstances, each program could be completed within a period of two calendar months. Allowing a further two months for data analysis and report preparation, the schedule of activities is outlined in Table 5.

The Phase I (pilot) study will require 6 months to complete and will cost in the neighborhood of $160,000. The experience gained in the conduct of this study will, together with the data themselves, provide a necessary input for the Phase II effort.

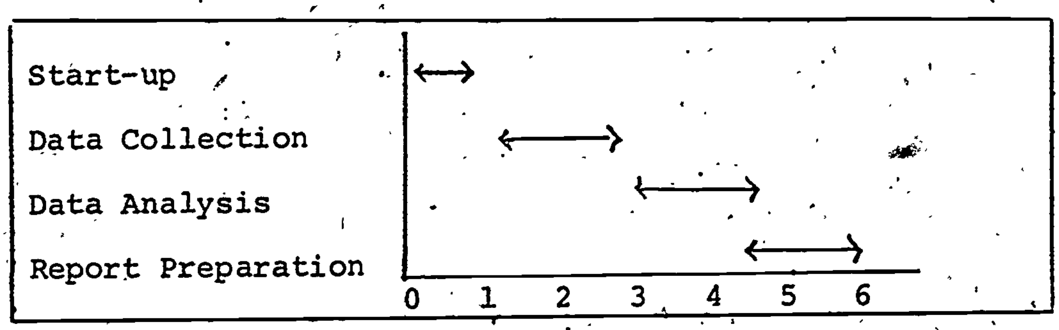


Table 5

Schedule of Activities--Phase I (Pilot) Investigation

3.4 Non-Experimental Design Phase II (Cross-Sectional)

3.4.1 Objectives of the Phase II Study

The objectives of the second phase of the non-experimental design are the "explanation" of significant effects observed in the pilot study and the assembly of baseline data for the third phase. If there are no differences, then either the hypothesis of environmental change is false or the length of time elapsed since entry into the program is insufficient for these effects to become apparent. The existence of differences does not, however, necessarily lead into the second phase. If the entering populations are not homogeneous in the sense that there are significant variations in the developmental levels of children entering different programs, then the observed treatment effects cannot be causally attributed to variations in the characteristics of the sampled programs. Necessary preconditions for the second phase investigation are the acceptance of Hypothesis 1.1 of the Pilot Study and the rejection of Hypothesis 1.2. If only the first of these conditions holds, it may be necessary to go directly to the longitudinal investigation.

In stating the objectives of the second phase study, it will be assumed that the above preconditions are satisfied; the homogeneity of the entering populations insures that the programs are comparable and the heterogeneity of the residents insures that there are some differences to "explain." The explanation of these differences will involve comparison of programs along two dimensions--the income mix of the program neighborhood and the income mix of the program itself. The principal objectives of the second phase study are therefore:

(1) Test of the hypothesis that the income mix of the neighborhood has an effect on the outcome measures

(2) Test of the hypothesis that the income mix of the program itself has an effect on the outcome measures.

Secondary objectives of the second phase study are:

(1) Analysis of the effect of the independent family variables

(2) Additional testing for self-selection amongst program applicants.

These objectives will be realized by an extension of the pilot study to include more programs, and by the elimination of those outcome measures will show little change in the pilot study. Procedural changes in the second phase should reflect the experience gained in the first place.

3.4.2 Selection of Programs for the Phase II Study

The programs required for the second phase reflect the need to achieve contrasts between programs along both of the dimensions of interest. The difference between the first and second phase studies in their treatment of programs corresponds to a change from a binary classification of programs to a fourfold classification. This is shown in Table 6.

This in no way prevents the subsequent analysis of data either through analysis of variance based on a 3-way classification of neighborhood SES or, as recommended in the Phase I study design, through use of the automatic interaction detector. By classifying neighborhoods 3 ways, a 3x2 design can be developed as shown in Table 7.

		Neighborhood Income	
		High/Moderate	Low
Program Income	Moderate		
	Low		

Table 6

Program Variables in Phase II Study

		Neighborhood Income		
		High	Moderate	Low
Program Income	Mod.			
	Low			

Table 7

Alternative 3x2 Form for 2nd Phase

The 3x2 design has many features to recommend it, particularly when the study is viewed as an investigation of acculturation effects; it may, however, be difficult to implement without the inclusion of more programs. For this reason, the remainder of the discussion will focus on the basic 2x2 program classification.

3.4.3 Selection of Outcome Measures in Phase II

The selection of outcome measures in Phase II is entirely dependent on the results of the pilot study. Measures which show very little evidence of effects can be eliminated from the test battery, and in this way an opportunity is provided for either extension of core measures which do show some effect, or for reduction in the length (and administrative cost) of the test battery. It is tempting to suggest formal rules for this process of elimination and substitution, but reflection on the nature of such rules suggests that they would be very complex and ultimately a poor substitute for good judgment. It should be noted, however, that where extended measures are introduced, it will be necessary to carry out additional measurements at the Phase I sites. This additional cost is not incurred if no extensions are made. With this reservation, the costs of survey and measurement will be assumed to be the same as in the Phase I study.

3.4.4 Design and Analysis in Phase II

The methods of analysis employed in Phase II are essentially the same as those of the pilot study. Instead of a 2^3 factorial design, however, it is now a 2^5 design with two independent program classifications. It is now reviewed in the framework of the analysis of variance.

It is convenient to being by considering the design as applying to a single measure on children of one age grouping. The associated 2^5 factorial layout is shown in Table 8. The pilot study, although laid out as a 2^3 factorial design, was principally analyzed in only three dimensions. Similar considerations lead to the reduction of the second phase study to four dimensions for most of the analysis. The choice of dimensions to be eliminated reflects the nature of the particular hypotheses of interest.

Hypotheses 2.1-2.2. There are no significant differences in the development levels of non-resident children families with respect to the program variables.

This is the test of homogeneity or absence of self-selection amongst the entering population. If the test was passed in Phase I, this represents a reconfirmation of the validity of inference in the second phase. These tests--one for each of the program classifications--are the conventional tests for main effects in a 2^4 factorial design.

Hypotheses 2.3-2.4. There are no significant differences in the development of resident children with respect to the program variables.

As in the pilot study, if Hypothesis 2.1 is accepted (i.e., no self-selection on the first program variable), then rejection of 2.3 is equivalent to confirmation of the hypothesis of environmental change.

Hypotheses 2.5-2.8. Within each of the 2^2 program classifications, there are significant gains in child development.

If there are significant self-selection effects discernible when the developmental levels of program entrants are compared, it will be impossible to make useful comparisons between programs. Nevertheless, it may still be possible to identify gains due to environmental change within each of the eight program groups. This is done by comparing residents with entrants for each group.

Program residence		No				Yes			
Family income		Moderate		Low		Moderate		Low	
Race of family		White	Black	White	Black	White	Black	White	Black
Low	Low								
	Mod.								
High/Mod.	Low								
	Mod.								
Nbhd. Income	Program Income								

Table 8

2^5 Fully Replicated Factorial Design for Phase II Study

Hypotheses 2.9-2.10. Family income and race have no significant effects on child development.

Both these hypotheses will be tested in the full 2^5 factorial design.

Hypotheses 2.11-2.12. Family race and income variables do not interact significantly with length of residence.

This is the compensatory hypothesis outlined in the discussion of analysis in the pilot study. Significant two-way interaction between these variables would suggest that gains are not equally distributed between program entrants or different income and/or racial background.

As in the Phase I study, the analysis should not be restricted to the analysis of variance alone. Partitioning of independent and nuisance variables for use with interaction detection methods will be used to uncover significant inter-relationships not included in the basic hypotheses.

3.4.5 Sample Sizes in Phase II

The size of the samples collected in Phase II is constrained by the number of children resident in the programs selected; again the relatively small size of the high income programs is the binding constraint. The pilot investigation called for a total of 108 children to be drawn from families resident in these three programs. Since these represent only one-eighth of the cell in the design, the total sample size will be 864, or exactly double the size of the Phase I sample. These will be distributed amongst age groups as follows: pre-school - 288; elementary school - 284; high school - 192. The adequacy of these sample sizes is now considered in terms of their ability to discriminate between alternative hypotheses.

Number of Replicates Per Cell

Size of treatment effect under alternative hypotheses		6	9	12
	0.50σ	0.67	0.83	0.91
	1.00σ	0.99	1.00	1.00

Table 9

Power of Test in 2^4 Factorial Design[1]

---

[1] Computed from Table of Tang, P.C., op. cit.

In the context of the proposed 2^5 factorial design, there will again be 9 replicates per cell in the pre-school group. The power of the test of the null hypothesis at the 5% level is shown in Table 9 for each level of replication in a 2^4 factorial design. This is done for two levels of the alternative hypotheses.

The Phase II design has considerably more power with respect to the alternative hypothesis than the Phase I design. This reflects the increase in the number of factors which more than offsets the reduction in the number of replicates.

3.4.6 Costs and Scheduling

The Phase II design requires the administration of tests to 864 children. Of these, 432 children at six locations will already have been surveyed in the pilot investigation. The incremental costs of the Phase II investigation reflect the additional survey requirements, together with a similar budget for project direction and data analysis.

Estimated Incremental Costs of Phase II Investigation

1. Direct Labor

Project Direction and Data Analysis				
1 Senior Principal Scientist	6mm	@	$2,500	$15,000
1 Principal Scientist	6mm	@	1,500	9,000
1 Clerical	6mm	@	750	4,500
Survey and Testing				
432 children (includes coordination, training, coding, but no payment to family)		@	80	34,560
TOTAL DIRECT LABOR				$63,060

2. Overhead at 100% of Direct Labor			63,060
3. Travel			
36 trips @ $150 average	5,400		
240 per diems @ $25	6,000		
TOTAL TRAVEL			11,400
4. Computer Rental			
Includes re-analysis of Phase I data 20 hours @ $425			7,020
TOTAL DIRECT COST		$147,020	
5. General and Administrative Expense			
at 15% of Total Direct Cost		22,053	
TOTAL COST			$169,073

The estimated total cost of carrying the study through both Phases is then:

Phase I	$163,035
Phase II	169,073
TOTAL PHASE I AND PHASE II	$332,108[1]

[1] If payments to families are included, this would increase the cost by $8,640 for $10 payments and by $17,380 for $20 payments.

The second phase study can be carried out on a schedule which is essentially the same as that of the first phase. Allowing for a period of three months between first and second phases in which to circulate and review the findings, both phases can be completed within a period of 15 months at a total cost of less than $350,000, not including payment of fees.

3.5 Non-Experimental Design Phase III (Longitudinal)

3.5.1 Objectives of the Phase III Study

The second phase of the non-experimental design differed from the first in that it enlarged the number of hypotheses which could be investigated. The differences between the second and third phases are of a different nature. Instead of enlarging the design to include more programs and more program dimensions, the third phase extends the second over time to permit a more accurate and powerful investigation of the findings to date. There are two principal reasons why the longitudinal investigation of Phase III can lead to much more convincing findings. The first relates to the within cell variance of the design. The second concerns the length of time over which treatment effects are conjectured to take place. Those are now reviewed in turn.

In the cross-sectional analyses of both the first and second phases of the design, the within cell variances correspond to the variances across the population of interest. Thus, in a simplified model, the scores of resident six-year olds are contrasted with the scores of non-resident six-year olds. In the longitudinal design, however, effects are estimated by comparing the <u>gains</u> of matched groups of six-year olds over a specified period of time, and the within cell variances are the test-retest variances rather than the population variances. This reduction in within-cell variation permits a considerable increase in the power of the tests used in the basic analysis of variance design.

The second advantage of the longitudinal design depends upon judgmental rather than statistical considerations. The programs included in the second phase of the study are comparatively new, and few of the better programs being more than three years old. This raises the question of whether or not 18-36 months is a sufficient length of time to allow for discernible effects on children to take place. If it is not felt to be long enough, and if the second phase study provides only limited support for the hypothesis of environmental effects, then the longitudinal study becomes critical to the success of the non-experimental investigation. If, on the other hand, the second phase provides considerable support for the basic hypotheses, the corroborative aspects of the longitudinal study are of considerable value.

The objectives of the longitudinal study can be stated in terms of the reinforcing or corroborative value of the analysis it permits. It also brings with it certain problems, such as attrition, which are not present in the cross-sectional designs. These are dealt with in the sections that follow.

3.5.2 <u>Design and Analysis in Phase III</u>

The basic method of analysis applied in the third phase of the non-experimental design remains the analysis of variance. This is applied to comparisons of residents in different programs who have been exposed to the treatment of interest for a similar length of time. Comparisons are made on the basis of measured gains over the period intervening between the Phase II measurements and a second wave of measurements (Phase III). There are, as will become apparent, problems in controlling for the number of observations in each cell, because of mobility in the subject populations. These problems will be deferred momentarily in order to explore the effects of controlling for individual differences in a two-wave longitudinal study.

In the cross-sectional studies, the basic test for homogeneity in the entering populations is equivalent to a test that the children and families of program applicants are all drawn from the same population, regardless of the program to which the family has applied. Gains are then estimated implicitly as the differences between the measured development of residents and non-residents when all other variables have been controlled for.

Maintaining the length of residence variable constant, and assuring the variance in scores to be equal for both populations, it is possible to compare the within cell variance of the cross-sectional and longitudinal designs. Let y_{ik} denote the measured development of the k^{th} child in the i^{th} time period (age is controlled for). Assume two time periods (i=0.1); let the (common) population variance within each time period be denoted by σ^2 and let the test-retest correlation coefficient be denoted by ρ. Now compare the variances of gains estimated by (1) differences in measured development between the same individual at different times, and (2) differences in measured development between different individuals at different points in time:

$$\sigma^2 (y_{1k}-y_{2j}) = 2\sigma^2(1-\rho) \qquad j=k$$
$$= 2\sigma^2 \qquad j\neq k$$

The ratio of these two, $(1-\rho)$, gives the relative efficiency of the two estimates of the cell mean. Thus, for a test-retest correlation of +0.8, the within cell variance is reduced by 80% and the standard deviation by almost 60%. This dramatically increases the power of the test with respect to alternatives specified in terms of population standard deviation.

As the length of time between separate survey waves increases, the correlation between test and retest will characteristically decline and in this way there will be some loss in efficiency. There will, however, be a more powerful influence working in the other direction, since it is reasonable to suppose that the effects of environmental change will be greater over a longer time period. Since the effects, if there are any, will probably increase more consistently than the test-retest correlation will decline, the net effect of these two influences will fall, other things being equal, on the side of a longer run study.

Unfortunately, things are not entirely equal for two reasons. In the first place, "gains" can only be estimated if the measurements on individuals are carried out on the same scale or if overlapping tests are used to consolidate broader age ranges. If overlapping tests are not available for all dimensions of interest, and if the limits of applicability are exceeded, then programs can only be compared on the basis of raw scores for matched groups in different programs. What remains is essentially a cross-sectional study which cannot be done for several years until the children get older. These considerations strongly suggest that the second wave should be carried out within 2-3 years of the first wave.

The second constraint on the length of time between survey waves is the problem of attrition. Attrition interferes with the design and analysis in two ways: one of these problems is present, although heavily concealed, in the cross-sectional design, and both are present in the longitudinal design.

The first difficulty created by attrition is the problem of self-selection over time. If child development is correlated with mobility--with the more (or less) advanced children being more (or less) likely to move either out of (or into) the programs, then there is a bias in estimates of gains based on comparison of residents and non-residents. This is a source of bias present in both the cross-sectional and longitudinal designs. The second difficulty is created even if mobility is completely uncorrelated with the outcome measures, because attrition will lead to smaller and more unbalanced samples, with consequent loss of power to discriminate. This is a problem which belongs only to the longitudinal design.

The second problem is more serious and can be compared exactly to the problem of family self-assignment to programs. Where families entering different programs differ with respect to variables important to child development, it is not possible to attribute observed differences to the effects of the program. Similarly, when the families which exit programs are not comparable across programs, it may not be possible to attribute differences in the outcome variables to the effect of the program working itself out over time; rather, it may simply reflect high turnover amongst a non-random subset of the sample.

This is a potential problem in the cross-sectional as well as in the longitudinal design. Since gains are estimated by comparing residents with non-residents, self-selection across time may undermine cross-sectional inference. In the longitudinal design, although the intervening time period is longer and the potential self-selection problem is correspondingly more serious, there is a simple and direct means of testing for the presence of self-selective effects. This method provides a method of validating the findings of both longitudinal and cross-sectional analyses.

It is useful to begin by considering the baseline data provided by the Phase II investigation as divided into two groups--residents and non-residents. At the time of the second wave of measurements, the original resident sample will consist of families who are still resident and families who have moved. Reanalysis of the baseline data with mobility introduced as a binary factor can then serve as a test of inter-temporal self-selection. If significant differences in the development levels of those who moved vs. those who remained are found, this will have the effect of compromising the findings of the study. If they are not found, this will serve to validate the findings of all three phases of the experiment. This exercise, which requires no further measurements to be undertaken, is therefore recommended whether or not the full Phase III design is implemented.

The non-resident group will itself have a pattern of mobility, and this is a more complicated problem in some respects. Program applicants included in the baseline (Phase II) survey may either, by the time of the second wave, have entered the program, have not entered the program and remained in their original location, or have not entered the program, but moved elsewhere (including possibly another housing program). In each of these cases, the role of the baseline applicant group in providing necessary contrasts requires special treatment. The alternative approaches to this problem are now reviewed.

It should be remembered that the principal objective of the Phase III design is to compare gains in child development across programs. In the Phase II analysis, the role of the applicant group was twofold: (1) to provide a test of homogeneity in the entering populations, and (2) to provide a means for estimating gains on a program by program basis if the test of homogeneity is rejected. Since the Phase III design should not be implemented if the test for homogeneity is rejected, and since the Phase III design contributes no new insights into the question of homogeneity amongst entrants, it is not clear why further measurements on applicants should be undertaken. Those applicants who do enter the program in the intervening period will differ from the original resident group with respect to length of residence, and therefore the two groups cannot be pooled in any convenient way. It is therefore sensible to eliminate the original non-resident group from the second survey wave.

The effect of the elimination of the non-resident group from the longitudinal design is to limit the hypotheses of interest to questions of relative gains in different housing environments. It is no longer possible, therefore, to make judgments with respect to absolute gains, in the sense of a comparison between the development levels of children in a given housing program with what they would have been if the family had not entered the program. To do this it would be necessary to resurvey all applicants, whether or not they moved in the interim period. Not only would this be an extremely costly and doubtful undertaking, but it is also not clear why better alternative housing opportunities in certain metropolitan markets should be reflected in reductions in the estimated impact of good housing programs. Nor can this be avoided by retesting only those applicant families who fail to move, since this merely introduces a further element of self-selection.

To summarize the basic strategy of the Phase III design in terms of the collection of data, it is useful to identify three defining choices:

(1) The Phase III design involves a second wave survey of the Phase II programs

(2) The second wave should be carried out within 2-3 years of the first wave

(3) The second wave should only include original program residents.

Three basic types of analysis should be carried out in the third phase of this study. There are reviewed in turn:

<u>Hypothesis 3.1. There are significant differences in baseline levels of development between children of the original resident group who remain in the program and children who are no longer in the program.</u>

This test involves reanalysis of the Phase II data on two-year residents. Substitution of "moved" vs. "not moved" for program residents in the Phase II design permits a test for this main effect in the framework of a 2^5 factorial design. Classification of families according to mobility is done on the basis of data collected on the second wave. The unbalanced nature of the design makes the test statistics somewhat more complicated.

Hypotheses 3.2-3.3. There are differences in mean development gains along the dimensions of neighborhood income and program income.

The tests will be conducted within the analysis of variance framework on the original resident group only. The dependent variable in all cases is the measured gain over the period. These are the tests of the basic environmental hypotheses.

Hypotheses 3.4-3.5. There are differences in the mean gains for families of different race and income.

These again are conventional tests of the impact of family variables, but this time they test for gains rather than absolute levels. The compensatory hypotheses which were originally investigated through two-way interactions between family variables and length of residence are not tested as main effects. Family income will be defined as the family's income at the time of entry to the program.

3.5.3 Sample Sizes in Phase III

The size of the populations sampled in Phase III of the design is essentially determined by the Phase II design. Given that only residents will be tested in the final phase, the number of children to be sampled is given by the size of the original resident group less losses from the original resident group. This would place the maximum sample size in the final phase at 432 children, distributed as follows: pre-school - 192; elementary school - 256; and high school - 128. The actual number will be significantly less by reason of attrition.

The power of the tests of main effects within the 2^4 factorial design is shown in Table 10 for each level of replication and assuming zero attrition. Alternative hypotheses are again stated in terms of the population variance, but the power reflects the within cell variance of gains. This is equal to $(1-\rho)$ times the population variance, where ρ denotes the test-retest correlation coefficient. This is assumed to be 0.75 for the purposes of estimating power.

The inevitable lack of balance in the design when mobility has taken its toll may reduce the power of these tests to some degree. Nevertheless, it is good to see that these tests have very high power to discriminate between the null and alternative hypotheses.

3.5.4 Costs and Scheduling

The final phase of the design resembles the first and second phases in the sense that the number of target children remains the same, although mobility and the previous sample selection will reduce the total number of families to be surveyed. The final phase does require more sites to be visited, since all the Phase I and II programs are included.

Size of treatment effect under alternative hypotheses	Number of Replicates per Cell		
	6	9	12
0.25	0.67	0.83	0.91
0.50	0.99	1.00	1.00

Table 10

Power of Test in 2^4 Factorial Design

Estimated Incremental Costs of Phase III Investigation

(1972 dollars)

1. Direct Labor

Project Direction and Data Analysis

1 Senior Principal Scientist	6 mm @ $2,500	$15,000		
1 Principal Scientist	6 mm @ 1,500	9,000		
1 Clerical	6 mm @ 750	4,500		
Survey and Testing				
300 children @ $80[1] (includes coordination, training, coding but no payment to family)		24,000		
TOTAL DIRECT LABOR				$52,000

2. Overhead @ 100% of Direct Labor				52,000

3. Travel

72 trips @ $150 average	$10,800			
360 per diems @ $25	9,000			
TOTAL TRAVEL		$ 19,800		

4. Computer Rental

20 hours @ $425		9,500		
TOTAL DIRECT COST			$133,300	

5. General and Administrative Expense

@15% of Total Direct Cost			20,095	
TOTAL COST				$153,395

---

[1] No adjustment has been made for the elimination of the short parent interview. This will be somewhat offset by the increased percentage of coordination time as fewer children are included on more sites.

The estimated cost of carrying the study through all three phases is then:

Phase I	$163,035
Phase II	169,073
Phase III	153,395
Total	$485,503

The second phase study can be carried out within a six-month period. It should be initiated within 2 and 2 1/2 years of the Phase II data collection effort.

3.6 Simulation of the Non-Experimental Design

In presenting the non-experimental design described in this section, attention was given to the possible presence of self-selection biases and to the need for a sequential investigation which could be terminated if intermediate results prove disappointing. These are worthy sentiments, but they are of little value as long as the effect of self-selection bias remains obscure and as long as no well-defined rules exist for discontinuing or redirecting the investigation on the basis of data collected during the first and second phases. To meet these needs, computer simulations of the entire design have been performed, and the results of these simulation runs have been used to determine both the robustness of the design with respect to potential impurities and the rules for the conduct of the three-phase investigation. The simulation described here provides not only important insights into the operation of the design, but as a by-product working computer programs for the analysis of the real data as it becomes available.

The remainder of this section is devoted to a description of the way in which simulated data are generated and analyzed, together with the results of the analysis. In the discussion of the generation of simulated data, emphasis is placed upon the methods used to simulate environmental effects and self-selection biases both between programs and across time. In the discussion of analysis, emphasis is placed upon the (positive) payoff of finding environmental effects when they really exist and the (negative) payoffs of finding effects when none exist. These payoffs, together with the estimated cost of each phase of the investigation, permit optimization of the design with respect to the critical values of the test statistics. These values provide the rules for the conduct of the investigation; they constitute an integral part of the design proposed.

3.6.1 Simulation Design

Simulation of the three-phase investigation consists of the generation of test scores for each level of replication proposed; these test scores are assigned to combinations of factors within a 2^5 factorial design. Scores are then modified to reflect inter-program differences between residents and non-residents attributable to the effect of residence (environmental impact) or to systematic differences between entering populations (self-selection type I) or to exiting populations (self-selection type II). The simulated scores aredrawn originally from a normal population with mean 100 and standard deviation of 20.[1]

Simulation of Environmental Impact

Environmental impact is characterized as the change in the mean scores of residents vs. non-residents attributable to program characteristics. In the cross-sectional phases of the investigation, this is achieved by simply adding an effect equal to 25% of the population standard deviation to the scores of residents in a selected program classification. In the longitudinal phase, environmental effects are slightly more complicated, involving modification of the mean of the distribution from which the second wave of observations is made. The subjects' second wave score then reflects both his first wave score, the environmental effect of the program, and a random element.

Simulation of Self-Selection Type I (Inter-program)

Inter-program self-selection is simulated in the same fashion as environmental effects, except that the changes are applied to residents and non-residents alike. The tests for self-selection type I are applied

---

[1]Twelve uniformly distributed random numbers are summed; when standardized, this sum approximates the normal distribution with zero mean and unit variance.

within the first and second phases of the investigation, and both involve analysis of the non-resident scores.

Simulation of Self-Selection Type II (Inter-temporal)

This is the most complicated aspect of the simulation experiment. Subjects' probability of moving in a two-year interval is correlated with his score. Thus the average probability of moving is modified either upwards or downwards according to the subjects' developmental level. By simulating mobility in this way, the test for self-selection of the second kind involves analysis of variance in a 2^4 factorial design with unequal cell frequencies.

3.6.2 Analysis of Simulation Experiments

The basic outputs of the simulation are sets of test statistics. These statistics will vary according to the assumptions made (i.e., environmental impact and no self-selection type I and no environmental impact) and to some extent randomly. The test statistics which are generated in each run of the simulation are all tests for main effects on program variables and on residence; tests of family income and race variables are not generated. The test statistics generated are as follows:

Phase I: F_{11} - Test of Self-Selection Type I

F_{12} - Test of Environmental Impact

Phase II: F_{21} - Test of Self-Selection Type I

F_{22} - Test of Environmental Impact

Phase III: F_{31} - Test of Self-Selection Type II

F_{32} - Test of Environmental Impact

For each level of replication (6, 9, and 12) there are eight possible combinations of environmental impact, self-selection type I and self-selection Type II, each of these being treated as binary variables. The test statistics are generated in the simulation for each of these eight possibilities.

For any given set of critical values corresponding to these test statistics, a given simulation will lead acceptance of some of the hypotheses and rejection of others. For a set of simulation runs, it is possible to estimate the probability (percentage of the time) that a particular hypothesis will be accepted or rejected. Thus, in the case where there are no environmental effects but self-selection of both kinds is present, it would be possible with 100 or 500 simulation runs to estimate the probability of concluding that environmental effects were present but there was no self-selection; these probabilities will vary with the critical values of the test statistics.

The simulation experiment now provides us with an instrument for the selection of critical values for this experiment. Tests carried out at the 5% level tend, as was demonstrated in the discussion of sample sizes in Phase I, to be extremely conservative. Although the null hypothesis is seldom rejected when true, it is frequently accepted when it is false. The danger which this creates is that of concluding that no environmental effects exist on the basis of a test which rejects the hypothesis of effects 40% of the time, even when they exist. To avoid this problem, and to take advantage of the sequential nature of the investigation proposed, the simulation has been used to generate rules for the conduct of the investigation.

3.6.2.1 Rules and Payoffs in the Three-Phase Study

It is reasonable to assume that the Government's sponsorship of this investigation reflects an interest in finding out whether or not environmental change has an effect on child development. If the investigation concludes that an effect exists, then the Government may initiate programs or policy to take advantage of this knowledge. If environmental change does, in fact, have an effect, these programs will be successful; if environmental change has no effect (conclusions were false) the programs will fail. In the first case positive net benefits will be realized; in the second case, money will be wasted. The investigation may, of course, conclude that no effects exist in which case no new programs will be initiated; under these circumstances, and regardless of whether or not effects really exist, there will be no benefits and no costs. These payoffs are summarized in Table 11. In this table, B denotes the benefits of programs designed to take advantage of environmental change effects, and C denotes their cost. It is reasonable to assume that B-C is positive.

It is now possible to express the expected value of the three-phase investigation in terms of these payoffs, and the expected cost of conducting the study (K). Denoting this expected value by V, we have

$$V = (B-C)\ \text{Prob (Effect found and no effect exists)}$$
$$-C\ \text{Prob (Effect found and no effect exists)}$$
$$-K$$

We now define what is meant by "finding an effect" in terms of the test statistics. Denoting with asterisks the critical values of the test statistics, an effect is said to be found if <u>all</u> the following inequalities are satisfied:

	Effect	No Effect
Effect Found	B-C	-C
No Effect Found	0	0

Table 11

Pay-off Matrix for Investigation

Phase I	$F_{11} \leq F_{11}^*$	No self-selection type I
	$F_{12} > F_{12}^*$	Effect exists
Phase II	$F_{21} \leq F_{21}^*$	No self-selection type I
	$F_{22} > F_{22}^*$	Effect exists
Phase III	$F_{31} \leq F_{31}^*$	No self-selection type II
	$F_{32} > F_{32}^*$	Effect exists

If any one of these inequalities is not satisfied, the investigation should be terminated; clearly there is no benefit to be achieved by continuing and a further phase of the investigation will only cost more money.

The rules of the investigation are now defined except for the choice of critical values for the test statistics. This choice is now performed by finding those critical values which maximize V. In carrying out this maximization, all eight possible combinations of effect and self-selection were assumed to be equally likely to obtain. The expected costs of the investigation itself (K) is computed as simply the cost of each phase times the probability of undertaking that phase.

3.6.2.2 Optimization Results

Separate optimization runs for 6, 9, and 12 replicates have been performed; these correspond respectively to the designed levels of replication for high-school, pre-school, and elementary school subjects. In each case, the benefits of finding an effect when one exists were assumed to be $20 million. The costs required to realize these benefits were assumed to be $10 million. At some additional expense, it would be possible to investigate the sensitivity of the results to alternative benefit and cost assumptions. As already mentioned, eight alternative combinations of effects and self-selection types, all assumed equally likely, were used to generate simulated tests statistics. In only 12.5% of the runs, therefore, were effects present without any form of impurity. In view of this, the results are extremely encouraing.

It is useful to consider the results as providing two kinds of information. In the first place, for each level of replication, optimal values of the test statistics in each phase are found. These are presented in Table 12. In the second place, for each level of replication, the power of the investigation to discriminate between alternative hypotheses and its probability of terminating at each phase are found. Those results are presented in Table 13.

It is interesting to note several aspects of these findings. In the first place, it is never proper to terminate the investigation for absence of effect in the first two phases of the study. This reflects the added power of discrimination which the longitudinal design provides. It is, however, possible to terminate the design because of self-selection effects. As the number of replicates increases, the critical values also increase; since

	Phase I		Phase II		Phase III	
Test	Effect	SS1	Effect	SS1	Effect	SS1
6 Replicates	.12	14.00	.25	5.75	4.25	7.125
9 Replicates	.19	15.50	.42	6.90	6.50	8.00
12 Replicates	.31	20.00	.56	9.40	7.25	9.00

Table 12
Optimal Critical Values of the Variance Ratio

this is true even for the self-selection effects, the power of the longitudinal phase must be quite sensitive to sample size.

The second way of looking at these results involves the probabilities that the experiment will find an effect given that one exists vs. the probability that it will find an effect when one does not exist. These probabilities together summarize the power of the design, in the expected presence of self-selection bias of both kinds, to discriminate between the primary hypotheses. The results, together with the probabilities of stopping at each phase, are presented in Table 13.

Level of Replication	Truth	Probability of finding an effect	Probability of Termination by Phase		
			III	II	I
6	Effect	0.83	0.11	0.01	0.05
	No Effect	0.10	0.69	0.04	0.17
9	Effect	0.88	0.07	0.00	0.04
	No Effect	0.05	0.75	0.01	0.19
12	Effect	0.92	0.05	0.00	0.03
	No Effect	0.04	0.79	0.01	0.16

Table 13

Discriminating Power of the Investigation and Probability of Termination at each Phase

CHAPTER FOUR

EXPERIMENTAL DESIGNS FOR HOUSING PROGRAM IMPACT

The non-experimental investigation proposed in Section 3 is vulnerable to impurities of two kind; in the first place there is the danger of self-selection or self-assignment of subjects both between programs and across time; in the second place there is the danger that the variables which characterize the housing treatments are not adequately controlled. If either or both of these impurities are present, significant treatment effects may reflect either differences between the subject populations, or the ommission of influential variables. In either case, the value of the investigation is severely compromised. Awareness of these pitfalls has prompted the investigation of truly experimental approaches to the design of this study.

In attempting to meet the requirements of an ideal experiment in the context of this study, certain difficulties arise immediately. In the first place, the random assignment of subjects to housing treatments requires a major committment of funds, however it is to be accomplished. On practical grounds alone, it is not reasonable to expect that short-term undertakings to subsidize families living in selected housing environments will be effective. In the second place, the housing treatments themselves can never provide the pure contrasts of the laboratory, since the controlled variation of the variables of interest can never be entirely independent of uncontrolled variation in others, such as schools, health care, local unemployment, and so on. Two experimental approaches to this investigation have been explored. Each of these approaches goes some way toward the removal of one--but not both of these difficulties.

The first alternative to be considered has significant advantages in terms of cost, but provides only limited insight into the dimensions of environmental impact. This is the housing allowance experiment which the U.S. Department of Housing and Urban Development is proposing to implement. The incremental cost of piggy-backing on to this experiment is quite small, but because of the way in which the project is designed, its value is somewhat peripheral to this study. The second alternative, which is the most powerful design of all those proposed, is also by far the most expensive. The only real way to achieve both planned combinations of housing conditions and random assignment of subjects is to build the programs and to manage admissions. By selecting locations to remove unintended local variations in unemployment, school quality, supportive services and so on, it will be possible to achieve much purer contrasts between programs than is possible with existing facilities. Support from agencies interested in other dimensions of environmental impact may defray some of the costs of this approach, but it will never be cheap.

These experimental approaches are seen as backups to the three phase non-experimental investigation outlined in the preceding section. As will become clear, a valid non-experimental investigation will provide more information than the Housing Allowance experiment at a cost which is roughly equivalent; it will provide almost as much a full-blown experiment at a considerably reduced cost. Only if the non-experimental designs fail to meet the necessary criteria of validity, should they be abandoned in favor of the designs which follow.

4.1 The Housing Allowance Experiment

The U.S. Department of Housing and Urban Development is currently planning to conduct a major housing allowance experiment; present plans call for the experiment to be conducted on nine sites with 600 households on each site selected to receive payments and an additional 400 households per site to serve as controls.[1] Payments to families will be geared to an experimental plan involving both family income and the rental payment. The principal objective of the experiment is to relate patterns of housing mobility to the structure of treatments called for in the design.

This experiment clearly provides an opportunity to investigate the relationships between housing and locational change on the one hand, and measures of child development on the other. It is important to emphasize, however, that the inclusion of child development measures in the context of this experiment will lead to hypotheses which are qualitatively different from those which the Non-Experimental Approach is designed to test. The Housing Allowance Experiment differs from the Non-Experimental Approach in that its emphasis is primarily upon locational change within both public and private sectors of the housing market rather than upon relocation into public housing. Other things being equal, this makes it generally more interesting than designs developed exclusively for public housing programs. Other things are not equal, however, and as will become clear, the Housing Allowance Experiment is not a particularly promising vehicle for this investigation. To understand why this is so, it is necessary to examine the design of this experiment in some detail.

The housing allowance experimental design calls for the experimental group to be assigned to ten basic treatments. These treatments involve different payment formulae, cost standards, contribution rates, and earmarking methods. There is, in addition, a control

---

[1]Not all these families will have children, naturally.

group which will receive no allowance. Special provision of information on housing opportunities may be provided to some, but not all, of the experimental group.

The basic objectives of the housing experiment cannot be realized unless there is some variation in the response of the different groups to alternative levels and combinations of allowance. To this extent there is a coincidence of interest between the primary objectives of this study and the secondary (or child development objectives). Here the coincidence ends, the objectives of the primary study relate to the response of individual housing choice to alternative combinations of incentives; the objectives of child development study on the other hand, relate to the response of child development to the induced housing change and not to the incentives which bring the change about. The outcome measures are, therefore, once removed from the actual treatments administered; intervening between treatment and outcome is the family's choice amongst the available housing alternatives; this is just the self-selective mechanism which the experimental approach is designed to avoid.

The implications of this are clear. It is not possible, despite the appearance of experimental assignment, to attribute observed developmental gains to the housing change of the family; within the treatment groups both housing change and child development will be jointly determined by the individual characteristics of the family; statistical association between the two cannot be given a causal interpretation. This problem would not arise if families were assigned directly into different types of housing, as in the second of the experimental designs proposed.

4.1.1 Design and Analysis in the Housing Allowance Experiment.

The self-selection problems of the Housing Allowance Experiment can be avoided if the attempt to associate developmental gains with the individual's housing choice is abandoned. Instead the analysis is directed toward the comparative impact of different allowance schemes on the child and family variables of interest. If significant differences between treatments are detected, it may then be possible to "explain" these differences by reference to the mean changes in housing and in disposable income for each treatment group.

It is useful to begin by considering the relationship between a selected child development variable and the independent variables of housing and disposable income. Denoting the measured gain of the i^{th} child assigned to the k^{th} treatment by Δx_{ik}, we can conjecture the existence of a linear relationship between this gain on the one hand, and changes in the family's housing (Δh_{ik}) and disposable income (Δy_{ik}) on the other:

$$\Delta x_{ik} = \alpha_i \Delta y_{ik} + \beta_i \Delta h_{ik} + \varepsilon_{ik} \qquad (4.1)$$

ε_{ik} denotes measurement error, and the gain may be defined either as relative to baseline measurements or to the control group mean. Notice that the coefficents of the independent variables are permitted to vary across individuals.

By taking expectations over individuals, equation (4.1) can be written:

$$\underset{i}{E}\{\Delta x_k\} = \alpha\Delta\bar{y}_k + \beta\Delta\bar{h}_k + \mathrm{Cov}\{\alpha_i \Delta y_{ik}\} + \mathrm{Cov}\{\beta_i \Delta h_{ik}\} \qquad (4.2)$$

α and β now denote the mean response of the dependent variable to changes in disposable income and housing respectively; they are the unknowns of interest.
y_k and h_k denote respectively, the average changes in disposable income and housing for the k^{th} treatment group.

This equation can only be estimated if the covariance terms are the same for each treatment group--or if their variation can be represented by some simple expression. The covariance term measures the degree of association between the responsiveness of the child development measure to housing choice and the responsiveness of housing choice to the experimental incentives provided. Thus if families who choose more housing improvement than the average of their group have children whose response to housing improvement is also above the average, the covariance term will be positive.

Unfortunately it is not reasonable to assume that the covariance terms will be constant for each treatment group. Consider the simplest form of subsidy system in which a lump sum payment, varying across treatment groups, is made. Let z_k denote the rate of subsidy to the k^{th} group, and let γ_i denote the fraction of the subsidy spent on housing by the i^{th} individual; it is assumed that γ_i does not vary with the level of the subsidy. Defining housing change in terms of expenditure, this gives:

$$\Delta h_{ik} = \gamma_i z_k \qquad \Delta y_{ik} = (1-\gamma_i) z_k$$

The covariance terms of interest can then be written:

$$\mathrm{Cov}\{\beta_i \Delta h_{ik}\} = z_k \, \mathrm{Cov}\{\beta_i \gamma_i\}$$

$$\mathrm{Cov}\{\alpha_i \Delta y_{ik}\} = -z_k \, \mathrm{Cov}\{\alpha_i \gamma_i\}$$

In this simple instance, the covariance is proportional to the rate of subsidy and will vary across treatments; it would nevertheless be simple to adjust for such variation by normalization. A similar normalization could be carried out in the context of the proposed Housing Allowance Experiment, although the implicit behavioral assumptions remain obscure. The total income change resulting from participation in the experiment can be separated into two effects--a change in disposable income, defined as increased purchasing power over non-housing goods, and a change in housing, defined as the increased expenditure on housing. Treatment groups will differ amongst each other with respect to the total income change and with respect to its breakdown between housing and non-housing expenditures. "Normalization" can then be carried out by dividing both sides of the equation (4.2) by the total income change:

$$\frac{\Delta x_k}{\Delta y_k+\Delta h_k} = a_o + a_1 \left(\frac{\Delta y_k}{\Delta y_k+\Delta h_k}\right) + a_2 \left(\frac{\Delta h_k}{\Delta y_k+\Delta h_k}\right) + \varepsilon_k \quad (4.3)$$

The constant a_o corresponds to the covariance terms of equation (4.2), which, now being "normalized," are not expected to vary between treatment groups. The dependent variable is now the <u>rate</u> of change of the (development) variable with respect to changes in income; the equation can be interpreted as saying that the rate of change of the development variable with respect to income depends on the way in which gains in income are allocated between housing and other expenditures. Incidentally, normalization can be justified on grounds relating to the distribution of the error term, which would otherwise tend to be heteroscedastic.

The equation cannot, of course, be estimated in this form because normalization reduces the dimensions of independent variation to one. If we denote the mean percentage of the increase in income for the k^{th} growth which is spent on housing by p_k (4.3) can be written as:

$$\frac{\Delta x_k}{\Delta y_k + h_k} = (a_o + a_1) + (a_2 - a_1) p_k + \varepsilon_k \qquad (4.4)$$

or more simply:

$$\frac{\Delta x_k}{\Delta y_k + \Delta h_k} = b_o + b_1 p_k + \varepsilon_k \qquad (4.5)$$

The hypothesis that housing change influences child development, over and above the effect of income change, can now be investigated through tests on the coeffecent b_1. Rejection of the null hypothesis that b_1 is zero is equivalent to a finding that housing change has an independent (and beneficial) effect on child development.

4.1.2 Sample Sizes and Confidence Levels

A possible application of simple regression analysis to the relationship between child development and housing expenditures has been presented in the previous section. This approach, while extremely limited in its scope, appears to avoid most of the problems of inference which result from self-assignment of families to housing within the treatment groups. In weighing the advantages and disadvantages of this approach, however, consideration must also be given to the confidence levels which might be attained and the sample sizes which would be required to attain them.

The most convenient way to approach this problem is through the length of the confidence interval for b_1; the $100(1-\alpha)$ percent interval for b_1 is given by:

$$\hat{b}_1 \pm t_{\alpha/2} \frac{\hat{\sigma}_\varepsilon}{\sum_i p_i^2}$$

where $t_{\alpha/2}$ denotes the critical value of the students t distribution, and where b_1 and $\hat{\sigma}_\varepsilon$ denote the least squares estimates of b_1 and the standard error of estimate respectively. Convenient rules of thumb for this purpose might be to require the length of this interval on one side to be less than 5×10^{-3} and less than 2×10^{-3}. (These would correspond, for every $1000 of income spent exclusively on housing, to 5 and 2 point gains on a scale with population standard deviation of 20). The p_i's correspond to the mean fraction of incremental income spent on housing for each of the treatment (and control) groups. Evidently the larger the variation of the p_i's the smaller will be the required sample size. Unfortunately these cannot be known in advance of the actual experiment. Table 17 provides required sample sizes for confidence intervals of this length for alternative values of Σp_i^2. In each case, a single-wave study is contemplated with mean scores, rather than gains, compared between treatment groups. The error variance is given by the population variance of the measure divided by the number of observations in each treatment group (i.e., the variance of the sample means of the treatment groups). It is assumed that 11 observations are available (10 treatment groups plus one control group).

These calculations are not particularly encouraging. In the worst case, where there is only limited variation between groups with respect to housing choice and where the one directional confidence interval spans only 10 percent of the standard deviation of the measure, the experiment would require over 50,000 observations. In the event that the allowances lead to quite wide variation in the choice between housing and other income ($\Sigma p_i^2 = 0.5$), a 5 point interval would require around 2000 observations.

These calculations have been deliberately restricted to dependent variables based on single observations on families and children. The required sample sizes could be reduced if, instead, <u>differences</u> between baseline and subsequent measurements were used (since test-retest correlations will offset the need to make two observations on each child).

Length of confidence interval / $\sum_{i=1}^{11} p_i^2$	5×10^{-3}	2×10^{-3}
0.10	820	5000
0.25	330	2000
0.50	165	1000

Table 14

Required Sample Sizes (observations per treatment group) for alternative 95% confidence intervals and alternatives Σp_i^2's.

If this strategy were followed, however, it would be impossible to know whether the variation in the independent variable p_i will be sufficient to make reliable inferences, in advance of the baseline survey. For this reason, a longitudinal approach is not recommended for the Housing Allowance alternative.

Considerations similar to those which lead to rejection of a longitudinal approach suggest that a single-wave cross-sectional analysis should not be undertaken until it becomes clear that adequate variation in the independent variables exists. Furthermore, by waiting to find out the magnitude of the effect of the allowance scheme on family housing divisions, it is possible to leave more time for the putative effects on child development to take place--and this in turn will permit the use of less rigorous tests for significance.

4.1.3 The Housing Allowance Experiement as a Vehicle for the Investigation of Housing/Child Development Interaction--Summary.

It has been demonstrated that the Housing Allowance Experiment does not lend itself readily to the purposes of this study. Families are assigned experimentally not to planned combinations of housing, but to planned combinations of incentives; the intervening response of families constitutes a self-selective mechanism which largely eliminates the desirable experimental aspects of the design--at least for the purposes of this study.

A partial solution to this problem can be found by aggregating across treatment groups and performing a simple linear regression of the child development variables of interest on the mean percent of family income spent on housing. This approach is, itself, somewhat tainted by aggregation problems and, in any case, it provides no insight into which dimensions of housing quality are influential in determining the rate of child development. Furthermore, unless the differential incentives of the allowance scheme are powerful enough to provide a wide range of response in terms of housing choice, the sampling variances of the least-squares estimates will be unacceptably high.

This approach, nevertheless, provides a useful back-up to the non-experimental study. If the non-experimental study fails by reason of self-selection the only alternative will be an experimental approach of one kind or another. The Housing Allowance Experiment provides an opportunity to take advantage of an existing study at, at least relatively, to other experimental designs, a reasonable cost.

One pre-condition is necessary if this approach is to be sensible. This condition is the success of the Housing Allowance Experiment in achieving significant variation between treatment groups with respect to their housing choices, and in particular with respect to the extent to which housing allowances are used to upgrade as opposed to the extent to which they are simply a subsidy to existing rental payments. Depending on the extent of this variation and this will be readily available when the experiment has been operational for a year, it will be possible to estimate required sample sizes and costs, and to make a decision to collect data in the second or third year of the experiment. To wait longer would be to run into problems of attrition, to be less patient would provide inadequate time for the effects of environmental change to work themselves out.

4.2 Alternative Experimental Designs

As a vehicle for the investigation of the relationship between child development and the housing environment, the Housing Allowance Experiment has been shown to be unsatisfactory. The principal reason for this finding was the lack of an experimentally accurate mechanism for assigning families to housing types. Attempts to avoid this difficulty lead to secondary difficulties; in particular, the need to collapse the characterization of housing change into one dimension is particularly undesirable in the context of this study. Those problems naturally suggest the possible development of alternative experimental approaches in which the child/housing development investigation would be the primary, and not the secondary purpose of the study.

To achieve a better approximation to the properties of an ideal experiment, two elements are required. In the first place, there is a need to insure that subjects are randomly assigned to alternative treatments. In the second place, the treatments themselves should represent planned combinations of the variables of interest in which there is no confounding of independent variables either with each other, or with variables excluded from the design. Both considerations suggest the desirability of conducting the experiment in a limited geographical area (this area could of course constitute a single block in a larger experiment); in this way it is possible to avoid confounding housing variables with regional differences, and an it is also possible to attempt randon assignment of individuals to housing treatments when the separating distance between alternaive treatments is not large. The design of the experiment is now condidered within the constraints imposed by limitation of the experiment to an area small enough to attempt random assignment of families to alternative housing programs.

4.2.1 Methods of Subject Assignment

The need to assign subjects at random to alternative housing projects places severe constraints on this experiment. In the first place, subject families must perceive all the housing opportunities provided by this experiment as preferable to other available private or public housing (including their present location). If this condition is not met, families will not agree to enter their assigned projects--or will only agree if they are assigned to preferred alternatives; this would then re-create the self-selection problem in an experimental setting. In practice, this may be a hard condition to meet, particularly in view of the planned variations in the environmental quality of the treatments themselves.

There are several ways of attempting to insure that the acceptance rate of assignments is adequate. In the first place, as already mentioned, the treatments must be located in reasonable proximity to one another, and all must have adequate transportation access. If the incremental costs associated with transportation to and from the assigned housing treatment exceed the dollar value of the improved housing quality, it is rational behavior for the family to decline the assignment. A second method of insuring a high acceptance rate would involve screening applicants to determine their willingness to move into any of the experimental projects. Two potential difficulties are associated with such a screening process: firstly, families may not give reliable answers in part because of the hypothetical nature of the question, and in part because they wish to retain the opportunity to move into one or more of the desirable projects; secondly, the process of screening further reduces the generalizability of the findings to that limited group of subjects for whom any of the alternatives are preferred. This can only really be overcome if some effort is made to insure that relocation into even the least desirable alternative is still acceptabel to a significant percentage of the population of interest. Efforts to insure the general acceptability of all the alternative treatments may, in addition, include some form of special payment over and above the standard rent supplement. To the extent that associated real income change exercises its own independent effect on the dependent variables of interest, however, problems of inference will be correspondingly complicated.

The selection and assignment of experimental families to housing treatments cannot be discussed without some reference to the possible use of control families. If the experimental design calls for comparisons of measured development in families assigned to selected housing programs vs. families not assigned, it would be necessary to select from amoungst applicants a set who would be denied admission to any of the experimental programs. In the context of a longitudinal design, some means for insuring their continued participation will then be required.

4.2.2 Design of Experimental Programs

The conduct of a truly experimental study provides an opportunity to achieve a much sharper contrast between alternative housing environments than is possible in a non-experimental setting. The non-experimental approach limits the investigation to a contrast between available housing environments, and this necessarily leads to some confounding between included and excluded variables. In the experimental setting on the other hand, the freedom exists to match experimental programs almost exactly on physical characteristics, and to a lesser degree, to match them on excluded school and neighborhood variables. In the latter case, the freedom is less than absolute, since once again the experimenter is forced to work within the constraints fo what exists. It is not realistic, for example, to build and populate new neighborhoods for the purpose of this experiment!

The number of experimental programs to be included in a design of this kind will depend upon the number of dimensions of interest. Consistency with the objectives expressed in the non-experimental design would be achieved by having four programs distinguished from one another in terms of the socio-economic characteristics of the neighborhood and mix of income within the project. Thus two of the programs should be located in high-income neighborhoods and two in low-income neighborhoods; within each of these pairs, there should be one project with a high percentage of moderate income families and one project with very few moderate income families. Attempts to achieve further distinction between programs on variables such as school "quality" or ethnic mix do not appear to offer much promise of success, given the necessity to live within the constraints of proximity and existing neighborhoods.

The exclusion of physical characteristics of the housing from the independent variables of interest suggests that each of the four projects should be similar in their design. Futhermore, since the size of the project is not in itself of intrinsic interest, the number of units to be built should be determined on statistical grounds. In two of the programs, half the resident population should be moderate income families, and therefore not experimental subjects. The programs should be at least the size required to house an adequate subject population. If adequate is defined in terms of the sample sizes required in the non-experimental design and allowing for some attrition, this would require each project to contain between 125 and 175 units.

4.2.3 Costs and Conduct of the Experimental Alternative

The experimental alternative should be conducted as a longitudinal investigation, with one survey wave performed prior to admission and one subsequent wave after a pierod of around three years. Test for significant differences in mean gains are then carried out exactly as in the third phase of the non-experimental investigation. Because of randomization in assignment, checks for self-selection at the time of the first wave are redundant. Checks for self-selection amongst exiting populations (which may be larger in the less desirable projects), should still be carried out.

In attempting to arrive at an estimate of the costs of this alternative, it is worth emphasizing that the actual costs of measurement and data analysis remain essentially unaffected. The real cost elements relate to the construction and maintenance of the projects and to the possible subsidization required to insure stability and absence of self-selection amongst the resident populations. The construction and maintenance costs should only be charged in total to the experiment if the programs possess no social value over and above the results of the experiment. This is clearly not true. At a minimum, the social value of the projects must equal the discounted value of future rental payments, leaving the real cost of the experiment to be determined by the discounted stream of rent supplements. This is probably a somewhat exaggerated estimate of the costs of the experiment, because to the extent that the resources used to develop the experimental projects were applied to an alternative development elsewhere, similar rent supplement charges would be incurred. Any way to assess the opportunity costs of this experiment are by nature arbitrary, but if a relatively high estimate of around $40/month per family is used, the undiscounted housing costs of the experiment would be around $1,500,000 over a period of five years;[1] five years might be an average

[1] $40.00 a month is the estimated interest subsidy for 235 and 236 programs for a family of 5 paying an average annual rent of $960. "Improving Federal Housing Subsidies: Summary Report," Bernard J. Frieden. Paper submitted to Subcommittee on Housing Panels, Part 2, Committee on Banking and Currency, House of Representatives, June 1971.

length of residence for experimental families. To this should be added costs for survey design and administration of around $500,000, for a total cost of $2,000,000. It is this cost which makes the experimental approach less desirable.

There is, of course, an additional reason for exercising caution with respect to demonstration programs of this kind. The experimental programs contemplated involve the construction of medium-sized low-income housing projects in middle and upper income communities. Even when such projects are sponsored within communities of this kind, effective community opposition is almost invariably forthcoming; when the project is sponsored outside the community and the basic design and location decisions reflect research interests, the likelihood of obtaining community support in any suitable location may be extremely small. These considerations, only touched on here, suggest that an experimental project may be infeasible, as well as expensive.

4.2.4 Alternative Experimental Approaches - Summary

Truly experimental approaches to this investigation which involve site-selection, design and construction of new housing projects, together with randomized admission procedures, are conceptually possible, and possess desirable properties from the viewpoint of statistical inference. They do, however, possess severe practical drawbacks. In the first place, satisfactory admission procedures may require payments to families over and above the basic rent subsidy for publicly supported housing. Secondly, even with expensive admission procedures, there will be some confounding of included neighborhood variables with omitted neighborhood variables and some self-selection amongst subjects. The experiment would not only cost about five times as much a a comparable non-experimental approach, but its feasibility is quite problematical. For these reasons, this approach appears to be rather unpromising.

CHAPTER FIVE

MEASUREMENT

The study of the effects of environmental change on children and their families relocating into new housing and new neighborhoods is broader in scope than studies of more limited intervention, and presents greater problems in isolating and controlling variables. If these obstacles can be effectively overcome, however, this study presents an opportunity, through its very broadness, to make further differentiations between the dimensions of child development and their response to environmental change. For these reasons, an attempt has been made to develop as comprehensive a battery of instruments as the constraints of time and cost allow.

The basic criteria applied to the solution of instruments can be summarized as follows:

(1) Is the area of child development measured by this instrument one which could reasonably be expected to be affected by the described environmental changes; is it reasonable to expect that developmental gains in this area will become precursors of adequate adult functioning?

(2) Has this test been subjected to adequate validity and reliability testing and do norms exist for both middle and low income children? Furthermore, is this instrument being used for the longitudinal evaluation of children so that useful comparisons can be made?

(3) Can this test be accomodated within a battery which satisfies reasonable constraints in terms of time and cost of administration?

These criteria have been applied to the selection of instruments for children in three separate age categories. These categories effectively correspond to pre-school, elementary school, and high school age ranges. Measures of change in parental attitudes and family situation are obtained through use of a parent interview instrument. Descriptions of all the selected measures, together with interview instruments, are provided in Appendices A and B.

Comprehensive review of the literature of environmental change provides mixed support for the hypotheses of developmental affects.[1] Although theory and intuition suggest that the basic lifestyle changes associated with major family relocation should contribute to the cognitive, socio-emotional, and physical development of children, previous research is inconclusive, except perhaps in the area of physical development where significant reductions in accidents and poisoning have been found. The methodological difficulties of most of these studies, together with the highly limited extent of the environmental change studied, conspire to leave the questions of environmental impact in cognitive, socio-emotional, and general health unresolved. For these reasons, measures are proposed in each of these areas. Measures of objective gains in income, employment, and housing quality, together with measure of parental aspirations and community involvement are also attempted.

Final selection amongst alternative instruments must necessarily reflect the subjective preferences and interests of the investigation and this work is no exception. It is important, however, to distinguish between biases of this kind, which reflect special interest in particular dimensions of development, and biases of another kind, which reflect culturally determined definitions of development and under-development. It is important, given the subject populations in these studies, to recognize the difficulty of transcending cultural differences in the measurement of cognitive development

---

[1]See Volume III, Literature Review.

in particular, and to maintain an awareness that "under-developed" by the standards of a dominant culture is not necessarily under-developed by the standards of another culture. The safeguard lies in the caution with which results are to be interpreted; thus, depending on professional judgment of the validity of a particular instrument, gains may either be interpreted as "development" or "assimilation." In either case they are results of interest.

5.1 Recommended Measurement Instruments for Children

Recommendations for the measurement of children are presented in three categories corresponding to the cognitive, socio-emotional, and physical dimensions of development. Age group appropriateness is indicated as follows:

1 = pre-school

2 = elementary school

3 = junior and senior high school

Descriptions of each of the recommended instruments are provided in Appendix B.

5.1.1 Cognitive Development

Four separate aspects of cognitive development are addressed: verbal skills, concept formation, school achievement, and general intelligence. Recommendations are as follows:

Verbal Skills

1. Peabody Picture Vocabulary Test - size of vocabulary (1, 2, 3)

2. ITPA Subtest of Verbal Expression - Verbal fluency and ideational flexibility (1, 2)

3. * ETS Story Sequence Test - ability to understand and produce a simple story (1, 2)

4. ITPA Subtest of Grammatic Closure (1,2)

Concept Formation

1. * Sigel Object Categorization Test - consistency and level of classificatory skills (1, 2)

2. Zimiles Matrix Test - abstraction of similarities (1, 2)

<u>School Achievement</u>

1. Metropolitan Achievement Test - reading and math achievement (2, 3)

<u>General Intelligence</u>

1. Pre-school Inventory (1)

2. Stanford-Binet Intelligence Test (1, 2, 3)

5.1.2 <u>Socio-Emotional Development</u>

Five distinguishable dimensions of socio-emotional development are addressed: mental health, self-concept, future goals and expectations, locus of control, and peer interactions. Recommendations are as follows:

<u>Mental Health</u>

1. Selected Subtests from California Test of Personability - measures of personal and social adjustment (2, 3)

2. Behavioral indicators of psychological and social adjustment (1, 2, 3)

<u>Self-Concept</u>

1. Brown IDS Self Concept Referents Test (Par I, Self Referent) (1, 2)

2. Coopersmith Self Esteem Inventory (2, 3)

Future Goals and Expectations

1. Battle's Achievement Expectancy Scale (2, 3)

Locus of Control

1. Stephens-Delys Reinforcement Contingency (1, 2)
2. Crandall's Intellectual Achievement Responsibility Scale (2, 3)
3. Nowicki-Strickland Personal Reaction Survey (2, 3)

Peer Interactions

1. Ohio Social Acceptance Scale - child's standing in class and his perception of it (2, 3)
2. Parent and child interview items (1, 2, 3)

5.1.3 Physical Development

No general pediatric examination yielding comparable and quantifiable data was found appropriate for this study. It is thus recommended that a series of tests screening for handicaps and tests specific to poverty and housing related problems be given as a battery. This procedure will be about equal in cost to a comprehensive pediatric examination, but will provide more specific information. In addition to the tests, a parent questionnaire on previous immunizations, diseases, and other medical history, such as those used by school systems for mail responses, should be used.

The following tests are suggested:

(1) Hearing test (including screening for middle ear infections)

(2) Vision test (in depth for detected abnormalities)

(3) Nutrition (anemia test, height-weight, vitamin level test)

(4) Lead Poisoning (blood or urine test)

5.2 Supplementary Measures for Parents

In addition to the basic test batteries described above, it is desirable to collect additional information from the families of the target children. This is done by means of a Parent Interview Instrument.[1] The Parent Interview is designed to provide information on objective gains (housing quality, spendable income and employment), aspirations, and community involvement. Data collected along each of these dimensions will be treated as dependent variables in the analysis.

The Parent Interview serves two additional purposes. A section on Peer Interactions corroborates information from the Ohio Social Acceptance Scale gathered for school-age children, and provides the only source of information on the peer interaction of pre-schoolers. Secondly, the first part of the Parent Interview provides demographic data required for screening and sample selection prior to the first survey wave.

[1] The Parent Interview Instrument is presented in Appendix A.

5.3 Cost of Administration

Cost estimates are developed on the basis of 100 children and 100 parents. Since some parents will have more than one child under study, the estimates may be somewhat high. The selection of 100 children who meet matching variable requirements is assumed to required 250 short screening interviews with parents. Child interviews will not be administered to pre-school children.

Parent Interviews		
250 x P.I. Pt. 1@4/day	62 days	
100 x P.I. Pt. 2@4/day	25 days	
Cost of Interviews @ $20/day		$1,740
Child Interviews		
75 x C.I. @4/day	19 days	
Cost of Interviews @ $20/day		380
Child Tests		
100 x Battery @ 2/day[1]	50 days	
(Preschool 1 1/2 hours in 2 sessions School Age 4-5 hours in 3 sessions).		
Cost of Child Tests @ $35/day		1,750
Coding		
250 P.I. Pt. 1 @25/day	10 days	
100 P.I. Pt. 2 @12.5/day	8 days	
100 C.I. @20/day	5 days	
100 Batteries @20/day	10 days	
Cost of Coding @ $20/day		660

[1] Estimated average time per child for group and individual testing.

Health Examinations	
100 x Health Test Exam @ $20/test[1]	$ 2,000
Coordination	
1 coordinator for 30 man days	
Cost of coordinator @ $50/day	1,500
ESTIMATED TOTAL DIRECT COST FOR 100 FAMILIES AND CHILDREN	8,030

ESTIMATED DIRECT COST PER FAMILY AND CHILD: $80.00

These costs represent only the direct labor costs of carrying out the field work. There will, in addition, be some required travel and, no doubt, some indirect charges such as overhead and administration expense. For 100 families and children, the additional charges should be roughly as follows:

1 This is estimated as follows: Hearing Test ($5-$10) Vision Test ($5-$10), Anemia ($1.00), Vitamin Level ($5.00), Blood or Urine Test for Lead Poisoning ($1.00). An average cost estimate of $20.00, which is roughly equivalent to a standard pediatric examination therefore seems reasonable.

Travel (Coordinator)		
Round-trip to site average	$150	
30 per diems @ $25	750	
Total Travel		$ 900
Overhead (Coordinator & Coders)		
100% of Direct Labor		2,160
General and Administrative Expense		
15% of Total Costs ($8,030+$900+$2,160)		1,663
Fee (7% of Total Costs)		892
TOTAL COST PLUS FIXED FEE		13,635

There is, of course, considerable variation amongst contractors in the way in which they account for indirect charges; nevertheless, an average cost of around $136 per child should be adequate under almost any circumstances.

APPENDIX A

Parent and Child Interviews

Family Name ____________

PARENT INTERVIEW
Part I: Screening (either parent)

The information and opinions we will ask for in this interview are for a study of children's development as it is related to housing. Everything you say will be held in strictest confidence by the research team. The results of the study may become public, but only in statistical terms, such as group percentages. You as an individual will not be named or discussed, and this interview will not be used by anyone else or for any other purpose.

ADDRESS: ____________

TELEPHONE: ____________ CONVENIENT TIMES TO CALL: ____________

LENGTH OF RESIDENCE:

	Date Moved In	Length of Residence years	Length of Residence months
Current residence	______	___	___
Previous residences (going back in time to 1965)-addresses:			
____________	______	___	___
____________	______	___	___
____________	______	___	___
____________	______	___	___
____________	______	___	___
____________	______	___	___
____________	______	___	___
____________	______	___	___

NAME OF HEAD OF HOUSEHOLD: ____________ AGE:___ RACE:___

NAME OF MOTHER (if different): ____________ AGE:___ RACE:___

Employment status Head of Household

_____ Full time
_____ Part time
_____ Temporary unemployment
_____ Permanent unemployment

FAMILY INCOME AND EDUCATION INFORMATION:

	HEAD OF HOUSEHOLD	OTHER INCOME EARNERS (GIVE RELATION TO HEAD OF HOUSEHOLD)			
		rel'n:	rel'n:	rel'n:	rel'n:
Occupation (specific)					
Principal Employer					
Annual income from principal employment					
Other job (1)					
Annual income from (1)					
Other job (2)					
Annual income from (2)					
Total employment income (annual)					
Highest year education (or degree)					
Supplementary education (night or technical)					

ADDITIONAL INCOME:

Child Support		TOTAL ADDITIONAL INCOME	
Social Security			
AFDC, ADC, or other welfare payments		TOTAL JOB INCOME (from above:	
Unemployment compensation			
Disability		TOTAL ANNUAL FAMILY INCOME:	
Pension			
Other			
Total			

CHILDREN (in birth order):

Name	Age	Sex	School Grade	Name of Nursery or School	Names and dates of previous Schools, Nursery Schools or Day Care Centers

Do any of the children we have listed have physical or mental disabilities? (Fill in table below.)

Have any of them spent prolonged periods (a month or more) hospitalized or confined to bed? If so, why?

Are any of the children adopted or foster children?

NAME OF CHILD	DISABILITY	HOSPITALIZATION	ADOPTED/FOSTER

PARENT INTERVIEW

Part II: Mother (attach Screening Interview)

I. EMPLOYMENT (if time lag since Screening Interview, check current employment and update if changed)

(1) We would like to know how living here has affected your family's ability to make a living. First, can you remember the jobs held by your family's income earners before the present? We will take them one at a time back to 1965.

MEMBER OF HOUSEHOLD (starting with head)	PREVIOUS DATES OF JOBS	NAMES OF PREVIOUS JOBS	ANNUAL INCOME (FINAL RATE EACH JOB)

(2) For each job change or new job since the beginning of your residence <u>here</u>, was the opportunity or necessity to change jobs related to living here? (e.g., easier access, better job market, etc.) If so, how was it related? (Repeat for OTHER working members of household.)

MEMBER OF HOUSEHOLD	JOB	HOUSING RELATION

(3) What effect would you say, then, that living in this place has had on your family's ability to make a living?

Helped __________
No Change __________
More Difficult __________

II. SPENDABLE INCOME

(1) How much is your rent per month here? ______________

Annual rent (to be computed by inverviewer): ______________

(2) How much was your rent per month where you lived just before you moved here? ______________ Annual rent: ______________

(3) About how much money do you have to pay each month in bills directly related to housing?

(4) About how much do you spend on the average each month on things indirectly related to housing?

	PRESENT RESIDENCE	PREVIOUS RESIDENCE
heat (no. mo. per yr.)		
gas & elec.		
repairs (ave.)		
parking		
other (specify)		
MONTHLY TOTAL		
transportation		
furniture		
higher shopping costs		
other (specify)		
MONTHLY TOTAL		
MONTHLY TOTAL (3) & (4) combined		

III. HOUSING AND ENVIRONMENTAL QUALITY

	PRESENT RESIDENCE		PREVIOUS RESIDENCE	
	YES	NO	YES	NO
Housing				
(1) Facilities				
Private outside entrance				
Dual egress				
At least one window per room				
Electric lights and outlets in every room				
Installed heating in every room				
Hot water				
Storage (at least one closet per bedroom)				
Laundry facilities in building				
Complete kitchen (sink, stove, refrigerator)				
Complete bath (basin, toilet, tub or shower)				
TOTAL (2 pts. each) yes				
(2) Maintenance				
Immediate attention to plumbing or electrical problems				
Immediate attention to heat/hot water failure				
Free from infestation (rodents)				
Free from infestation (roaches and other insects)				
Sanitary garbage collection system (ask how & when collected)				
Structural safety (holes, railings, etc.)				
No flaking paint				
Adequate security (locks)				
Lit entrance and hallways				
Repainting at least every three years				
TOTAL (1 pt. each) yes				
(3) Occupancy				
No shared facilities (e.g., kitchen, bath) (3 pts.)				
No more than 2 persons per bedroom (3 pts.)				
No more than 1 1/2 persons per room (excluding kitchen & bath) (2 pts.)				
TOTAL				
HOUSING TOTAL (Max.40)				

IV. ASPIRATIONS

(1) What kind of job(s) would you most like your child(ren) to have when he (they) grow up? (Try to elicit answer without reading categories first. Repeat for each child, starting with oldest.)

(2) Do you think your child(ren) are likely to get jobs as good as those you would like them to have? (Write Yes, No, DK in each box.)

(3) How far would you like your child (ren) to go in school? (Repeat 6-11 for each child.)

Names of children.				
(1)				
Services (hairdresser, mechanic, etc.)				
Businessman				
Steady work in a plant				
Teacher/Social Worker				
Minister (Rabbi, etc.)				
Office worker (clerk, secretary, etc.)				
Professional sports				
Military				
Housewife				
Professional (doctor, lawyer, etc.)				
Own a small store				
Other:				
(2)				
(3)				
Professional or graduate work				
Graduate from a 4-year college				
Some college but less than 4 years				
Technical, nursing, business school after high school				
Finish high school				
Don't care if finish high school				
DK				

Names of children

(4)	Do you think your child(ren) will get as far as you would like?	Probably yes				
		Maybe				
		Probably no				
		DK				
(5)	How good a student do you want your child(ren) to be with respect to his(their) studies? (Read off the choices.)	In the middle of the class				
		Above the middle of the class				
		One of the best in the class				
		Just good enough to get by				
		DK				

[Answer (6)-(10) for school-age children only.]

(6)	Do you think your child(ren) is(are) doing as well in school as you would like?	Probably yes				
		Maybe				
		Probably no				
		DK				
(7)	Do you want your child(ren) to study or read a regular amount of time for each school day?	Yes				
		No				
		DK				
(8)	Do(es) your child(ren) study or read as much as you would like each day?	Yes				
		No				
		DK				
(9)	Have you encouraged your child(ren) to take any lessons outside of regular school because you felt it would help him(them) do better in school? What kind? ______________________	Yes				
		No				
		DK				
(10)	Do(es) your child(ren) do any of these things at present?	Yes				
		No				

(11) How important are school grades to you?

	VERY IMPORTANT	IMPORTANT	NOT SO IMPORTANT	UN-IMPORTANT
to you?				
to your husband?				
to teachers?				
to ____________?				
____________?				
____________?				

(names of children)

V. COMMUNITY - SUBJECTIVE EVALUATION AND PARTICIPATION

Hand Card

Very good
Good
Neither good nor poor
Poor
Very poor

(1) a. How would you rate this neighborhood on the following facilities and characteristics? Please choose your answers for each one from the card.

b. Now let's go through the same categories for your previous neighborhood.

	PRESENT NEIGHBORHOOD					PAST NEIGHBORHOOD				
	VG	G	NG NP	P	VP	VG	G	NG NP	P	VP
a. Shopping: food										
b. Shopping: other than food										
c. Churches (Synagogues, etc.)										
d. Medical facilities										
e. Transporation										
f. Day Care										
g. Schools										
h. Parks & Playgnds.										
i. Police protection										
j. Recreation (movies, museums, libraries)										
k. Garbage collection, sanitation										
l. Street clearance and repair										
m. Juvenile delinquency; drugs, vandalism, etc.										
n. Adult crime rate										

(2) a. Using the same card, how would you rate the people in this community?

b. Now let's go through the same categories for your previous neighborhood.

	VG	G	NG NP	P	VP	VG	G	NG NP	P	VP
a. Friendliness of people living in this project (bldg)										
b. Helpfulness of people living in this project (bldg.)										
c. Friendliness of (bldg.) project personnel										
d. Friendliness of people living in this neighborhood outside the project										
e. Helpfulness of people living in this neighborhood outside the project										

(3) a. About how many families in this project do you know by name? Please list as many names as you can remember. The families you list will not be contacted. Listing them just helps us get more accurate information about how well people can get to know each other here.

FAMILY NAME	FRIENDS

PROJECT -
Total families: ____
Total friends: ____

b. About how many families in this project would you call your friends? ________ (Check off list in Question 3a.)
Total: __________

c. About how many families in this neighborhood, not living in the project, do you know by name? Please list as many names as you can remember. The families you list will not be contacted. Listing them just helps us get more accurate information about how well people can get to know each other here.

FAMILY NAME	FRIENDS

NEIGHBORHOOD
Total families: ______
Total friends: ______

d. About how many families in this neighborhood, not living in the project, would you call your friends? (Check off on list in Question 3c.) Total: ________

e. Do most of your friends live either in this project or this neighborhood? ____________ If no, do most of your friends live in your former neighborhood? ____________

f. Do you have as many friends now as you did where you lived before moving here? ____________

			Weekly	Once/ month	Several times/ year	Once/ year	Takes Children
(4) Do you go to meetings, services, or other events for groups of people in this neighborhood? About how often?	a.	Church (Synagogue)					
	b.	PTA (or other school group)					
	c.	Clubs					
	d.	Political groups					
	e.	Boy or Girl Scout mothers					
	f.	Volunteer services					
	g.	Tenant organizations					
	h.	Sports teams					
	i.	Other ________					
(5) What places in this neighborhood do you visit? About how often?	a.	School					
	b.	Library					
	c.	Parks & Playgrounds					
	d.	Church (Synagogue)					
	e.	Movies					
	f.	Clinics (medical facilities)					
	g.	Museums					
	h.	Bowling alleys (or other sports)					
	i.	Other ________ ________					

(6) Do you usually take your children to the meetings or places we have just talked about? (Check each item in last column.)

		Weekly	Once/ Month	Several times/ year	Once/ year	With whom
(7) Are there meetings or places your children go to by themselves or with people in this neighborhood? About how often? Who do they usually go with? (With friends, alone, with other adults?)	a. Libraries					
	b. Scouts					
	c. Church/(Synagogue)					
	d. Sports teams					
	e. Lessons					
	f. YM(W)CA					
	g. Movies					
	h. Sports events					
	i. Playgrounds & parks					
	j. Other ____________					
	____________					

(8) Who usually takes care of your children when you cannot?

____ a. Older sibling
____ b. Relative (in child's home)
____ c. Relative (outside home)
____ d. Non-relative (in home)
____ e. Non-relative (outside home)
____ f. Daycare center
____ g. Other ____________

(9) About how many hours a week is ____________ (name of child) cared for this way? ________ (hours/wk.)

(Repeat for each child.) ____________ ________

____________ ________

____________ ________

(10) Do you think your children would have had good opportunities to get ahead in life if you had stayed in your previous neighborhood? ____________

(11) Do you think your children will have good opportunities to get ahead in life if you stay in this neighborhood? ____________

(12) Do you think your children would have a better chance to get ahead in life if you moved to another home and neighborhood?

____

(13) Do you think you will move anywhere with better opportunities for your children while they are still in school? ______ Yes

______ No

______ DK

VI. CHILDREN: PEER RELATIONSHIPS

(1) About how much time do(es) your child(ren) spend with children living inside this project and with children living outside the project, but in this neighborhood?

	Project	Neighborhood
Sometime every day	______	______
Sometime several days a week	______	______
Sometime once a week	______	______
Less than once a week	______	______
Rarely	______	______
Never	______	______

(2) a. Do you think you know who your child(ren)'s friends are? ______________

b. Can you try to list as many of your child(ren)'s friends by name as you can (for each child in turn)?

Name of child	Name of friend	School of friend	Grade in school	LOCATION: In project	In neighborhood	Outside	TIME SPENT WITH: Sometime every day	Sometime several days a week	Sometime once a week	Less than once a week	Rarely	Never

(3) Do you think your child(ren) has enough opportunity to make friends here? ____________

(4) Do you think your child(ren) has(have) as many friends here as they did where they lived before? ____________

(5) Are you pleased with the kinds of friends your children have here? ____________ If no, why not? ____________
____
____

(6) Can you describe any differences in the kinds of friends your child(ren) has(have) here and where you lived before?
____
____
____
____

CODING OF PARENT INTERVIEW

PART I: Screening (10 data points)
(duplicate care required for each subject child)

CARD I

Columns		
1-5	Project name code	
6-10	Family name code	
11	No. residences since 1965 (9=9 and above)	
12	Sex of head of household (M or F)	
13-14	Age of head of household	
15	Race of head of household	1=white 2=black 3=Spanish 4=Oriental 5=other
16	Occupation head of household	9=professional, technical 8=managers, proprietors 7=clerical 6=sales 5=service workers 4=craftsmen, foremen 3=operatives (skilled) 2=laborers (unskilled) 1=household
17	Occupational status of head of household	1=full time employment 2=part time employment 3=1 plus 2 4=temporary unemployment 5=retirement (permanent unemployment)

18	Educational level of head of household	8=Advanced degree 7=graduate of 4 year college 6=college (less than 4 years) 5=technical, nursing, business training after high school 4=graduate high school 3=some high school 2=8th grade 1=less than 8th grade
19	Total Annual Family Income	1=under $3,000 2=$3,000-$4,500 3=$4,500-$6,000 4=$6,000-$7,500 5=$7,500-$9,000 6=$9,000-$10,500 7=over $10,500
20-24	Absolute Family Income (Annual)	
25-26	Total No. Members Household	
27	No. Children Household	

CODING OF PARENT INTERVIEW

Part II: Mother

[Coding will be done separately for each child in the study. Where questions are answered for each child (e.g. Sec. IV), code only answer for specific child; where collective answers are given, they will be entered as collective scores for every child (e.g., V(7)].

CARD II

Columns		
1-5		Project Name Code
6-10		Family Name Code
11-13		Child Name Code
14	I(1)	No. Jobs Since 1965: Head of Household
15	I(2)	No. Jobs since Moving to Current Address
16	I(3)	Relocation Effect on Earning a Living 1=Helped 2=No Change 3=More difficult
17-20	II(1)	Annual Current Rent
21-25		Annual Current Income
26-27		Annual Current Rent/Annual Current Income (Express fraction as 2 digit number, e.g., .25=25)
28-31	II(2)	Annual Previous Rent
32-36		Annual Previous Income
37-38		Previous Rent/Previous Annual Income
39-40	II(3)	II(2)-II(1) (Signed difference, e.g. -14, in rent/income fractions)
41-43		Additional Housing Expense (Absolute)
44-45		Additional Housing Expense (Annual)/ Annual Income
46-47	III(1)	Total Housing: Facilities Score (Max 22)

48-49	III(2)	Total Housing: Maintenance Score (Max 10)
50	III(3)	Total Housing: Occupancy Score (Max 8)
51-52		Housing Total (Max 40)
53-54	III(1)	Previous Housing: Facilities Score
55-56	III(2)	Previous Housing: Maintenance Score
57	III(3)	Previous Housing: Occ. Score
58-59		Previous Housing Total
60-62		Present-Past Housing Total (Signed Difference)
63	IV(1)	0=No 1=Yes
64	IV(1)	Classify job aspiration for Subject Child by Code for Part I: Occupation Head of Household (Code: Card I, Col. 16) 3=9 or 8 2=7 to 3 1=2 or 1
65		Difference Between Occ. H. of H. and Aspiration for Child: 3=child higher 2=equal 1=child lower
56	IV(2)	Job Expectancy 3=Yes 2=DK 1=No
67	IV(3)	Educational Aspiration for Child (Code: Card I, Col. 18) 3=8, 7, or 6 2=5 or 4 1=3, 2, or 1 Add: 9 Don't Care or Don't Know
68	IV(4)	3=Prob. yes 2=Maybe or DK 1=Prob. no

69	IV(5)	best in class 3= above middle 2= middle 1= get by DK	(Class rank Aspiration).
70	IV(6)	3=Prob. yes 2=Maybe or DK 1=Prob no	(School performance, expectancy)
71	IV(7)	3=yes 2=DK 1=No	(Study aspiration)
72	IV(8)	3=yes 2=DK 1=no	(Study expectancy)
73	IV(9) IV(10)	4=(9)yes, (10)yes 3=(9)yes, (10)no 2=(9)no, (10)yes 1=(9)no, (10)no	(Educational motivation)
74	IV(11)	4=Very important 3=Important 2=Not so important 1=Unimportant	(Educational motivation)
75-76		Total Aspiration Score: Sum Scores in Cols. 54, 55, 57, 59, 61 (range 5-15)	
77-78		Total Expectancy Score: Sum Scores in Cols. 56, 58, 60, 62 (range 3-12)	
79		Total Motivation Score: Sum Cols. 65-66, 67-68 (range 2-8)	

CARD III

1-5		Project Name Code	
6-10		Family Name Code	
11-12	V(1)a	5=VG 4=G 3=NG/NP 2=P 1=VP	Total all present neighborhood facilities (range 14-70)
13-14	V(1)b	5=VG 4=G 3=NG/NP 2=P 1=VP	Total all past neighborhood facilities (range 14-70)
15-17	V(1)a-b	Present Neighborhood - Past Neighborhood (Negative if present poorer, e.g., -23)	
18	V(2)a	5=VG 4=G 3=NG/NP 2=P 1=VP	Average Rating of People this neighborhood (a-e)
19	V(2)b	5=VG 4=G 3=NG/NP 2=P 1=VP	Average rating of people previous neighborhood
20-21	V(2)a-b	Present-Past (Negative if present poorer)	
22-23	V(3)a	No. Project Families Known by Name	
24-25	V(3)b	No. Project Families Called Friends	
26-27	V(3)c	No. Neighborhood Families Known by Name	
28-29	V(3)d	No. Neighborhood Families Called Friends	
30	V(3)b,d,e	1=More friends in project and neighborhood 0=More friends outside	
31		1=More friends now	
32	V(3)f	0=More friends previously	

33-34	V(4)	4=weekly 3=once/month 2=several times/year 1=once/year	Total score group activities (sum a-i)
35-36	V(5)	4=weekly 3=once/month 2=several times/year 1=once/year	Total score places visited (sume a-i)
37-38	V(6)	Total No. of Checks in "Takes Children" column in (4) and (5)	
39-40	V(7)	4=weekly 3=once/month 2=several times/year 1=once/year	Sum independent activities of children
41	V(8)	1=a 2=b 3=c 4=d 5=e 6=f 7=g	
42-43	V(9)	Hours per week subject child is cared for	
44	V(10)-(12)	3=this neighborhood best 2=no difference 1=this neighborhood poorer	
45	V(13)	3=yes 2=DK 1=no	(Aspiration to relocate better)
	VI(1)	6=some time/day 5=several times/week 4=once/week 3=less than once/week 2=rarely 1=never	

46		Time spent with project children (1-6 above)
47		Time spent with neighborhood children (1-6 above)
48		1=36 is greater 0=35 is greater
49-50	VI(2)	No. friends listed for subject child
51-52		No. friends in same school
53		Time spent with project friends
54		Time spent with neighborhood friends
55		Time spent with outside friends
56	VI(3), (4),(5)	4=all e "yes" 3=2 "yes" 2=1 "yes" 1=no "yes"

Score 1-C as in cols: 35&36

(Satisfaction: children's friends)

PARENT INTERVIEW: DEPENDENT VARIABLES

	QUESTION	CARD NO. AND COLUMN
Parents:		
Material Gains:		
Employment	Screening Interview	II 17
Income (spendable)	II(1)(2)	II 39-40
Housing	III(1)-(3)	II 60-62
Community	(Community Evaluation Scale Present-Past Difference Score)	
Social Integration:		
Friendships	V(3)a.-f.	III 22-32
Community Participation	V(4),(5)	III 33-36
Self Concept-Outlook		
Aspirations for Children	IV(1)-(11)	II 75-76
Expectancy for Children	IV(1)-(11)	II 77-78
Subjective Gains:		
Employment	I(3)	II 16
Friendship	V(3)	III 22-32
Community:		
People	V(2)	III 18-21
Facilities	V(1)	III 11-17
Participation	V(4),(5)	III 33-36
Sum Evaluation	V(10)-(12)	III 44
Children:		
Social Integration:		
Friendship	VI(1),(2)	III 46-55
Community Participation	V(6),(7)	III 37-40

Family name ________________
Child's name _______________
Age:___ Sex: ___ Race:___
Address: ___________________

CHILD INTERVIEW

I. FRIENDS

(1) a. About how many kids in this project do you know by name? Can you name them for me? Will you also tell me what school they go to and which grade they are in?

LOCATION AND NAME OF CHILD	SCHOOL	GRADE	check one: BEST FRIEND	FRIEND	TIME SPENT WITH — Outside of school hrs.: Sometime every day	Sometime several days a week	Sometime once a week	Less than once a week	A few times a year	In school only
Project										
Neighborhood										
Outside Neighborhood										

b. Which of the kids you have named would you call your friends? (Fill in table above.)

c. Which of those friends would you call "best friends"?

(2) a. How many kids in this neighborhood, but not living in the project, do you know by name? Can you name them? (Neighborhood is roughly defined as elementary school district.) Do they all go to your school? What grades are they in? (Fill in table above.)

b. Which of these kids would you call friends? (Fill in table above.)

c. Would you call any of them "best friends"? (Fill in table above.)

(3) a. Do any of your friends live outside this neighborhood? Can you name them and tell me what schools they go to? (Fill in table above.)

b. Which of them live in your old neighborhood? (Before you moved here?) (Mark old neighborhood friends with asterisks.) (Fill in table above.)

c. Are any of them "best friends"? (Fill in table above.)

(4) Now I will read off all the kids' names we listed, and I would like you to tell me about how much time you spend with each of them. If you see them only a couple of times a year, tell me that, or every day, week, and so on. (Fill in table above.)

II. COMMUNITY FACILITIES AND PARTICIPATION

(1) Do you belong to any clubs, teams, or other groups that have regular meetings or activities?

GROUP (check)	FREQUENCY OF MEETINGS	REGULAR ATTENDANCE
Scouts		
Church		
YM(W)CA		
Boys Club		
Little League		
Other teams		
School clubs		
Outside clubs		
Other (specify)		

(2) What kinds of places in this neighborhood do you go to? About how often? With whom do you usually go? (Cross out places that do not exist in neighborhood.)

PLACE	FREQUENCY: Daily	Weekly	Few times a month	Few times a year	Never		WITH WHOM: Alone	W/friends	W/parents	W/other adults
Library										
YM(W)CA										
Parks/Playgrounds										
Churches/Synagogues										
Movies										
Bowling alley										
Skating rinks										
Swimming pools										
Clinics										
Special stores (specify)										
Gyms										
Other (specify)										

III. ACTIVITIES AND ASPIRATIONS

(1) a. Do you have any paid part-time jobs? What are they?

JOB	TIME hrs./wk.

b. (If jobs) About how many hours a week do you spend working?

(2) a. What three things would you say you spend the most time doing after school and on weekends?

	TIME SPENT	WITH WHOM
watching television		
organized sports		
unorganized sports		
working (paid jobs)		
visiting friends		
visiting relatives		
studying		
helping around the house		
taking care of younger children		
messing around		
going to movies		
going places (YM(W)CA, libraries, museums, etc.)		
school related activities (clubs, newspaper, etc.)		
hobbies		
Other: ______		
______		
______		

b. About how many hours a week do you spend doing each of these things and with whom? (Fill in above table.)

(3) Can you tell me three things you like doing best after school and on weekends? They might be things other than those we already mentioned--just think about things you like doing best, even if you don't often get to do them.

______

______

______

(4) a. What kind of job would you most like to have when you are grown up?

b. What kind of job do you think you will have when you are grown up?

	MOST WANT	EXPECT
Services (hairdresser, mechanic, etc.)		
Businessman		
Steady work in a plant		
Teacher/Social worker		
Minister (Rabbi, etc.)		
Office worker (clerk, secretary, etc.)		
Professional sports		
Military		
Housewife		
Professional (doctor, lawyer, etc.)		
Own a small store		
Other:		

(5) a. How far would you like to go in school?

b. How far do you think you will go in school?

Advanced degree		
Graduate from four-year college		
Some college (but less than four years)		
Technical, nursing, business course		
Finish high school		
School until age 16		
Would quit now if could		

(6) What are all the subjects you have in school this year?

____

____

____

____

____

____

____

(Make sure all are listed.)

(7) [Interviewer: Fill in names of all English/Language and Math Subjects for each quadruple set of questions which follow.]

a. What grade do you expect to get in __________ on your next report card? (Circle one.)

A A- B+ B B- C+ C C- D+ D D- F

b. What is the lowest grade you could get in __________ on your next report card, that you would be satisfied with?

A A- B+ B B- C+ C C- D+ D D- F

c. How certain on a scale from 1-10, with 10 being most certain, are you that you could get a _____ [Fill in answer to (7)b.] in __________?

1 2 3 4 5 6 7 8 9 10
Uncertain Certain

d. How important is __________ to you, if 10 is very important and 1 is not at all important?

1 2 3 4 5 6 7 8 9 10
Not at all important. Very important.

NEXT SUBJECT

e. What grade do you expect to get in __________ on your next report card? (Circle one.)

A A- B+ B B- C+ C C- D+ D D- F

f. What is the lowest grade you could get in __________ on your next report card, that you would be satisfied with?

A A- B+ B B- C+ C C- D+ D D- F

g. How certain on a scale from 1-10, with 10 being most certain, are you that you could get a _____ [Fill in answer to (7)f.] in __________?

1 2 3 4 5 6 7 8 9 10
Uncertain Certain

h. How important is __________ to you, if 10 is very important and 1 is not at all important?

1 2 3 4 5 6 7 8 9 10
Not at all important. Very important.

NEXT SUBJECT

i. What grade do you expect to get in __________ on your next report card? (Circle one.)

A A- B+ B B- C+ C C- D+ D D- F

j. What is the lowest grade you could get in __________ on your next report card, that you would be satisfied with?

A A- B+ B B- C+ C C- D+ D D- F

k. How certain on a scale from 1-10, with 10 being most certain, are you that you could get a _____ [Fill in answer to (7)j.] in __________?

1 2 3 4 5 6 7 8 9 10
Uncertain — Certain

l. How important is __________ to you, if 10 is very important and 1 is not at all important?

1 2 3 4 5 6 7 8 9 10
Not at all important. — Very important.

NEXT SUBJECT

m. What grade do you expect to get in __________ on your next report card? (Circle one.)

A A- B+ B B- C+ C C- D+ D D- F

n. What is the lowest grade you could get in __________ on your next report card, that you would be satisfied with?

A A- B+ B B- C+ C C- D+ D D- F

o. How certain on a scale from 1-10, with 10 being most certain, are you that you could get a _____ [Fill in answer to (7) n] in __________?

1 2 3 4 5 6 7 8 9 10
Uncertain — Certain

p. How important is __________ to you, if 10 is very important and 1 is not at all important?

1 2 3 4 5 6 7 8 9 10
Not at all important. — Very important.

NEXT SUBJECT

q. What grade do you expect to get in __________ on your next report card? (Circle one.)

A A- B+ B B- C+ C C- D+ D D- F

r. What is the lowest grade you could get in __________ on your next report card, that you would be satisfied with?

A A- B+ B B- C+ C C- D+ D D- F

s. How certain on a scale from 1-10, with 10 being most certain, are you that you could get a _____ [Fill in answer to (7) r.] in __________?

1 2 3 4 5 6 7 8 9 10
Uncertain Certain

t. How important is __________ to you, if 10 is very important and 1 is not at all important?

1 2 3 4 5 6 7 8 9 10
Not at all important. Very important.

NEXT SUBJECT

u. What grade do you expect to get in __________ on your next report card? (Circle one.)

A A- B+ B B- C+ C C- D+ D D- F

v. What is the lowest grade you could get in __________ on your next report card, that you would be satisfied with?

A A- B+ B B- C+ C C- D+ D D- F

w. How certain on a scale from 1-10, with 10 being most certain, are you that you could get a _____ [Fill in answer to (7) v.] in __________?

1 2 3 4 5 6 7 8 9 10
Uncertain Certain

x. How important is __________ to you, if 10 is very important and 1 is not at all important?

1 2 3 4 5 6 7 8 9 10
Not at all important. Very important.

(8) How important are good school grades to you?

	Very important	Important	Not so Important	Unimportant
YOU				
MOTHER				
FATHER				
TEACHER				

How important are good school grades to your mother, your father, your teacher? Fill in above table.

CODE BOOK - CHILD INTERVIEW

CARD IV

Columns	Question	
1-5		Project name code
6-10		Child name code
11-12		Age
13		Sex
14		Race 1=white 2=black 3=Spanish 4=Oriental 5=other
15-16	I(1)-(4)	No. project kids known
17-18		No. project kids called friends (inc. best)
19-20		No. neighborhood kids known
21-22		No. neighborhood kids called friends (inc. best)
23-24		No. outside kids known
25-26		No. outside kids called friends (inc. best)
27		3 = most best friends in project 2 = most best friends in neighborhood 1 = most best friends in outside
28		3 = most time (outside of school) spent with kids in project 2 = most time (outside of school) spent with kids in neighborhood 1 = most time (outside of school) spent with kids outside
29		2 = most friends (inc. best) in same school 1 = most friends (inc. best) in different school

30	II(1)	No. groups belonging to
31		No. groups belonging to - regular attendance
32	(2)	No. community facilities used at least several times per month
33-34		No. community facilities used at least a few times per year
35		1 = facilities mostly used alone 2 = facilities mostly used with friends 3 = facilities mostly used with parents 4 = facilities mostly used with other adults
36	III(1)a.	No. part time jobs
37-38	b.	Work hours per week
39	(2)a.	1 = sports and other purposeful social activity 2 = work (jobs, studying, chores, etc.) 3 = passive activity (movies, TV) 4 = visiting 5 = school rel. activity 6 = educational activity (libraries, museums) 7 = aimless activity ("messing around") 8 = solitary purposeful activities (hobbies, etc.) 9 = other
40	b.	1 = activity primarily solitary 2 = activity primarily in family 3 = activity primarily with friends 4 = activity primarily with other adults
41	(3)	1st preferred activity 2nd preferred activity 3rd preferred activity (Code to be developed from answers obtained)
44	(4)a.	Job aspiration 9=professional, technical 8=managers & proprietors 7=clerical 6=sales 5=service workers

4=craftsmen, foremen
3=operatives (skilled)
2=laborers (unskilled)
1=household

45	III(4)b.	Job expectancy
46		Diff: 3 = aspiration higher 2 = equal asp/exp 1 = exp higher
47	(5)a.	Educ. aspir.
48	b.	Educ. expect.
49		Diff: 3 = aspiration higher 2 = equal asp/exp 1 = exp. higher
50-51	(7)	Average grade expectancy language subjects 12 = A 11 = A- 10 = B+ 9 = B 8 = B- 7 = C+ 6 = C 5 = C- 4 = D+ 3 = D 2 = D- 1 = F
52-53	III(7)	Average lowest grade satisfied with (language)
54-55		Average certainty of satisfaction (language)
56		3 = expectancy higher than satisfaction 2 = expectancy equal to satisfaction 1 = expectancy lower than satisfaction
		Average importance rating (language)
57-58		Average grade expectancy math 12 = A 11 = A- 10 = B+

9 = B
8 = B-
7 = C+
6 = C
5 = C-
4 = D+
3 = D
2 = D-
1 = F

59-60 Average lowest grade satisfied with (math)

61-62 Average certainty of satisfaction (math)

63 3 = expectancy higher than satisfaction
2 = expectancy equal to satisfaction
1 = expectancy lower than satisfaction

Average importance rating (math)

64 III(8) Grade importance: you
4 = very important
3 = important
2 = not so important
1 = unimportant

65 Grade importance: mother
4 = very important
3 = important
2 = not so important
1 = unimportant

66 Grade importance: father
4 = very important
3 = important
2 = not so important
1 = unimportant

67 Grade importance: teacher
4 = very important
3 = important
2 = not so important
1 = unimportant

CHILD INTERVIEW: DEPENDENT VARIABLES

	QUESTION	CARD NO. AND COLUMN
Social Integration:		
Friendship (Peer Relations)	I(1)-(4)	IV 15-29
(Summary Scores)		IV 27-29
Community Participation	II(1),(2)	IV 30-35
Activities	III(1)-(3)	IV 36-43
(Summary Scores)		IV 39,41
Self-Concept-Outlook:		
General: Aspirations	III(4)a., (5)a.	IV 44,47
Expectancies	III(4)b., (5)b.	IV 45,48
Aspiration-Expectancy Differences	III(4)c., (5)c.	IV 46,49
School: Expectancy, Motivation, Satisfaction (Battle)		
Average Expectancy		
(Language)	III(7)	IV 50,51
(Math)	III(7)	IV 57,58
Average Satisfaction		
(Language)	III(7)	IV 52,53
(Math)	III(7)	IV 59,60
Difference		
(Language)	III(7)	IV 56
(Math)	III(7)	IV 63
Confidence		
(Language)		IV 54,55
(Math)		IV 61,62

Self-Concept-Outlook: (continued)

School: (continued)

School Importance	III(8)	IV 64
Perception of School importance to others	III(8)	IV 65-67

APPENDIX B

Recommended Measurement Instruments

BATTLE'S MEASURE OF

ACHIEVEMENT EXPECTANCIES, GOALS, AND VALUES

BEHAVIOR MEASURED: Achievement goals and expectancies.

AGE-GRADE RANGE: Used with Grades 7-9 in two research studies (Battle, 1965, 1966). No reason why it could not be used with elementary school children if they received letter grades in Math and English.

ADMINISTRATION: Group administered written questionnaire. No time given, but very short.

OPERATIONAL DESCRIPTION: Child circles point on a letter grade scale to indicate his Minimal Goal, Minimal Goal Certainty, Expectancy, Absolute Attainment Value, and Relative Attainment Value in Mathematics and English. Each questionnaire also contained filler items.

FORMS: Still a research instrument.

NORMS: None. As any norms would be relative to the grading system of a particular school, standardization would be meaningless.

VALIDITY:

Content. Questions directly tap desired variables in two school subjects, Math and English.

Construct. Minimal goal certainty and expectancy are correlated in 7th to 9th graders (r=0.46, N=74) (Battle, 1965). In larger samples (N=250-500) from same source, attitudes to Math and English were correlated:

Expectancy: $r_p = 0.51$

Minimal Goal: $r_p = 0.62$

Minimal Goal Certainty: $r_p = 0.49$

Absolute Attainment Value: $r_p = 0.47$

(Battle, 1966)

VALIDITY (continued):

Concurrent. In a town of 28,000, a sample of seventy-four seventh, eighth, and ninth graders was drawn; it was distributed in the following way:

20	Minimal Goal	=	Expectancy
20	Minimal Goal	>	Expectancy
20	Minimal Goal	<	Expectancy, Attainment Value High
14	Minimal Goal	<	Expectancy, Attainment Value Low

In this sample, Expectancy correlated 0.47 and Minimal Goal Certainty correlated 0.42 with persistence in an experimental math task. Minimal Goal, Absolute Attainment Value, and Relative Attainment Value did not predict task persistence (Battle, 1965).

Predictive. See Table on following page.

RELIABILITY: Because each construct is measured by a single item, internal consistency and alternate form reliability are irrelevant concepts. Test-retest reliability is irrelevant because test is supposed to be sensitive to what happens in school. Rater reliability is irrelevant because answers are precoded.

SOURCE: Battle, 1965, 1966 (basic items outside of original format)

REFERENCES:

Battle, E.F. Motivational determinants of academic task persistence. Journal of Personality and Social Psychology, 1965, 2, 209-218.

Battle, E.F. Motivational determinants of academic competence, Journal of Personality and Social Psychology, 1966, 4, 634-642.

Correlation of Achievement Attitudes With Grade Attained by Subject Matter and Sex

	Male		Female		Total	
	N	r	N	r	N	r
Mathematics						
Expectancy + grade	245	.74	213	.78	458	.76
Minimal goal + grade	253	.56	220	.66	473	.60
Minimal goal certainty + goal	253	.42	220	.52	473	.46
Attainment value (absolute) + grade	247	.31	213	.34	460	.31
Attainment value (relative) + grade	241	.23	211	.24	452	.20
English						
Expectancy + grade	243	.81	208	.85	451	.84
Minimal goal + grade	252	.67	220	.70	472	.68
Minimal goal certainty + grade	251	.32	219	.39	470	.36
Attainment value (absolute) + grade	246	.28	211	.42	457	.37
Attainment value (relative) + grade	241	.03	210	.09	451	.13

1 All of the variables, except English relative attainment value had independent effects on grades, beyond what is accounted for by IQ. Expectancy of success more powerful determinant than IQ when the two factors are in opposition.

BROWN IDS SELF-CONCEPT REFERENTS TEST

BEHAVIOR MEASURED: Self-concept--own and perception of teacher's image of him or herself.

AGE-GRADE RANGE: 3 1/2- grade 3

ADMINISTRATION: Individual, oral. Test is untimed.

OPERATIONAL DESCRIPTION: Full-length picture (color) taken of child as he stands against off-white background. No posing instructions given. After recognizing him or herself in the picture, the child is asked a standardized set of 15 questions about characteristics of the person in the picture. Questions are in "either-or" format so that bipolar characteristics of various dimensions are represented. For example, "Is (child's name) happy or is he sad?" For children attending school, the same 15 questions are then asked from the teacher's perspective. For example, "Does (child's) teacher think that (child's name) is happy or that he is sad?" At the end, the child is given another picture of himself to take home. (While Brown included "other kids" perception in another set of questions, the Manual [Educational Testing Service, 1969] does not.]

FORMS: One.

NORMS: Sample was composed of 44 children, mean age 5.7 years, from upper-lower to lower-middle class backgrounds. The ethnic composition was Anglo and Chicano. The results for particular items are presented in the Table on the following page.

Another sample consisted of 1851 children participating in the ETS-Headstart Longitudinal Study (Educational Testing Service, 1970). They were tested at age four prior to school attendance. Because children this age have difficulty understanding the difference between self and other referents, only scores referring to the child's perception of self are reported: mean = 10.6, SD = 2.7. (These scores are positively skewed.)

TEST STATISTICS FOR PILOT STUDY OF BROWN IDS
SELF CONCEPT REFERENTS TEST*

Self Referent Item	Percent of Each Choice			
	Child thinks he is			
1	Happy	73%	27%	Sad
2	Clean	80%	20%	Dirty
3	Ugly	4%	96%	Good looking
4	Play with others	78%	22%	Not play
5	Talk a lot	73%	27%	Not talk a lot
6	Own things	29%	71%	Other's
7	Bad	7%	93%	Good
8	Smart	89%	11%	Stupid
9	Scared of a lot of things	18%	82%	Not scared
10	Like clothes	80%	20%	Not like clothes
11	People scared	24%	76%	Not people scared
12	Strong	78%	22%	Weak
13	Sick	18%	82%	Healthy
14	Like face	69%	31%	Not like face
15	Like school	80%	20%	Not like school
Teacher Referent Item				
	Teacher thinks child			
1	is ugly	9%	91%	good looking
2	is bad	4%	96%	good
3	is smart	82%	18%	stupid
4	is strong	71%	29%	weak
5	has lots of friends	78%	22%	not a lot of friends

*Stanford Research Institute, 1971

VALIDITY:

Content. "The test's rationale uses Mead's notion of the evolvement of self-concept from one's perception of significant others' perception of self." (Educational Testing Service, 1970, Vol. 1, p. 177).

Construct. Brown (1966) gave 38 lower class black and 36 middle class white children the same 14 bipolar questions which were repeated for their mothers', teachers', and "other kids'" perceptions of them. In that study and a later replication, Brown found black children's self-perceptions were significantly less favorable than those of white children. Black children perceived their teachers as seeing them in less favorable positions. (No difference between the two groups in either their perceptions of their mother's or peer's evaluation of them.) No race differences appeared in the ETS-Headstart study (Educational Testing Service, 1970).

In the ETS sample, self-concept scores were no different for males and females. There was a significant relationship between age and number of omitted items: the youngest children (42-44 months) omitted an average of 2.10 items; the oldest (57-59 months) an average of 0.42 items. Artifactual variables may have entered into the response to certain questions. For example, subjects had less difficulty responding to items using opposite adjectives than to those using negatives.

RELIABILITY: Internal consistency - In Stanford Research Institute sample described above; $r_{Kr20} = 0.72$.

SOURCE: Educational Testing Service
Princeton, New Jersey

REFERENCES:

Educational Testing Service. Disadvantaged children and their first school experiences: ETS-Headstart Longitudinal Study. Vol. I, Princeton, New Jersey, 1970.

Manual for test administration: Brown IDS Self-Concept Referents Test. Educational Testing Service, Princeton, New Jersey, 1970.

Stanford Research Institute. Report of the pilot studies for the Follow Through evaluation, Fall 1971 test battery. Menlo Park, California, 1971.

SUBTESTS OF THE

CALIFORNIA TEST OF PERSONALITY

(Sense of Personal Worth and Antisocial Tendencies)

BEHAVIOR MEASURED: Mental health--personal and social adjustment.

AGE-GRADE RANGE: Kindergarten to adult.

ADMINISTRATION: Group administered, written test. Maximum of 45 minutes for whole test; untimed.

OPERATIONAL DESCRIPTION: Child answers 8, 12, or 15 (depending on age) yes-no questions for each subtest. 1B-Sense of Personal Worth is part of the personal adjustment battery; 2C-Antisocial Tendencies is part of the social adjustment battery.

FORMS: AA and BB. Different test booklets for grades K-3, grades 4-6, grades 9-college, and adult. (There is also an Intermediate booklet for grades 7-10.)

NORMS: The standardization sample represented every geographic area of the U.S.

Subtest	Age or Grade	Number	Score at 50th Percentile
1B	K-3	4,500	6/8
	4-8	4,562	8/12
	7-10	2,812	12/15
	9-14	3,331	12/15
2C	K-3	4,500	6/8
	4-8	4,562	8/12
	7-10	2,812	12/15
	9-14	3,331	12-13/15

VALIDITY:

Content. "The original sources of the items...were the publications of psychologists and original research by the authors. The specific adjustment patterns which they held to be the best indicators of adjustment (or lack of it) were selected from this literature as criteria of adjustment. However, since these psychologists were not in agreement on many points, five other educational and five other clinical psychologists were given the task of judging the appropriateness and eliminating the least desirable of these criteria," (Thorpe, Clark, and Tiegs, 1953, p. 7).

Test items were devised to correspond to the surviving criteria on each of five age levels.

Some of the subtests have outdated content. The two subtests described here were chosen because both content and correct answers seemed reasonable at the present time.

Construct. The intercorrelations of subtest scores within the Personal Adjustment area range from 0.27 to 0.58 for children in grades 1-3 (N=237). For the Social Adjustment area, the range is from 0.18 to 0.58 (Thorpe, Clark, and Tiegs, 1953).

Concurrent. The California Test of Personality was found to correlate more closely with clinical findings than any other personality test (California Test Bureau, 1949).

In a comparison of the test with three rating methods and an interview technique, the California Test of Personality was most valuable in minimizing personality variance within individuals (Jackson, 1946).

RELIABILITY:

Internal consistency. 1B-Sense of Personal Worth was selected because it has the highest correlations with the total personal adjustment score (0.71-0.82) of any subtest with modern content. Similarly, 2C-Antisocial Tendencies, has the highest correlations with the total social adjustment score (0.66-0.82) of any subtest with modern content. The correlations of the two subtests with each other, are fairly low, ranging from 0.33 to 0.36 (Thorpe, Clark, and Tiegs, 1953).

RELIABILITY (continued):

Alternate forms. "For all levels, the items of each component...have been made equivalent by the process of having each item of Form AA matched with an equivalent item of Form BB as to difficulty, discriminative power, and internal consistency. Thus, the means and standard deviations are identical and the [other] reliability data apply equally to Forms AA and BB." (Thorpe, Clark, and Tiegs, 1953, p. 4).

Test-retest. No information about the time interval between test and retest was given when the following reliability coefficients were determined:

Subtest	Age or Grade	Number	K-R (20) Coefficient
2B	K-3	255	0.82
	4-8	648	0.64
	7-10	1136	0.75
	9-14	2262	0.77
3C	K-3	255	0.82
	4-8	648	0.77
	7-10	1136	0.86
	9-14	2262	0.84

SOURCE: California Test Bureau
Division of McGraw Hill Book Company
Del Monte Research Park
Monterey, California 93940

REFERENCES:

California Test Bureau. Summary of investigations number one, enlarged edition, Los Angeles, 1949.

Jackson, J. The relative effectiveness of paper-pencil test, interview, and ratings as techniques for personality evaluation. Journal of Social Psychology, 1946, 23, 35-54.

Thorpe, L.P., Clark, W.W., and Tiegs, E.W. Manual: California Test of Personality. Monterey, California: CTB/McGraw-Hill, 1953.

COOPERSMITH SELF-ESTEEM INVENTORY

BEHAVIOR MEASURED: Self-concept - self-esteem.

AGE-GRADE RANGE: Age 8 through adult.

ADMINISTRATION: Group, written. Untimed; no time given, but probably about ten minutes for Form A, five minutes for Form B.

OPERATIONAL DESCRIPTION: Form A -- a list of 58 items for which subject checks "like me" or "unlike me". Items broken down into 5 subscales: general self, social self-peers, home-parents, social-academic, lie scale (8 items). Form B -- 25 items having highest item-total score relationship in Form A. Form B has no subscales (Coopersmith, undated).

FORMS: A and B.

NORMS:

Sample	Age or Grade	Mean %	SD
43 girls	Grades 5-6	83.3	16.7
44 boys	Grades 5-6	81.3	12.2
1,748 school children in central Connecticut	Ages 9-15		
Girls		72.2	12.8
Boys		70.1	13.8

(Coopersmith, 1968)

VALIDITY:

Content. Items reflect personal judgments of worthiness in four areas: peers, parents, school, and personal interests. Five psychologists sorted a pool of items into two groups--those indicative of high self-esteem, those indicative of low self-esteem. Repetitious and ambiguous items were thus eliminated.

Construct. Children from higher social class background have significantly higher self-esteem. Jewish children expressed higher self-esteem than Catholic and Protestant (Rosenberg, 1965).

The following findings are based on 85 children from Coopersmith's standardization sample, systematically selected to represent different levels of self-esteem. Children of fathers with irregular work history are most likely to have low self-esteem. There is a significant association of mother's self-esteem and stability and child's subjective self esteem.

> "The mothers of children with high self-esteem are more self reliant and resilient in their attitudes and actions concerning maternity and child care. They are also more likely to accept their roles as mothers and carry them out in a realistic and effective manner.... Fathers of high self-esteem subjects are more likely to be attentive and concerned with their sons and... the sons are more likely to confide in their fathers. The interaction between husband and wife in the families of children with high self-esteem is marked by greater compatibility and ease than is the case in the families of children with less self-esteem. There are more instances of previous marriages and rearing by step-parents in the families of low self-esteem children... From evidence on the decision making process employed within the family, we gain the impression that the high self-esteem families establish clearer patterns of authority and areas of responsibility," (Coopersmith, 1968, p. 116)

> "The most general statement about the antecedents of self-esteem can be given in terms of three conditions: total or nearly total <u>acceptance</u> of the children by their parents, clearly defined and enforced <u>limits</u>, and the respect and latitude for individual action that exist within the defined limits." (Coopersmith, 1968, p. 236)

VALIDITY (continued):

Construct (continued).

"Self-esteem is higher among first and only children... Children high in self-esteem are less likely to have experienced frequent, nonserious problems during the early years of their childhood." (Coopersmith, 1968, p. 163)

Mother's certainty about method of feeding is associated with high self-esteem. Children with high self-esteem are less likely to have been "loners" in their childhood and are more likely to have had good relationships with siblings and peers (Coopersmith, 1968).

Concurrent. The following associations are not based on the calculation of correlation coefficients, but rather, x^2 statistics from contingency tables. The sample is the same as above.

"High self-esteem is associated with high IQ and medium grades. Medium self-esteem is associated with high grades. High self-esteem is also associated with low manifest anxiety, a low level of destructive behavior, and high level of aspiration (beanbag measure)." (Coopersmith, 1968)

RELIABILITY:

Alternate forms. r=0.86

Test-retest. Five week interval; 30 5th grade children, r=0.88. Three year interval; 56 children tested in grades 5 and 8, r=0.70.

SOURCE: Coopersmith, 1968.

REFERENCES:

Coopersmith, S. The antecedents of self-esteem. San Francisco, Freeman, 1968.

Coopersmith, S. Instructions for scoring and interpreting the Self-Esteem Inventory, University of California, Davis, undated.

Rosenberg, M. Society and the Adolescent Self-image. Princeton: Princeton University Press, 1965.

CRANDALL'S INTELLECTUAL ACHIEVEMENT RESPONSIBILITY SCALE

BEHAVIOR MEASURED: Locus of control (external vs. internal control of events) - in the intellectual-academic domain only.

AGE-GRADE RANGE: Grades 3-12

ADMINISTRATION: Individual oral administration for Grades 3-5; written administration in group for Grades 6-12. Untimed; average time not given.

OPERATIONAL DESCRIPTION:

Child is asked to pick the answer that best describes how he or she feels in 34 forced binary choice questions concerning the causation of positive or negative outcomes in the intellectual-academic domain (Crandall, Katkovsky, and Crandall, 1965). For example: "If a teacher passes you to the next grade, would it probably be (a) because she liked you; (b) because of the work you did." Answer (b) is the Internal Control alternative.

FORMS: The following research refers to the original 34-item form. There are, however, two newer short forms (20 items), one for Grades 3-5, one for Grades 6-12. These eliminate items having the lowest item - subscale point biserial correlation (NIMH Progress Report, 1966-1968).

NORMS: The sample was composed of 923 elementary and high school students drawn from five different schools: a consolidated country school, a village school, a small-city school, a medium city school, and a college laboratory school. (See Table , on following page.)

VALIDITY:

Content. For each question, one alternative states that the event was caused by the child and another states that the event was caused by the behavior of someone else in the child's immediate environment. Because there is no evidence that beliefs about locus of control generalize across different behavioral areas, the Intellectual Achievement Responsibility Scale limits items to the single domain of intellectual-academic achievement. The scale was constructed to sample an equal number of positive and negative events, so that separate subscores can be obtained for beliefs in internal responsibility for successes and failures.

Intellectual Achievement Responsibility

Internal Control Scores

Subjects and Grades	Number	Total Score Mean	Total Score SD	Positive Outcome Mean	Positive Outcome SD	Negative Outcome Mean	Negative Outcome SD
Boys:							
3	44	23.16	3.80	12.32	2.02	10.84	3.08
4	59	24.83	3.00	12.41	2.07	12.42	2.08
5	52	24.04	3.69	12.38	2.52	11.65	2.46
6	93	24.74	4.57	12.99	2.54	11.75	2.79
8	68	25.38	3.51	13.07	1.97	12.31	2.23
10	90	25.27	4.62	13.13	2.60	12.13	2.83
12	52	24.38	3.71	11.85	2.83	12.54	1.96
Girls:							
3	58	23.22	4.00	12.88	2.08	10.35	3.01
4	44	24.75	3.81	12.66	2.20	12.04	2.65
5	47	24.36	3.96	12.47	2.54	11.85	2.92
6	73	26.93	3.71	13.88	2.21	13.05	2.43
8	93	26.64	3.86	13.27	2.35	13.38	2.27
10	93	26.50	3.93	13.29	2.22	13.22	2.40
12	57	27.33	2.98	13.40	2.15	13.93	1.94
Total:							
3	102	23.20	3.92	12.64	2.08	10.56	3.05
4	103	24.80	3.37	12.51	2.13	12.26	2.35
5	99	24.19	3.83	12.42	2.53	11.75	2.69
6	166	25.70	4.35	13.38	2.44	12.32	2.72
8	161	26.11	3.77	13.19	2.20	12.92	2.31
10	183	25.90	4.33	13.21	2.41	12.68	2.68
12	109	25.93	3.66	12.66	2.62	13.27	2.07

VALIDITY (continued):

Construct. Correlations between the Intellectual Achievement Responsibility subscores for positive and negative outcomes are positive but low, achieving significance in grades 6, 8, and 10, but not in grades 3-5 or 12. Social class accounts for a much smaller proportion of the variance in Intellectual Achievement Responsibility than it does in other measures of locus of control which sample a wide range of social experiences, not just school-associated activities. Correlations with IQ (Lorge-Thorndike and the California Test of Mental Maturity) were small but statistically significant.

In grades 6-12 Internal Intellectual Achievement Responsibility Scores were significantly associated with being first born in a family and coming from a small family. To assess "social desirability" (the tendency with which children will dissemble in order to put themselves in a socially desirable light), the children's IAR scores were correlated with their scores on the Children's Social Desirability (CSD) Questionnaire (Crandall, Crandall, & Katkovsky, 1964). An absence of relationship between these two measures would suggest that IAR scores are independent of social desirability tendencies. Social desirability tendencies do not account for much of the variance in Intellectual Achievement Responsibility scores (Crandall, et al., 1965).

In a sample of 40 first, second and third graders, girls scored significantly more Internal on the Intellectual Achievement Responsibility scale. There was a correlation of 0.38 in the total sample between this score and the child's intellectual attainment value, that is, the relative importance for the child of intellectual achievement in comparison with other kinds of achievement (Crandall, Katkovsky, and Preston, 1962).

Congruent. In grades 1-3, Stanford-Binet IQ and Intellectual Achievement Responsibility scores were highly correlated for boys but not girls ($r=0.52$, $n=20$).

Concurrent. In grades 3-5 Internal Achievement Responsibility scores correlated positively and significantly with almost all parts of the Iowa Tests of Basic Skills (reading, mathematics, and language subscores and total achievement test scores) and with report card grades. In grades 6-12, Intellectual Achievement Responsibility correlated less strongly but still positively with grades, but most of the correlations with the California Achievement Test were not statistically significant (Crandall, et al., 1965).

VALIDITY (continued):

In 50 grade school children from 6 years, 10 months to 12 years, 5 months, there was a significant relationship between Internality on the Intellectual Achievement Responsibility Scale and speed and correctness on the Witkin Embedded Figures Test, with age partialled out. In a sample of 40 first, second, and third graders, degree of Internal Responsibility for intellectual achievement was correlated with persistence, efficiency (concentration and purposefulness), and conceptual approach (hierarchical, from subunits to whole, vs. trial and error) in solving a 3-D Chinese puzzle. (NIMH Progress Report, 1963-1965.)

Predictive. In first, second and third grade boys, Intellectual Achievement Responsibility scores were highly correlated with the amount of time the boys (n=20) chose to spend in intellectual activities during free time (r=0.70) and the intensity with which they were striving in these activities (r=0.66). They were also highly correlated with boys' reading achievement test scores (r=0.51). These correlations were not significant for girls (Crandall, et al., 1962).[1]

RELIABILITY: Standardization sample (Crandall, Katkovsky, and Crandall, 1965).

Type	Grade	Number	Coefficients: Total Scores	Coefficients: Positive Outcome	Coefficients: Negative Outcome
Test-retest (2 months)	Grades 3-5 Grade 9	47 70	r=0.69 r=0.65	r=0.66 r=0.47	r=0.74 r=0.69
Internal Consistency (odd-even)	Grades 3-5 Grades 6, 8, 10, 12	130 130	r=0.60	r=0.54 r=0.60	r=0.57 r=0.60

SOURCE: Dr. Virginia Crandall
Fels Research Institute
Yellowsprings, Ohio 45387

[1]Although the authors talk about "predicting" from the independent to the dependent variables, the time interval between them is not specified, and they may actually be concurrent measures.

REFERENCES:

Crandall, V.C., Katkovsky, W. and Crandall, V.J. Children's beliefs in their own control of reinforcements in intellectual-academic achievement situations, Child Development, 1965, 36, 91-109.

Crandall, V.J., Katkovsky, W., and Preston, A. Motivational and ability determinants of young children's intellectual achievement behaviors, Child Development, 1962, 643-661.

NIMH Progress Report, Grant No. MHO-2238, National Institute of Mental Health, National Institute of Health, 1963-1965, pp. 110-119 and 1966-1968, pp. 60-67.

ETS STORY SEQUENCE TASK

BEHAVIOR MEASURED: Verbal skills comprehension (Part I) and imitation (Part II) of narrative; visual sequencing (Parts III and IV).

AGE-GRADE RANGE: Ages 4-8 (upper limit is questionable as this has not been established).

ADMINISTRATION: Individual, oral. Untimed; 15 minutes (Part I).

OPERATIONAL DESCRIPTION: The materials consist of six sets of cards with drawings of animals in various situations. There is no apparent sequence in the pictured situations--the sequence is provided by the verbal cues used in the stories. One (Educational Testing Service, 1970) or three (Tanaka, 1968) other sets is used for instructional purposes.

The first part is a comprehension test and requires no verbal response. The child is presented with an array of three or four picture cards, and he is asked to select a sequence of cards to go with a story he is hearing.

In the second part, the child listens to a story and then recalls, i.e., retells, the story.

In the third part, the child produces his own story, selecting and ordering cards from an array (Shipman, 1970). (There is also a fourth part of the test, used by Tanaka, 1968: the child is asked to choose one of three answer choices for the final card needed to finish the sequence. Part IV involves four more sets of cards.)

FORMS: One.

NORMS: See Table on following page.

VALIDITY:

Content. The test has obvious face validity, but no discussion of item selection is available.

Part	Sample	Age or Grade	Mean	SD	Study
I	962 children sampled from school districts near Head Start centers in four geographic regions	42-59 months	4.00*	2.23	Educational Testing Service, 1970
II	28 lower class Head Start children	Age 4	2.71**	1.62	Tanaka, 1968
	35 middle class nursery school children	Age 4	3.77	0.63	Tanaka, 1968
	44 lower class children	Grade K	3.81	0.71	Tanaka, 1968
	77 middle class children	Grade K	3.81	0.58	Tanaka, 1968
III	Lower class Head Start children	Age 4	4%***		Tanaka, 1968
	Middle class nursery school children	Age 4	39%		Tanaka, 1968
	Lower class children	Grade K	22%		Tanaka, 1968
	Middle class children	Grade K	46%		Tanaka, 1968

* out of 7
** out of 4
*** percentage of children correctly completing picture sequence

VALIDITY:

Construct. For Part I there is a small sex difference in mean scores in favor of females (disadvantaged four-year-olds). The great majority of this sample needed help in the instructional items, in sharp contrast to a more advantaged group (Educational Testing Service, 1970).

As for referring to the original sequence in Part II, there are significant age differences between ages four and five, mostly manifest in the low SES group (See Table 7). There were no sex differences. As for time words, middle-class children did better than lower-class, girls did better than boys, and five-year-olds did better than four-year-olds.

RELIABILITY:

Internal consistency. In the ETS-Headstart Longitudinal Study (1971), the overall product-moment correlation coefficient for 1448 children was 0.33 for the two items in Part I.

SOURCE: Dr. Virginia Shipman
Dr. Masako Tanaka
Educational Testing Service
Princeton, New Jersey

REFERENCES:

Educational Testing Service, Disadvantaged children and their first school experiences -- ETS Headstart Longitudinal Study: Preliminary description of the initial sample prior to school enrollment. Vol. I and II, Princeton, New Jersey, 1970. Structure and development of cognitive competencies and styles prior to school entry, 1971.

Tanaka, M.N. Time sequence task. Paper presented at the meeting of the American Educational Research Association, Chicago, Illinois, 1968.

ILLINOIS TEST OF PSYCHOLINGUISTIC ABILITIES

Verbal Expression Subtest

BEHAVIOR MEASURED: Verbal skills -- spontaneous sentence production.

AGE-GRADE RANGE: Ages 2 1/2 to 9.

ADMINISTRATION: Oral, individual. Requires about 5 minutes, untimed.

OPERATIONAL DESCRIPTION: The Verbal Expression Subtest assesses the child's ability to express his own concepts vocally. The child is shown four familiar objects one at a time and is asked "Tell me about this."

The score is the number of discrete relevant and approximately factual concepts expressed. Starting point varies with age. Basal and ceiling levels are determined by performance.

FORMS: One.

NORMS:

Sample	Age	Mean	SD	Study
Approximately 1000 average children between two and ten	2.7-3.1 4.7-5.1 8.7-9.1	4.9 12.9 36.3	2.22 4.74 9.12	Kirk, McCarthy, & Kirk, 1969.
40 upper-lower and lower middle class children	5.7	12.8		Stanford Research Institute 1971

VALIDITY:

Content. "Objects selected were more or less regular in shape, of solid color, without pattern, and homogeneous in composition; this selection allows for a wide variety of characteristics and a minimum of minor details." (McCarthy and Kirk, 1953, p. 9.)

Between 6-0 and 7-0 scores are significantly correlated with social class.

Construct. This subtest correlates 0.82 with the total ITPA score and 0.78 with the grammatic closure subtest (based on 700 children in the 1959-1960 standardization sample, McCarthy and Kirk, 1963).

Congruent. This subtest had a correlation of 0.80 with Stanford-Binet mental age, 0.17 with Stanford-Binet IQ (based on same sample, McCarthy and Kirk, 1963).

RELIABILITY: The following reliabilities are based on the 1959-60 standardization sample, consisting of 700 children between two and nine, randomly selected from those attending school in Decatur, Illinois and their preschool siblings (McCarthy and Kirk, 1963).

Internal consistency. The coefficient of internal consistency is 0.92. The split-half reliability coefficient is 0.91.

Test-retest. After an interval of three months or more, $r_p = 0.37$ for the restricted age range from 6-0 to 6-6 (69 children). The estimate for the full age range is 0.73.

SOURCE: University of Illinois Press
Urbana, Illinois 61801

REFERENCE:

Kirk, S.A., McCarthy, J.J., and Kirk, W.D. Examiner's Manual: Illinois Test of Psycholinguistic Abilities, Revised Edition. University of Illinois Press, Urbana, 1968.

Stanford Research Institute. Report of the pilot studies for the Follow Through evaluation, Fall 1971 test battery. Menlo Park, California, 1971.

ILLINOIS TEST OF PSYCHOLINGUISTIC ABILITIES

Grammatic Closure Subtest

BEHAVIOR MEASURED: Verbal skills--production of grammatical inflections.

AGE-GRADE RANGE: Ages 2 1/2 to 9.

ADMINISTRATION: Individual, oral. Requires about five minutes; untimed.

OPERATIONAL DESCRIPTION: The Grammatic Closure Subtest assesses the child's ability to make use of the redundancies of oral language in acquiring automatic habits for handling syntax and grammatic inflections. There are 33 orally presented items accompanied by pictures which portray the content of the verbal expressions. Each verbal item consists of a complete statement followed by an incomplete statement to be finished by the child. Starting point varies with age. Basal and ceiling levels are determined by age.

FORMS: One.

NORMS: The standardization sample consisted of approximately 1000 average children between two and ten.

Age	Mean	SD
2.7-3.1	2.8	2.17
4.7-5.1	10.2	3.39
8.7-9.1	27.1	3.11

(Kirk, McCarthy, Kirk, 1969)

VALIDITY:

Content. Berko (1958) found that young children could regularly supply currect inflections to nonsense words. Therefore, this ability is located at the non-meaningful level. Nevertheless, the test used pictures of "meaningful"

VALIDITY (continued):

Content (continued).

and familiar objects because children showed a tendency to make nonsense syllables into words. The pictures provide support to younger children but are not thought to be informative.

Construct. At ages 6-6 and 7-0 only, scores are significantly correlated with social class. Scores on this subtest correlate 0.90 with total ITPA score, 0.78 with the verbal expression subtest (based on 700 children in 1959-1960 standardization sample, McCarthy and Kirk, 1963).

Congruent. This subtest had a corrleation of 0.89 with Stanford-Binet mental age, 0.17 with Stanford-Binet IQ (same sample, McCarthy and Kirk, 1963).

RELIABILITY: The following reliabilities are based on the 1959-1960 standardization sample, consisting of 700 children between two and nine randomly selected from those attending school in Decatur, Illinois and their preschool siblings (McCarthy and Kirk, 1963).

Internal consistency. Kuder-Richardson (20) coefficient is 0.93. The split-half reliability coefficient is 0.95.

Test-retest. After an interval of three months or more, r_p=0.72 within the restricted age range from 6-0 to 6-6 (69 children). The estimate for the full age range is 0.92.

SOURCE: University of Illinois Press
Urbana, Illinois 61801

REFERENCE:

Kirk, S.A., McCarthy, J.J., and Kirk, W.D. Examiner's Manual: Illinois Test of Psycholinguistic Abilities, Revised Edition, University of Illinois Press, Urbana, 1968.

McCarthy, J.J. and Kirk, S.A. The construction, standardization, and statistical characteristics of the Illinois Test of Psycholinguistic Abilities, University of Illinois Press, Urbana, 1963.

MATRIX TEST

BEHAVIOR MEASURED: Concept formation - nonverbal classification.

AGE-GRADE RANGE: Ages 4-9.

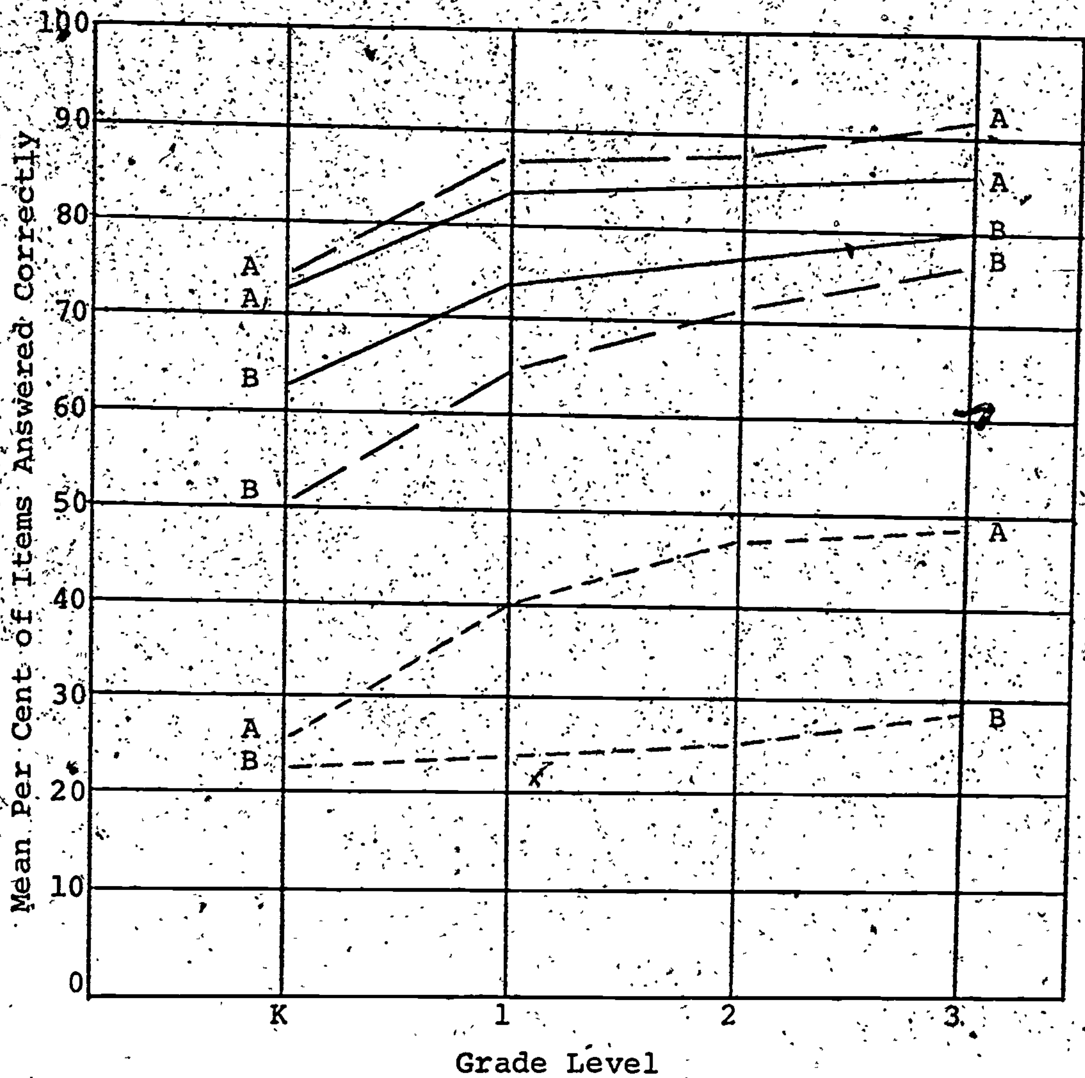
ADMINISTRATION: Individual, oral. Untimed; no time given.

OPERATIONAL DESCRIPTION: The test presents representational as well as abstract items individually on separate cards. Each of 44 items of the test presents a matrix of 2 x 2 or 2 x 3 squares. One square is empty in each case; the others contain two-dimensional geometric figures or pictures of familiar objects which bear some relationship to each other on the basis of their appearance, content, or spatial position in the matrix. The subject is expected to find the figure or pictures for the empty square on the basis of the relationship established by the figures in the other remaining squares. He makes his choice from among four figures or pictures presented alongside the matrix. The subject is asked to point to the figure that he believes to belong in the empty square. The 44 items are divided into four classes of problems. There is no attempt to probe the child's response to any item. (Zimiles, 1968)

FORMS: One.

NORMS: The sample included 320 children. 40 children (20 boys and 20 girls) were tested in kindergarten and each of the first three grades of two public schools. One school was located in a middle class neighborhood and its children came from predominantly white middle class families. All children attending the other school were from lower class black families. The sample was drawn in equal numbers from the upper, middle and lower levels of ability in those grades in which there was homogeneous grouping. The resultant norms are presented in Figure 1 according to age and social class background (Zimiles, 1968).

PER CENT OF ITEMS ANSWERED CORRECTLY AS A FUNCTION OF AGE AND SOCIAL CLASS BACKGROUND



A = School A (advantaged)
B - School B (disadvantaged)

——— Class membership items
— — — One-way classification items
- - - - Two-way classification items

VALIDITY:

Content. Items fall into four classes.

(1) Perceptual matching items (3) present a 2 x 2 matrix in which the figures in all three occupied squares are identical.

(2) Class membership items (18) present a 2 x 2 matrix in which the three occupied squares contain different figures that have a common feature.

(3) One-way classification items (11) present 2 x 2 or 2 x 3 matrices in which all the members of each column or each row are the same. Thus the identity of the missing item is given by its column or row membership.

(4) Two-way classification items (12) present 2 x 3 or 2 x 2 matrices in which row and column membership, in combination, determine the nature of the missing figure (Zimiles, 1968).

Construct. There were no sex differences in the standardization sample (Zimiles, 1968).

Analysis of incorrect responses on the one-way classification task shows that most result from failing to make the transition from horizontal (rows alike) to vertical (columns alike) arrays.

Concurrent. Not available.

RELIABILITY: Reliability data might be contained in Zimiles and Asch, 1967, but Zimiles' office could not locate a copy.

SOURCE: Dr. Herbert Zimiles
Bank Street College of Education
New York, New York

REFERENCES:

Zimiles, H. Classification and inferential thinking in children of different age and social class. Paper presented to the American Psychological Association, San Francisco, 1968.

Zimiles, H. and Asch, H. The development of the Matrix Test. Document 1, Progress Report (1966-1967), Headstart Evaluation and Research Center. Research Division, Bank Street College of Education, New York, December, 1967.

METROPOLITAN ACHIEVEMENT TEST
(1970 Edition)

BEHAVIOR MEASURED: Reading and math achievement.

AGE-GRADE RANGE: Kindergarten through grade 9.

ADMINISTRATION AND OPERATIONAL DESCRIPTION: Group administered, timed, written test. Multiple choice format. The following chart shows areas tested at each grade level and testing time for each.

Summary of Levels and Tests

Level:	Primer		Primary I		Primary II		Elementary		Intermediate		Advanced	
Grades:[a]	*K.7–1.4*		*1.5–2.4*		*2.5–3.4*		*3.5–4.9*		*5.0–6.9*		*7.0–9.5*	
Tests	*Items*	*Time*[b]	*Items*	*Time*	*Items*	*Time*	*Items*	*Time*	*Items*	*Time*	*Items*	*Time*
Word Knowledge			35	15	40	18	50	15	50	15	50	15
Word Analysis[c]	39	20	40	15	35	15						
Reading	33	20	42	30	44	30	45	25	45	25	45	25
Language							50	30	103	50	90	45
Spelling					30	10	40	20	50	15	50	15
Math. Computation					33	18	40	35	40	35	40	35
Math. Concepts[d]	34	20	62	30	40	20	40	25	40	25	40	25
Math. Problem Solving					35	25	35	30	35	25	35	25
Science									78	35	80	35
Social Studies									93	45	94	45

[a] Indicates grades for which each battery is *primarily* intended.
[b] Time in minutes.
[c] Referred to as "Listening for Sounds" at Primer level.
[d] At Primer and Primary I levels, only a Total Mathematics score, based on both concepts and computation, is reported. At the Primer Level, Mathematics is referred to simply as "Numbers."

FORMS: F, G, H

NORMS: Tests standardized on 250,000 pupils in grades K-9 throughout continental U.S. Report No. 8 (Test Department, Harcourt Brace Jovanovich, 1971) presents normative data for every subtest at every age level for every form.

VALIDITY:

Content. On the basis of curricular analysis, items were constructed to reflect what was being taught in the classroom at the time of the 1970 revision (Test Department, Harcourt, Brace, Jovanovich, Report No. 1, 1971).

Construct. Intercorrelations among the three subtests of the Primer battery range from 0.64 to 0.79. For the other batteries they range from 0.60 - 0.89.

Concurrent. Correlations between Metropolitan subtests and the Otis-Lennon Mental Ability Test tend to increase systematically up to grade 5. Above grade 5, the correlations level off in the range of 0.75 - 0.80 for both raw score and deviation IQ (Test Department, Harcourt Brace Jovanovich, Report No. 12, 1971). (Median correlations for each grade are presented in Report No. 10, 1971.)

RELIABILITY:

Internal consistency. Median odd-even split-half reliabilities ranged from 0.91 to 0.94 for nine different age levels (Test Department, Harcourt Brace Jovanovich, 1971).

SOURCE: Harcourt, Brace, Jovanovich, Inc.
757 Third Avenue
New York, N.Y. 10017

REFERENCES:

Test Department, Harcourt Brace Jovanovich, 1970. Report No. 1, Content development; Report No. 8, Summary statistics for national standardization groups; Report No. 10, Reliability estimates and standard errors of measurement; Report No. 11, Intercorrelations among subtests; Report No. 12, Correlations between scores on the Metropolitan Achievement Tests and the Otis-Lennon Mental Ability Test.

NOWICKI-STRICKLAND PERSONAL REACTION SURVEY

BEHAVIOR MEASURED: Locus of control (external vs. internal control of events).

AGE-GRADE RANGE: Grades 1-6, 7-12 (two sets of partially overlapping questions). First and second graders have some difficulty taking test.

ADMINISTRATION: Group administered written test. Questions can be asked orally or in writing. Untimed.

PROCEDURE: Forty forced choice "yes" or "no" questions such as "Are you often blamed for things that just aren't your fault?" "Do you feel that it's easy to get friends to do what you want them to?" (One point is scored for internal control response.)

FORMS: One for each age range.

NORMS: The sample consisted of 1017 mostly white students from four communities bordering on a large southern city (Nowicki and Strickland, 1971).

Internal Orientation Scale

Grade	Males Number	Males Mean	Males SD	Females Number	Females Mean	Females SD
3	44	22.03	4.67	55	22.62	3.06
4	59	21.56	3.58	45	21.20	3.63
5	40	21.68	4.38	41	23.00	4.03
6	45	26.27	5.16	43	26.68	4.58
7	65	26.85	4.87	52	26.06	4.23
8	75	25.27	4.35	34	27.71	3.58
9	43	26.19	4.06	44	27.75	3.75
10	68	26.95	5.34	57	27.02	5.31
11	37	27.52	4.81	53	27.99	5.15
12	39	28.62	4.74	48	27.63	5.05

VALIDITY:

Content. Items sample reinforcement situations across interpersonal and motivational areas such as affiliation, achievement and dependency. Readability at 5th grade level and agreement on classification of answers were also criteria for including items (Nowicki and Strickland, 1971).

VALIDITY (continued)

Construct: There is a small, consistent correlation between degree of Internality and occupational and educational level of the parents. Scores are not correlated with Otis Intelligence (Nowicki and Strickland, 1971). In a socio-metric study of third, fifth, seventh, tenth, and twelfth graders, Internal males received significantly more votes for class president than any other groups, while Internal female twelfth graders were significantly more involved in extracurricular activities (Strickland and Nowicki, 1971). In a camp situation, Nowicki and Barnes (1971) found that, for adolescent black males, Internal scores were related to total votes received for friend and camp president.

Strickland and Hill (1971) found that a group of eight to eleven year old males with severe reading problems were more likely to be External than a matched group of normal readers.

Nowicki and Roundtree (1971) found that ordinal position in family is related to Internal-External scores: first born males and later born females are more Internal.

Nowicki and Barnes (1971) increased Internal Control scores in a group of 291 black lower-class seventh, eighth, and ninth graders by making explicit the connection between the camper's behavior and resultant rewards.

Concurrent. In the standardization sample, the following correlations between locus of control and achievement test scores were found.

Grade	Number	Male	Number	Female
3	34	.284	27	.178
4	50	.118	31	.195
5	42	.398	45	.254
6	33	.272	32	.112
7	35	.335	34	.306
10	49	.442	38	.034
12	38	.451	48	.004

(Nowicki and Strickland, 1971)

In a mostly middle-class white school, Internal students from eight to eleven were significantly more likely to prefer a large reward later to a small reward immediately (delay of reinforcement). This relationship did not hold for black

VALIDITY (continued)

Concurrent (continued)

students in a lower-class school (Strickland and Nowicki, 1971).

In a sample of elementary school children, Duke and Nowicki (1971) found that white External males were significantly more likely than Internal males to distance black peers. The same results held for males and females in a high school sample.

RELIABILITY: Standardization sample (Nowicki and Strickland, 1971)

Type	Age or Grade	Coefficient
Test-retest (6 weeks)	Grade 3 Grade 7 Grade 10	r=0.63 r=0.66 r=0.71
Internal consistency (split half)	Grades 3-5 Grades 6-8 Grades 9-11 Grade 12	r=0.63 r=0.68 r=0.74 r=0.81

SOURCE: Drs. Bonnie R. Strickland and Stephen Nowicki, Emory University, Atlanta, Georgia.

REFERENCES:

Duke, M.P. and Nowicki, S: Perceived interpersonal distance as a function of the subject's locus of control and the race and sex of stimuli in elementary and high school children. Paper presented to the South-eastern Psychological Association, Meanu, Fla., 1971.

Nowicki, S. and Barnes, J. Effects of a structured camp experience on locus of control orientation. Unpublished paper, Emory University, Atlanta, Georgia.

Nowicki, S. and Roundtree, J. Correlates of locus of control in secondary school age students. Developmental Psychology, 1971, 4, 479.

REFERENCES (continued)

Nowicki, S. and Strickland, B. A locus of control scale for children. Paper presented to the American Psychological Association, Washington, D.C. 1971.

Strickland, B.R. and Hill, J. An investigation of some personality variables in male children with severe reading problems. Unpublished paper. Emory University, Atlanta, Ga., 1971.

Strickland, B.R. and Nowicki, S. Behavioral correlates of the Nowicki-Strickland locus of control scale for children. Paper presented to the American Psychological Association, Washington, D.C., 1971.

OHIO SOCIAL ACCEPTANCE SCALE

BEHAVIOR MEASURED: Peer interactions--child's <u>actual</u> acceptance of and by classmates and his <u>perception</u> of classmates' acceptance of him.

AGE-GRADE RANGE: Grades 4-9 (approximate).

ADMINISTRATION: Combination oral-written group administration (Lorber). Untimed; no time given.

OPERATIONAL DESCRIPTION: The test is made up of six numbered paragraphs designed to describe a continuum of relationships from very close association (1) to very definite rejection (6). Each pupil of a group assigns one of these paragraph numbers to every other member of the group (a class list must be provided). The child then assigns to himself the rating he thinks each other classmate would give him.

FORMS: Rath's original form is described above. A Revised Ohio Social Acceptance Scale requires ratings on a five-point scale. (The two most negative points on the original scale have been collapsed into one.) It has not been used to assess the child's perception of how others accept him (Forlano and Wrightstone, 1955; Justman and Wrightstone, 1951).

NORMS:

Number	Grade	% Accepted (1-2)	% Rejected (4-5)
37	7	36	30
37	7	33	36
36	7	46	22
32	8	42	25
32	8	36	37
30	8	40	21

These probably should not be considered norms, as independent observation indicated a tendency for a preponderance of acceptance ratings within a class to be associated with normal or good class spirit and absence of behavior problems.

VALIDITY:

Content. The content of the six paragraphs is based on information obtained from pre-adolescents by asking them to describe activities they liked to do with their very, very best friends, with their other friends, etc. (Raths, 1943).

Construct. A special type of class featuring a democratic, cooperative classroom setting with individual guidance (at the ninth grade level) produced a higher level of overall social acceptability within a classroom than conventional classrooms (Forlano and Wrightstone, 1955).

The test was found to reflect short term changes in student relationships occurring during a trip taken by a high school class (Raths, 1947).

After using the test to identify isolates, Frechs trained their teachers in procedures which might help their social acceptability. Eleven of the twelve made gains on the Ohio Social Acceptance Test administered as a post-test measure (Raths, 1947).

There is good agreement between teacher's judgments of social acceptability and the results of the scale. There is also a relationship between social status and social acceptance as measured by the Scale. In general, the results in a classroom reflect the children's prevailing criteria of social acceptance, e.g., respect for athletic prowess (Raths, 1947).

Concurrent.

Sample	Grade	Criterion	Correlation Coeff.	Study
41 suburban junior high school students	Grade 7	Sociometric measure	$r_s=0.90$	Young, 1947 (unclear which form of scale was used)
		Ohio Recognition Scale (reputation of children as viewed by classmates)	$r_p=0.83$	
5 classrooms	Grade 8	Moreno Sociometric Measure	$r_s=0.89$, 0.81, 0.78, 0.96, 0.89	Justman & Wrightstone, 1951

RELIABILITY:

Test-retest. r=0.72 for acceptance ratings; r=0.77 for perception of acceptance ratings (one-week interval between test and retest).

SOURCE: Raths, 1943.

REFERENCES:

Forlano, J. and Wrightstone, J.W. Measuring the quality of social acceptability within a class. Educational and Psychological Measurement, 1955, 15, 127-136.

Justman, J. and Wrightstone, J.W. A comparison of measuring pupil status in the classroom. Educational and Psychological Measurement, 1951, 11, 362-367.

Lorber, N.L. The Ohio Social Acceptance Scale, 240-243.

Raths, L. Identifying the social acceptance of children. Educational Research Bulletin, 1943, 22, 141-168.

Raths, L. Evidence relating to the validity of the Social Acceptance Test. Educational Research Bulletin, 1947, 26, 141-146, 167-168.

Young, L.L. Sociometric and related techniques for appraising social status in an elementary school. Sociometry 1947, 10, 168-177.

PEABODY PICTURE VOCABULARY TEST

BEHAVIOR MEASURED: Verbal skills--vocabulary comprehension.

AGE-GRADE RANGE: Ages 3-17.

ADMINISTRATION: Individual oral instruction. Untimed; 10-15 minutes.

OPERATIONAL DESCRIPTION: Examiner says word and asks child to point to referent picture (out of four alternatives). (Older children indicate picture by saying number.) Starting points vary with age. Basal and ceiling levels vary with individual performance.

FORMS: A & B (both use same set of 150 plates).

NORMS: The sample consisted of 4,012 white children randomly selected from public school population, Nashville, Tennessee (Dunn, 1965).

Age Levels	Raw Scores			
	Form A		Form B	
	Mean	S. D.	Mean	S. D.
17-6 to 18-5	108.99	14.97	110.57	14.77
16-6 to 17-5	105.11	14.13	106.60	13.98
15-6 to 16-5	102.25	14.27	102.70	13.09
14-6 to 15-5	96.88	13.45	99.10	13.65
13-6 to 14-5	93.36	11.99	95.03	13.33
12-6 to 13-5	90.69	11.07	91.96	12.70
11-6 to 12-5	84.62	10.20	86.23	10.79
10-6 to 11-5	82.23	9.95	83.22	10.90
9-6 to 10-5	75.49	9.61	76.11	9.29
8-6 to 9-5	71.29	9.03	71.52	9.29
7-6 to 8-5	65.81	8.60	65.92	8.69
6-6 to 7-5	60.82	7.77	60.75	7.61
5-6 to 6-5	55.37	7.52	55.20	7.00
4-9 to 5-5	50.22	8.17	49.62	8.05
4-3 to 4-8	45.58	8.04	45.05	7.74
3-9 to 4-2	42.08	10.34	42.02	8.80
3-3 to 3-8	35.67	9.72	35.16	9.45
2-9 to 3-2	29.28	9.66	29.68	8.80
2-3 to 2-8	20.39	8.45	20.23	8.54

VALIDITY:

<u>Content.</u> Item sample drawn from all words in Webster's New Collegiate Dictionary which could be depicted by a picture. Criteria for the selection of the four words to be used in making up any one plate were: (1) all four words were found at the leveling and pre-test stages to be of the same difficulty level; (2) all four words demonstrated good linear growth curves in terms of per cent passing at successive age levels; (3) words were used where no sex differences were found to exist; (4) primarily singular and collective nouns, some gerunds, and a few adjectives and adverbs were used; (5) words were omitted which seemed to be biased culturally, regionally, and racially, as were dated words, plurals, double words, scientific terms, etc.

The illustrations were drawn by an illustrator using the following criteria: (1) equal size; (2) equal intensity; (3) equal appeal; and (4) appropriateness to the age level of subjects most likely to view the plate in the test situation; i.e., in general, earlier plates were of young children, middle numbered plates were of elementary children, and later plates were of adolescents and adults. (Dunn, 1965)

<u>Construct.</u> "Item" validity established by selecting individual words where the per cent of subjects passing increased from one age group to the next. (Dunn, 1965)

Anglo-Saxon (English-speaking) group of 4th, 5th, 6th grade children scored 5.1 IQ points higher than Chicano (Spanish-speaking) group in same grades (Corwin, 1962).

In Grade 5, 18 poor readers had a mean PPVT IQ of 94, 18 average readers had a mean IQ of 97, 18 good readers had a mean IQ of 110.

VALIDITY (continued: Other Types

Type of Validity *	Sample	Age or Grade	Criterion	Coefficient	Study
Congruent.	92 students with IQs from 50 to 57	Age 12-14 Grades 7-9	IQ (WB-F)	$r_p=0.82$	Lavitt 1963
			IQ (WISC-V)	$r_p=0.86$	
			IQ (WISC-P)	$r_p=0.70$	
			IQ (SB)	$r_p=0.92$	
	140 children	Grades K-6	IQ (WISC-F)	$r_p=0.57$	Lindstrom 1961
			IQ (WISC-V)	$r_p=0.67$	
			IQ (WISC-P)	$r_p=0.39$	
			(The mean PPVT IQ averaged 10 points above mean WISC IQ)		
	150 students	Grade 7	Henmon-Nelson Test of Mental Ability	Form A $r_p=0.64$, Form B $r_p=0.61$	Tempero & Ivanoff 1960
			California Test of Mental Maturity	$r_p=0.57$ $r_p=0.58$	
Concurrent.			CAT-Reading & Vocabulary	$r_p=0.56$ $r_p=0.58$	
			CAT-Reading Comprehension	$r_p=0.60$ $r_p=0.68$	
			CAT-Arithmetic Reasoning	$r_p=0.50$ $r_p=0.57$	
			CAT-Arithmetic Fundamentals	$r_p=0.40$ $r_p=0.49$	
			CAT-Language Mechanisms	$r_p=0.44$ $r_p=0.45$	
			CAT-Language Spelling	$r_p=0.41$ $r_p=0.45$	
Predictive.	270 children	PPVT-August before Grade 1. MAT, Primary I following Spring	MAT-word knowledge	$r_p=-.39$	Klaus & Starke 1964
			MAT-word discrimination	$r_p=0.35$	
			MAT-reading	$r_p=0.39$	

TEST ABBREVIATIONS

WB-F = Wechsler-Bellevue - Full scale.
WISC-V = Wechsler Intelligence Scale for Children - Verbal scale
WISC-P = Wechsler Intelligence Scale for Children - Performance scale
SB = Stanford Binet
CAT = California Achievement Test
MAT = Metropolitan Achievement Test

*Tests under <u>congruent</u> are those measuring the same item. Tests listed under <u>concurrent</u> are those with "practical" criterion (achievement tests are supposed to be "practical" nontest criterion).

RELIABILITY:

Standardization Sample.

Age	Number	Alternate Form Pearson's Reliability Coefficients
18	227	0.84
17	305	0.84
16	354	0.80
15	307	0.83
14	269	0.81
13	287	0.70
12	250	0.78
11	384	0.81
10	319	0.77
9	259	0.74
8	100	0.79
7	100	0.74
6	183	0.67
5	133	0.73
4-6	122	0.72
4-0	110	0.77
3-6	119	0.81
3-0	92	0.75
2-6	92	0.76

Other Samples.

Type of Reliability	Sample	Grade	Coefficient	Author & Date
Alternate forms (both given same day)	150 students (group administration)	7th grade	$r_p=0.75$	Tempero & Ivanoff 1960
Alternate forms			$r_p=0.91$	Stanford Research Institute 1971
Test-retest			$r_p=0.81$	"
Internal Consistency			$r_p=0.88$	"

SOURCE: Lloyd M. Dunn
American Guidance Service, Inc.
720 Washington Avenue, S.
Minneapolis, Minnesota 5541

REFERENCES:

Corwin, B.J. The influence of culture and language on performance on individual ability tests. Unpublished paper, Division of Education, San Fernando Valley State College, Northridge, California, 1962.

Dunn, L.M. Peabody Picture Vocabulary Test Manual, Circle Pines, Minnesota: American Guidance Service, 1965.

Klaus, R.A. and Starke, C. Experimental revision of the Peabody Picture Vocabulary Test as a predictor of first grade reading ability. Unpublished paper, Psychology Dept., Peabody College, Nashville, Tennessee, 1964.

Lavitt, J. Comparison of the Peabody, Wechsler, Binet, and California Tests of intellectual ability among 7th to 9th grade pupils. Unpublished paper, Westfield Public Schools, Massachusetts, 1963.

Lindstrom, A.E. A comparison of the Peabody Picture Vocabulary Test and the Wechsler Intelligence Scale for Children, Studies in Minnesota Education, 1961, 131-132.

Stanford Research Institute, Report of the pilot studies for the Follow Through evaluation Fall 1971 battery. Unpublished paper, Menlo Park, California, 1971.

Tempero, H. and Ivanoff, J. Effectiveness of the Peabody Picture Vocabulary Test with seventh grade pupils. Unpublished paper, Dept. of Educational Psychology and Measurement, University of Nebraska, Lincoln, Nebraska, 1960.

PRESCHOOL INVENTORY

BEHAVIOR MEASURED: Achievement in areas regarded as necessary for success in school.

AGE-GRADE RANGE: Age 3-6 (but could be used with older children as six-year-olds pass only 2/3 of the items on the average).

ADMINISTRATION: Individual, oral; 15 minutes, untimed.

OPERATIONAL DESCRIPTION: Responses to some of the 64 items are verbal; others are nonverbal. Items measure "information about the self (e.g., name, age, parts of the body); number concepts ("more" vs. "less"); knowledge of basic sensory attributes (color names); spatial movement with respect to common environmental objects and phenomena ("which way does a ferris wheel go?"; a rudimentary understanding of social roles ("What does a dentist do?"); and the ability to follow simple directions as well as relatively complex directions that presume an understanding of prepositions ("behind," "under," "in," etc.)." (Educational Testing Service, 1970, p. 141.)

FORMS: One.

NORMS: Headstart Sample.

Age	Number	Mean*	SD
3.0 - 3.11	158	25.6	9.8
4.0 - 4.5	528	30.0	10.1
4.6 - 4.11	438	33.9	10.5
5 - 5.5	259	38.4	10.1
5.6 - 6.6	148	42.4	11.0

Headstart is this year using a 32 item version; no data is available yet.

* Out of 64 items.

NORMS (continued):

North Carolina Kindergarten children:

Social Class	Number	Mean	SD
Lower class	82	38.6	10.9
Middle class	136	46.5	9.1
Upper class	99	51.5	8.5

VALIDITY:

Content. Kindergarten curricula and classroom observation provided data on skills children implicitly assumed to possess, in order to function in kindergarten classroom. These skills, as well as areas of deficit in disadvantaged children, were the basis for item construction.

Construct. In the standardization sample the following correlations with Stanford-Binet IQ were obtained:

Age	r
3.0 - 3.11	0.39
4.0 - 4.5	0.59
4.6 - 4.11	0.64
5.0 - 5.5	0.65
5.6 - 6.11	0.63

These are low enough to indicate that the test measures something other than general intelligence.

Among 1875 children between 3.9 and 4.8, Preschool Inventory scores correlated .59 with the Picture Completion and .54 with Form Reproduction scores on the WPPSI (Shipman, 1971).

RELIABILITY:

Internal consistency. The Kuder-Richardson (20) subtest coefficient in the total standardization sample is 0.91.

The Preschool Inventory is not meant to be a homogeneous test, but average biserial correlations between each item and the total score (0.45-0.56) are higher than for most achievement tests covering a broad field of knowledge (Preschool Inventory Handbook, 1970.

SOURCE: Educational Testing Service, Princeton, New Jersey

REFERENCES:

Educational Testing Service, Disadvantaged children and their first school experiences: ETS-Headstart longitudinal study Vols. 1 and 2, Princeton, New Jersey, 1970.

Preschool Inventory Handbook. Educational Testing Service, Princeton, New Jersey, 1970.

Shipman, V.C. Structure and development of cognitive competencies and styles prior to school entry. Educational Testing Service, Princeton, New Jersey, 1970.

SIGEL OBJECT AND PICTURE
CATEGORIZATION TESTS

BEHAVIOR MEASURED: Concept formation.

AGE-GRADE RANGE: Ages 4-6.

ADMINISTRATION: Individual, oral. Untimed; 40-60 minutes.

OPERATIONAL DESCRIPTION: The Object Categorization Test utilizes twelve objects--matches, blocks, top, pipe, cup, notebook, ball, cigarettes, crayons, bottle opener (photograph of objects appears in Sigel and McBane, 1967, Sigel and Olmstead, undated).

The test has three sections: Identification, Active Sort, and Passive Sort. In the Identification part, the child is asked what each object is. In the Active Sort, the tester picks out an object, and asks the child to pick out the ones that are the same or like it. Then he asks "Why are the objects the same?" This procedure is repeated with each object. In the Passive sort, the tester picks out three items and asks the child how they are alike.

FORMS: One.

NORMS: Because no overall scores are derived from this test, norms are very complex. They are found in Tables 7, 8, 8:1, 9, 9:1, 10, 10:3, 10:4 in Sigel and Olmstead (undated).

VALIDITY:

Content. The task utilized an array containing twelve life-sized familiar objects classifiable into four classes, kitchen things, smoking things, toys, and writing things. All items were known and could be labeled or defined in some way.

Construct. Lower-class black and, to a lesser degree, white children, had greater difficulty classifying pictures than objects; and they used different types of category for each representational medium. Middle-class white and black children showed no differences in grouping three- and two-dimensional stimuli (Sigel, Anderson, Shapiro, 1966; Sigel and McBane, 1967; Sigel and Olmsted, undated). A later study (Headstart Evaluation and Research Center, 1970) did not find this discrepancy.

Training programs, modeled after Sigel, Roeper, and Hooper (1966) were successful in influencing quality and quantity of styles of categorization in five-year-old kindergarten children. The effects remained one year later. Training did not, however, influence the discrepancy in categorization performance with objects and pictures (Sigel and Olmsted, 1968). Headstart alone significantly improved scores on both tests in both the Active and Passive Conditions (Headstart Research and Evaluation Center, 1970).

The following pattern of intercorrelations between parts of the two tests was obtained in a sample of 160 Headstart children:

	RPC	AOC	DOC
APC	.56	.73	.61
PPC		.49	.61
AOC			.69

APC = Active Picture Categorization
RPC = Passive Picture Categorization
AOC = Active Object Categorization
POC = Passive Object Categorization

Congruent.
In the same sample, the pattern of correlations with the Stanford Binet and Preschool Inventory was as follows:

	Stanford-Binet	Preschool Inventory
APC	.22	.40
PPC	.28	.34
AOC	.35	.40
POC	.30	.40

RELIABILITY:

Internal Consistency. Measures of internal consistency are not given because Sigel and Olmsted feel that "such a reliability would reflect more the behavior of the subject than the consistency of the test" (p. 27, undated).

RELIABILITY (continued):

Test order does have an effect. Test order interacts with sex and social class of the child.

Test-retest. Grouping responses in Object and Picture Categorization Tests.

	Time Interval	Number	r
Active sort	6 months	81	0.69
Passive sort	8 months	51	0.44

(Sigel and Olmsted, undated)

SOURCE: Sigel and Olmsted, undated.

REFERENCES:

Headstart Evaluation and Research Center, Michigan State University and Merrill-Palmer Institute. A classification and attention training program for Headstart children. Project Report, May 1970.

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

Sigel, I.E. and McBane, B. Cognitive competence and level of symbolization among five-year-old children. In J. Hellmuth (ed.), Disadvantaged Child, Vol. 1. Seattle: Special Child Publications, 1967.

Sigel, I.E. and Olmsted, P.P. Analysis of the object categorization test and the picture categorization test for preschool children. Unpublished paper, Michigan State University Headstart Evaluation and Research Center, undated.

Sigel, I.E. and Olmsted, P.P. Modification of cognitive skills among lower-class Negro children: A follow-up training study. Final Report, Headstart Subcontract #4118 with Michigan State University Headstart Evaluation and Research Center, 1968.

Sigel, I.E., Roeper, A., and Hooper, F.H. A training procedure for acquisition of Piaget's conservation of quantity: A pilot study and its replication. British Journal of Educational Psychology, 1966, 36, 301-311.

STANFORD BINET INTELLIGENCE SCALE

BEHAVIOR MEASURED: General intelligence.

AGE-GRADE RANGE: Age 2 to adult.

ADMINISTRATION: Individual oral administration. The examination of a young child can usually be completed in half an hour to forty minutes; that of an older child frequently requires an hour and a half. There are no timed items before the 10-year-old level.

OPERATIONAL DESCRIPTION:

At every age level some items require verbal responses, others require non-verbal responses. There is a wide range of item types. For instance, the 4-year old items involve vocabulary, memory, and form discrimination. The test begins at an age level where the particular child is likely to succeed, with some effort. For an average child this point would be one year below chronological age. Basal and ceiling levels vary with performance.

FORMS: In the 1937 version there were two forms--L and M. But the 1960 revision changed to a single form, LM.

NORMS: In general the mean Mental Age equals Chronological Age, where each item passed contributes one month to Mental Age below year 6, two months above year 6. But standard deviations are totally variable. It is therefore necessary to use the IQ tables presented by Terman and Merrill (1962) in order to make comparisons with their normative sample. This sample, the 1937 standardization group, consisted of 3184 native-born white subjects, including approximately 100 subjects at each half-year interval from 1 1/2 to 5 1/2 years, 200 at each age from 6 to 14, and 100 at each age from 15 to 18. Each age group was equally divided between males and females and geographically distributed within the U.S. by region and degree of urbanization.

VALIDITY:

Content. The primary criterion for selection of items was to secure an arrangement of subtests which would fulfill "the assumption that general intelligence is a trait which develops with age," (Terman and Merrill, 1962, page 7). Only items which "would probably correlate well with acceptable criteria of intelligence" (page 9) could be included.

This intuitive, culturally-relative notion, whatever it means, is the only source of face validity for the test.

Construct. All items show an increase in percent passing at successive age levels. Only items having the highest correlations with total scores were retained in the 1960 revision, this contributing to the measurement of general intelligence rather than specific abilities (Terman and Merrill, 1962).

The results of McNemar's (1942) factorial analyses of all the items support the view that a single common factor would explain performance, but Jones' (1949) study reveals the operation of group factors. Hofstaetter's (1954) factor analyses of inter-age correlations in Bayley's eighteen-month growth study revealed three factors which account for a child's achievement: "'Sensory-motor Alertness' accounts for the variance of mental age scores for the first two years. Between 2 and 4 years a second factor, 'Persistence' is operative and after 4 years 'Manipulation of Symbols.'" (Terman and Merrill, 1962, page 35).

Kirk (1958) was one of the first investigators to show that ability level can be changed by radical changes in early environment. Many structured educational programs have produced large, temporary changes in Stanford-Binet IQ (e.g., Bereiter and Engelman, 1966; Blank, 1970; Caldwell, in press).

Stanford-Binet IQ's are lowered by an adverse reaction to the failures built into the standard test administration. According to this procedure, the tester begins at a point where the child must exert some effort to succeed. He continues up to that mental age level where the child fails all six tests. Thus, the child experiences increasing failure as the test progresses. Hutt (1947) raised poorly adjusted children's IQs 4.5 points by altering this aspect of the procedure. Well adjusted children's scores were not affected.

Many group differences in Stanford-Binet IQ have been observed. The following table shows the correlation between parental occupation and child's IQ found in one study.

Estimated Average IQ's for Different Occupational Levels (Goodenough, Tyler, 1959)	
Professional	116
Semi-professional, Managerial	112
Clerical, Skilled Trades, Retail	107
Semi-skilled, Minor Clerical	105
Slightly Skilled	98
Rural Owner Farmers	95
Day Laborers	96

VALIDITY (continued)

Congruent. Almost all intelligence tests developed since the Stanford-Binet have used it as a validity criterion. For example, the full scale WISC correlated .82 with the Stanford-Binet in one study (Cronbach, 1960). The correlation with the Peabody Picture Vocabulary Test is presented in the description of that test.

Predictive. The following correlations between IQ in grade 9 (one school) and achievement tests one year later constitute a representative sample of Stanford-Binet validity coefficients.

Reading comprehension	.73	
Reading speed	.43	
English usage	.59	
History	.59	
Biology	.54	
Geometry	.48	(Bond, 1940)

RELIABILITY:

Internal consistency. The mean biserial correlation for all the subtests (items) is .66.

Test-retest. The following table summarizes reliability studies with long intervals between administrations of the test. They assess IQ stability as much as test reliability.

	S-B Revision	Ages	Pearson Correlation Coefficient	Study
111 subjects from Standardization sample	1937	5 & 30	.64	Bradway, 1962
100 adopted children	1916	4 & 13	.58	Skodak, Skeels, 1949
50 subjects from Fels study	1937	6 & 13	.67	Sontag, Baker, Nelson, 1958

SOURCE: Houghton Mifflin Company, 2 Park Street, Boston, Massachusetts 02107

REFERENCES:

Bereiter, C. and Engelmann, S., Teaching disadvantaged children in the preschool, 1966.

Blank, M. A methodology for fostering abstract thinking in deprived children. Bulletin of the Ontario Institute for Studies in Education, 1970.

Bond, E.A. Tenth grade abilities and achievements. New York: Teachers College, Columbia University, 1940.

Bradway, D.P. and Thompson, C.W. Intelligence and adulthood: A 25-year follow-up. Journal of Educational Psychology, 1962, 53, 1-14.

Caldwell, B.M. Educational day care for infants and young children. Boston: Houghton-Mifflin, in press.

Cronbach, L.J. Essentials of psychological testing, New York: Harper and Row, 1960.

Goodenough, F.L. and Tyler, L.E. New York: Appleton-Century-Crofts, 1959.

Hofstaetter, P.R. The changing composition of intelligence: A study of the t-technique, Psychology, 1954, 85, pages 159-164.

Jones, L.F. A Factor analysis of the Stanford-Binet at four age levels, Psychometrika, 1949, 14, 299-331.

Kirk, S.A. Early education of the mentally retarded. Urbana: University of Illinois Press, 1958.

McNemar, Q. The revision of the Stanford-Binet Scale. Boston Houghton-Mifflin, 1942.

Skodak, M., Skeels, H. A final follow-up study of 100 adopted children. Journal of Genetic Psychology, 1949, 75, 85-125.

Sontag, L.W., Baker, C.T., Belson, V.L. Mental growth and personality development: A longitudinal study, Monographs.

Terman, L.W. and Merrill, M.A. Stanford-Binet Intelligence Scale: Manual for the Third Revision Form L-M, Boston, Houghton-Mifflin, 1962.

STEPHENS-DELYS REINFORCEMENT CONTINGENCY INTERVIEW

BEHAVIOR MEASURED: Locus of control (external vs. internal control of events).

AGE-GRADE RANGE: Grades 3-8.

ADMINISTRATION: Individual structured interview. Untimed; mean - 15-20 minutes; range - 10-30+ minutes.

OPERATIONAL DESCRIPTION: 40 open-ended questions, e.g., "What makes mothers happy?", "What makes teachers angry?" Answers are scored internal if they describe reinforcement as contingent either upon the subject's own behavior or upon social rules.

FORMS: One.

NORMS:

Sample	Age or Grade	Mean Internal* Control Scores (SD not given)
14 black Headstart girls	Pre-kindergarten	8.7
15 black Headstart boys		8.1
8 white Headstart girls		13.0
4 white Headstart boys		8.5
17 white middle-class girls		14.9
13 white middle-class boys		11.5
6 white middle-class girls		15.8
9 white middle-class boys		14.4

(Parker and Stephens, 1971)

VALIDITY:

Content. Eight questions are asked about each of five reinforcement agents--self, peers, mothers, fathers, teachers (Delys and Stephens, 1971).

*Out of 40.

VALIDITY (continued):

Construct. Chinese-American school children showed higher Internal Control scores than comparable middle-class Anglo-American 4- and 5-year olds. American Indian children from a Headstart program scored about the same in Internal Control as middle-class white nursery school children. Internal Control Scores consistently related to economic status within ethnic groups. Lower-class black and Appalachian white preschool children show the most extreme External Control expectancies of all groups tested (Stephens and Delys, 1971). Anglo- and Afro-American preschool girls have higher Internal Control scores than boys. Puerto Rican, Mexican, American Indian, and Chinese-American boys have more internal scores (Wang and Stephens, 1971). Regular increase in Internal Control scores with age from 3-8. Spanish-American children had significantly higher Internal Control scores when interviewed by a Puerto Rican interviewer, either in Spanish or English, rather than by an Anglo interviewer. This interviewer ethnicity effect did not show up with black children and a black interviewer (Stephens, 1971).

Congruent. Low but fairly consistent correlation between Internal Control score and Stanford-Binet, Preschool Inventory, and Peabody Picture Vocabulary Test score (Stephens, 1971).

Predictive. Relatively Internal scores associated with increasing school performance over time. Relatively External scores associated with decreasing school performance over time (Stephens, 1971).

RELIABILITY:

Rater reliability. r=0.98 (Delys and Stephens, 1971).

Internal consistency (split-half) and test-retest reliability. Two black Headstart classrooms; r= 0.65 (Delys and Stephens, 1971).

SOURCE: Dr. Mark Stephens, Purdue University, Lafayette, Indiana.

REFERENCES:

Delys, P. and Stephens, M.W. A structured interview procedure for preschool students. Research and clinical uses. Paper presented to the Indiana-Kentucky Psychological Association, Louisville, 1971.

Parker, L. and Stephens, M.W. Locus of control expectancies of preschool age disadvantaged children. Paper presented to the Indiana-Kentucky Psychological Association, Louisville, 1971.

Stephens, M.W. Cognitive and cultural determinants of early IE development. Paper presented to the American Psychological Association, Washington, D.C., 1971.

Stephens, M.W. and Delys, P. Subcultural determinants of locus of control (IE) development. Paper presented to the Midwestern Psychological Association, Detroit, 1971.

Wang, P. and Stephens, M.W. Cultural factors and philosophical aspects of locus of control expectancies. Unpublished paper, Purdue University, 1971.

APPENDIX C

Simulation

APPENDIX C--DESCRIPTION OF THE SIMULATION PROGRAM

Description of the simulation study of the proposed non-experimental design is most conveniently carried out in terms of (1) the generation of test statistics in the cross-sectional (Phases I and II) and longitudinal (Phase III) elements and (2) optimization of the three phase design with respect to critical values of the test statistics. Description of these components is provided in C.1 and C.2 respectively; C.3 provides the basic program listings.

C.1 Generation of Test Statistics

The relative complexity of the simulation is reflected in the somewhat cumbersome notation which follows:

X_{ijkt} ~ Test score of t^{th} subject with program residence category i, program income category j, neighborhood income category k.

The additional (unknown) conditions which attach to the existence of "effects," self-selection of the first kind (entering), and self-selection of the second kind (exiting) are denoted as follows:

$$\ell = \begin{cases} 1 & \text{effects} \\ 0 & \text{no effects} \end{cases}$$

$$m = \begin{cases} 1 & \text{self-selection I} \\ 0 & \text{no self-selection I} \end{cases}$$

$$n = \begin{cases} 1 & \text{self-selection II} \\ 0 & \text{no self-selection II} \end{cases}$$

The existence of effects and self-selection of the first kind are both characterized by adding increments of 0.25σ to the subject's score; σ denotes the population standard deviation of the measure. In the case of effects this is only added to the scores of residents ($i=1$); in the case of self-selection of the first kind it is added to the scores of non-residents ($i=o$). Effects and self-selection are both introduced through the mediating effect of the program variables. Thus when effects exist, for example, they only affect the scores of subjects resident in high/moderate income neighborhoods and/or moderate income programs.

The introduction of effects and self-selection is carried out by changing the mean of the population from which the scores are generated. This population is assumed to be normal with mean 100 and standard deviation 20.[1] The basic relationship for generating simulated scores in the cross-sectional phases is therefore as follows:

$$X_{ijkt} = 0.25\sigma \quad (j+k) \; [i(\ell+m)+(1-i)m] + u_t$$

Test scores generated by means of this equation are then used to compute the F statistics for tests of effects and self-selection of the first kind. This is done for all 4 combinations of the binary variables ℓ and m, and for three different levels of replication (6,9,12). The only difference between first and second phases is that the analysis in Phase I is carried our only for the j, k combinations (1,1) and (0.0); in the second phase all four combinations are used. 500 complete sets of test statistics were generated (including the longitudinal elements) and stored on tape.

---

[1]Normal random numbers are generated through transformation of uniform random numbers (CDC RNZ) by an argument of the central limit theorem.

$$u = \left(\sum_{i=1}^{k} x_i - k/2\right) / \sqrt{k/12} \qquad 0 \leq x_i \leq 1$$

R.W. Hanning, Numerical Methods for Scientists and Engineers, McGraw-Hill, New York, 1962.

The longitudinal phase differs from the cross-sectional phase in two respects. In the first place, the dependent variable is now the difference between the subjects original score and his score during the second wave. In the second place, self-selection of the second kind (correlation between mobility and achievement) is introduced. The methods of introducing these features are now described.

Self-selection of the second kind is introduced by simulating mobility amongst the subject population. A base mobility rate of 40% is assumed for all the programs over the two year period separating survey waves. The individual family's probability of moving is modified however, when self-selection of the second kind is present. Modification of carried out in a way designed to confound program classifications with the dependent variables. Thus subjects with high scores in the baseline survey will have a greater probability of staying in high income programs than in low income programs. Converesly, those with low scores at the baseline will have a higher probability of moving out of high income programs.

The expression relating the probability that a familty moves to its baseline score and the characteristics of the program is given by:

$$\Pr\{move/j,\ k,\ n,\ x_{ijkt}\}$$

$$= p_0 - \rho n(j+k)\ \sqrt{p_0(1-p_0)}\ \cdot\left(\frac{x_{ijkt}-\bar{x}_{ijk}}{\sigma}\right)$$

When n=0 (no self-selection) the probability of moving is simply p_0. When n=1, the probability of moving is additionally dependent on the program classification (j,k), on the baseline score relative to the mean for that group of subjects and on the correlation coefficient (ρ) between the binary random variable (move/not move) and the individual's score. This correlation coefficient was made to be 0.2 in the simulation. Using this method each subject was assigned a uniform random number and was deleted from the sample if this number was below the subject's calculated probability of moving. Variance ratios were then computed to test for the significant association between moving vs. not moving and baseline scores. These require, of course, the more complicated expressions appropriate for analysis of variance in factorial design with unequal cell frequencies.

It then becomes necessary to arrive at gain scores for the remaining subject population. In order to insure correlation between gains and probability of moving, gains are made to depend on past scores as well as on independent variables. Differentiating the second wave score by a prime, the relationship employed was:

$$x'_{ijkt} - x_{ijkt} = 0.25\,\sigma(j+k)\ell + \sqrt{r}\,(x_{ijkt} - \bar{x}_{ijk}) + \varepsilon_t$$

ε_t is a normal random variable with zero mean and variance of 1.3 and r is equal to 0.1. Thes numbers were chosen to insure (1) constancy of population variance, and (2) test-retest correlation of 0.75. This relationship was used to generate gain scores for 500 simulation runs of the longitudinal design. Test statistics for the existence of both effects and self-selection were generated and stored on tape.[1]

---

[1] The listing of the program used to generate all simulated scores and the associated test statistics is presented in Section C.3.1.

C.2 Optimization Methods

The basic simulation experiment described in the last section provides statistics for tests for the existence of effects and of self-selection bias through each phase of the investigation. By running the experiment 500 times at each level of replication, it is possible to arrive at good approximations of the multivariate distribution of these test statistics for each of the eight combinations of effects and self-selection of both kinds. These simulated test statistics can then be used to estimate the probability of failing each of these tests given a specified level of the critical value of the test. By varying these critical values, the associated probabilities of failing (and passing) these tests will vary. As described in Section 3.6.2.1, solution of the design for optimal values of these critical levels was carried out; the basic results were presented in Tables 14 and 15. The structure and methods of obtaining these results are described here.

The objective function in this optimization program is:

(B-C)Pr{Finding an effect and Effect exists}

- C Pr{Finding and effect and No effect exists}

-K.

B and C are real numbers designed to measure the benefits and costs associated with the alternative outcomes; K is the expected cost of the experiment itself, and is computed as the cost of successive phases times the probability of carrying them out.

The complex elements of the problem are the evaluation fo the probabilities. These expressions can be written as:

$$Pr\{E \cap \ell\} = \sum_{m=0}^{1} \sum_{n=0}^{1} Pr\{E/\ell, m, n\} \; Pr\{\ell, m, n\}$$

$$\ell = 0,1$$

The first terms in the simulation on the right are estimated by the relative frequency with which all the tests are passed in the 500 experimental runs for given values of the test statisitcs. The second term is the prior probability associated with the particular combination of effects and self-selection. The eight combinations were each expected to be equally likely for the purposes of optimization.

The actual optimization was carried out using a search method, since the approxiamtion to the objective is a grid method and convergence was obtained within a maximum of 150 iterations.[1]

---

[1] Hooke, R. and Jeeves, T.A. "Direct Search" Solution of Numerical and Statistical Problems. Journal of the ACM, Vol. 8, No. 2 (April, 1961), pp. 212-229.

C.3 Simulation

1. Simulation Program

```
                PROGRAM  STAT  (INPUT,OUTPUT,TAPE1,TAPE5,TAPE39,TAPE85,PUNCH)
           C    NOFT = THE NUMBER OF TIMES A DIFFERENT POPULATION IS TESTED.
           C    M = THE NUMBER OF SUBJECTS IN EACH CELL BEFORE DELETIONS.
           C    AA(I,J,K,) = THE MATRIX WHICH HOLDS ALL SUBJECTS+ DATA.
           C              THE 1ST INDEX IS THE NUMBER IN CELL I.
           C              THE 2ND INDEX IS THE CELL NUMBER, THAT IS 1-16.
           C              THE 3RD INDEX IS THE DIFFERENCE BETWEEN MEANS SPECIFICATION.
           C    XDIFF(I) = THE DIFFERENCES BETWEEN MEANS OF INTEREST TO EXPERIMENTER
           C    NUMBER OF SUBJECTS = 16*M OR THE ADJUSTED M FOR EACH CELL.
           C    XMEAN = THE MEAN OF THE SUBJECT POPULATION BEING TESTED.
           C    SDEV = THE STANDARD DEVIATION OF THE SUBJECT POPULATION BEING TESTED.
           C    MMIN = THE MINIMUM NUMBER OF SUBJECTS USED PER CELL.
           C    MMAX = THE MAXIMUM NUMBER OF SUBJECTS USED PER CELL.
           C    INTEV = THE INTERVAL BETWEEN NUMBER OF SUBJECTS USED IN A CELL.
           C    NOCU = THE NUMBER OF CELLS USED.
           C    NOMS = THE NUMBER OF DIFFERENT CELL SIZES USED:  1+((MMAX-MMIN)/INTEV).
           C    NOOS = THE NUMBER OF DIFFERENCES BETWEEN MEANS BEING TESTED.
           C    SC(I,J) = THE SUM OF INDIVIDUALS IN CELL I,J.
           C              THE 1ST INDEX IS THE NUMBER OF THE CELL BEING USED.
           C              THE 2ND INDEX IF THE DIFFERENCE BETWEEN MEANS SPECIFICATION.
           C    SSC(I,J) = THE SUMED SQUARES OF INDIVIDUALS IN CELL I,J.
           C               THE 1ST INDEX IS THE NUMBER  OF THE CELL BEING USED.
           C               THE 2ND INDEX IS  THE DIFFERENCE BETWEEN MEANS SPECIFICATION.
           C    SCS(I) = THE SUM OF INDIVIDUALS IN CELL I SQUARED.
           C    X(I) = SUMS OF NUMBERS USED IN CALCULATING F-RATIOS.
           C    FCRIT = THE F FOR A SPECIFIC CONFIDENCE LEVEL DETERMINED BY THE
           C            NUMBER OF DEGREES OF FREEDOM IN THE F-TEST.
           C    XNBF(I,J,K) = THE MATRIX WHICH HOLDS ALL F-RATIOS IN INTERVALS OF .5.
           C              THE 1ST INDEX IS THE INTERVAL SPECIFICATION, EX:  3.5-<4.0.
           C              THE 2ND INDEX IS THE POP.  DIFFERENCE BETWEEN MEANS INDICATOR
           C              THE 3RD INDEX IS THE NUMBER OF SUBJECTS PER CELL INDICATOR.
           C    XNBFP(I) = THE PERCENT OF F-RATIOS IN A SPECIFIC INTERVAL.
           C    XNBFPC(I) = THE CUMULATIVE PERCENTAGE OF F-RATIOS BP TO CELL I.
000003          DIMENSION  AAI(12,16,8),AAII(12,32,8),AAIII(12,16),AAIIIN(12,16)
000003          DIMENSION  AAIIIM(12,16)
000003          DIMENSION  AAFAMV(16),AASSI(16),AAE(16),ASSIDF(8),AEDIFF(8)
000003          DIMENSION SC1(16,8),SCS1(16),SSC1(16,8),SC2(32),SCS2(32),SSC2(32)
000003          DIMENSION  SC3S2(32),SCS3S2(32),SSC3S2(32),SC3S2D(32)
000003          DIMENSION  SC3E(16),SCS3E(16),SSC3E(16),SC3ED(16)
000003          DIMENSION  XX(12),X(12),X1(6),MMOVCT(16),MNTMCT(16)
000003          DIMENSION  PROBM(3), ROE(3)
000003          DIMENSION  FSE123( 9,24)
000003          DIMENSION  FCRIT(9), XNPFT(9)
000003          DIMENSION  ARAND(500)
000003          DATA    ISTORE      /0/
000003          DATA  NOFT, SDEV, XMEAN              /  500,  20.0, 100.0/
000003          DATA  NOCU1,NOCU2,NOCU3S,NOCU3E,NOMS,NOOS /16,32,32,16,3,8/
000003          DATA  MMIN1,MMIN2,MMAX1,MMAX2,INTEV1,INTEV2  /6,6,12,12,3,3/
000003          DATA  NOCP1T, NOCP2T  /8,16/
000003          DATA  PROBM, ROE  /.4,.4,.4,        -.2,-.1, .0/
000003          DATA  ASSIDF          /0.0,1.0,0.0,1.0,0.0,1.0,0.0,1.0/
000003          DATA  AEDIFF      /1.0,1.0,1.0,1.0,0.0,0.0,0.0,0.0/
000003          DATA JAKPSES              / 0.1/
000003          DATA FCRIT, XNPFT  /9*3.84,9*0.0/
000003          XS = RANF(5)
           C    READ IN MEAN EFFECTS DATA.
000006          READ 450, AAFAMV, AASSI, AAE
```

```
000017     450 FORMAT(3(16F5.2/))
          C    CALCULATION OF MEAN EFFECTS FOR FAMILY, SSI, AND DEVELOPMENTAL VARIABLES.
000017         FM= 0.0 $ EM = 0.0 $ SSIM = 0.0 $ J = 0
000023         DO 401 I = 1, 4
000024         FM = AAFANV(I) + FM
000026         SSIM = AASSI(I+J) +                     SSIM
000031         EM = AAE(I+J) +                 EM
000034         J = J+3
000035     401 CONTINUE
000037         FM = FM/4.0 $ SSIM = SSIM/4.0 $ EM = EM/4.0
          C    START OF A(ONE) TEST RUN.
000042         IT= ITIMER(0)
000045         PRINT 521, IT
000052     521 FORMAT(5X,O20)
000052         DO 1000  ISIM = 1, NOFT
          C    PHASE II CREATION ROUTINE.
000054         IINDIC = 0
000055         MDFS = 0
000056         00 303 NOC = 1, NOCU2
000057         DO 301 IRANN = 1,MMAX2
000060         PAA = 0.0
000061         00 302 I = 1,12
000063         RANON = RANF(0)
000065         PAA = RANON + PAA
000067     302 CONTINUE
000071         ISTORE = ISTORE + 1
000072         ARAND(ISTORE) = PAA
000074         IF(ISTORE .LT. 500)  GO TO  888
000076         ISTORE = 0
000076         WRITE(1)  ARAND
000103     888 IF(MDFS .EQ. 1)  GO TO 311
000105         IF(IINDIC .EQ. 1) GO TO 311
000107         DO 800 MDIFF = 1,NODS
          C    MEAN ADJUSTMENT ROUTINE FOR FAMILY,SSI, AND DEVELOPMENTAL EFFECT.
000110         IF(NOC .GT. 16)  GO TO 311
000113         AAII(IRANN,NOC,MDIFF)=((PAA-6.0)*SDEV) + XMEAN + (AAFAMV(NOC)*SDEV
              1) + (AASSI(NOC)*SDEV*ASSIDF(MDIFF))
000133         GO  TO 800
000134     311 AAH=((PAA-6.0)*SDEV)+XMEAN+(AAFAMV(NOC-16)*SDEV)+(AASSI(NOC-16)*SD
              1EV*ASSIDF(MDIFF))+(AAF(NOC-16)*SDEV*AEDIFF(MDIFF))
000152         IF(MDIFF .LE. 2)  GO TO 313
000154         IF(MDFS .EQ. 1)  GO TO 306
000156         IF(MDIFF .EQ. 5)  GO TO 305
000160         IF(MDIFF .EQ. 6)  GO TO 313
          C    SELF-SELECTION ROUTINE (TYPE II) FOR PHASES I AND II.
          C  CHECKING ROUTINE FOR S.S. TYPE2 MAGNITUDE  PHASE II.
000161         IF(NOC .GE. 17  .AND. NOC .LE. 20)  GO TO 317
000172         IF(NOC .GE. 21  .AND. NOC .LE. 28)  GO TO 316
000201         IF(NOC .GE. 29  .AND. NOC .LE. 32)  GO TO 315
000210     316 I = 2  $  GO TO 318
000212     315 I = 3  $  GO TO 318
000214     317 I = 1
000215     318 AAH=XMEAN+FM*SDEV+SSIM*SDEV*ASSIDF(MDIFF)+EM*SDEV*AEDIFF(MDIFF)
000226         WEIGHT = ABS(ROE(I))* (1.0 - ABS(ROE(I)))
000232         P=PROBM(I)+((ROE(I)*SQRT(  WEIGHT           )*(AAH-AAM))/SDEV)
000243         RAN = RANF(0)
000245         IF(RAN .GT. P)  GO TO 313
000250         IRANN = IRANN - 1
```

```
000251         IINDIC = 1
000252         GO TO   301
000253     305 IRANN = IRANN - 1
000255         MDF5 = 1
000256         GO TO 301
000256     306 MDF5 = 0
000257     313 AAII(IRANN,NOC,MDIFF)=AAH
000265     800 CONTINUE
000270         IINDIC = 0
000271     301 CONTINUE
000274     303 CONTINUE
        C      PHASE I CREATION ROUTINE.
000276         JSI = 0 $ LL1 = 1
000300         DO 30 J = LL1, NOCU2
000301         DO 50 I = 1, MMAX1
000302         DO 40 K = 1, NODS
000303         AAI(I,J+JSI,K) = AAII(I,J,K)
000317      40 CONTINUE
000322      50 CONTINUE
000324         GO TO 35
000325      38 LL1 = 13 $ J = 12 $ GO TO 30
000330      42 LL1 = 29 $ JSI = -16 $ J = 28 $ GO TO 30
000334      35 IF(J .LT. 4)  GO TO 30
000337         IF(J .EQ. 4 )  GO TO 37
000340         IF(J .GE. 17 .AND. J .LT. 20)  GO TO 30
000347         IF(J .EQ. 20)  GO TO 38
000350         IF(J .GE. 13 .AND. J .LT. 16)  GO TO 30
000357         IF(J .EQ. 16)  GO TO 42
000360         IF(J .GE. 29)  GO TO 30
000362      37 LL1 = 17 $ JSI = - 8 $ J = 16 $ GO TO 30
000366      30 CONTINUE
        C      INITIALIZATION FOR TESTING ROUTINES.
000371         MS1= 1
000372         MMM = 0
000373         DO 250 I = 1,NOCU1
000374         DO 250 J = 1, NODS
000375         SC1(I,J) = 0.0
000400         SSC1(I,J) = 0.0
000402     250 CONTINUE
        C      MAIN TESTING LOOP - ROUTINES FOR PHASES I, II, AND III.
        C      PHASE I TESTING ROUTINE.
000406         DO 700 MT = MMIN1, MMAX1, INTEV1
000410         MMM = MMM + 1
000412         DO 850 MDIFF = 1,NODS
000413         IPLACE = (NODS*(MMM-1)) + MDIFF
000420         DO 204 I = 1, NOCU1
000421         SCS1(I) = 0.0
000422     204 CONTINUE
000424         DO 201 IJ = 1, NOCU1
000426         DO 202 IT = MS1, M1
000430         SC1(IJ,MDIFF) = SC1(IJ,MDIFF) + AAI(IT,IJ,MDIFF)
000437         SSC1(IJ,MDIFF) =SSC1(IJ,MDIFF)+(AAI(IT,IJ,MDIFF)*AAI(IT,IJ,MDIFF))
000446     202 CONTINUE
000450         SCS1(IJ)  =  SC1(IJ,MDIFF)*SC1(IJ,MDIFF)
000454     201 CONTINUE
000456         DO 252 I = 1,8
000457         XX(I) = 0.0
000460         X(I) = 0.0
```

```
000461    252 CONTINUE
000463        J = 0
000464        DO 205 I = 1, 4
000465        DO 203 K = 1,4
000466        XX(I) = XX(I) + SC1(J+K,MDIFF)
000475    203 CONTINUE
000500        J = J+4
000501    205 CONTINUE
000503        J = 0
000504        DO 215 I = 2,3
000505        X(I) = XX(I-1+J) + XX(I+J)
000512        J = J+1
000513    215 CONTINUE
000515        DO 216 I = 1, NOCPIT
000516        X(4) = X(4) + SSC1(I,MDIFF)
000522        X(5) = X(5) + SCS1(I)
000524        X(6) = X(6) + SSC1(I+8,MDIFF)
000527        X(7) = X(7) + SCS1(I+8)
000531    216 CONTINUE
000533        K = 0
000534        DO 217 J = 1,2
000535        I = J+K
000537        X1(J) =  (XX(I)*(XX(I)) +  XX(I+1)*(XX(I+1)))/((NOCPIT/2)*M1)
000550        K = K+1
000552    217 CONTINUE
000554        K=3
000555        DO 218 I=1,2
000556        FSE123(I,IPLACE)       = ((X1(I)-(((X(I+1))*X(I+1))/(NOCPIT*M1))))*(
             1NOCPIT*(M1-1)))/(X(K+I)-(X(K+I+1)/M1))
000602        K=K+1
000603        IF(MDIFF .NE. 5) GO TO 218
000605        IF(FSE123(I,5) .GT. FCRIT(I))  GO TO 218
000611        XNPFT(I)  =  XNPFT(I) + 1.0
000613    218 CONTINUE
        C     PHASE II TESTING ROUTINE.
000615        M2 = M1
000617        DO 811 I = 1,NOCU2
000620        SC2(I) = 0.0
000621        SSC2(I) = 0.0
000622    811 CONTINUE
000624        DO 812 IJ = 1, NOCU2
000626        DO 813 IT = 1  ,M2
000627        SC2(IJ) = SC2(IJ) + AAII(IT,IJ,MDIFF)
000636        SSC2(IJ) = SSC2(IJ) + (AAII(IT,IJ,MDIFF)*AAII(IT,IJ,MDIFF))
000644    813 CONTINUE
000646        SCS2(IJ) = SC2(IJ)*SC2(IJ)
000650    812 CONTINUE
000652        DO 814 I = 1,12
000654        XX(I) = 0.0
000655        X(I) = 0.0
000656    814 CONTINUE
000660        DO 815 I = 1, 4
000661        XX(1) = XX(1) + SC2(I) +SC2(I+4)
000664        XX(2) = XX(2) + SC2(I+ 8) + SC2(I+12)
000667        XX(3) = XX(3) + SC2(I) +   SC2(I+ 8)
000672        XX(4) = XX(4) + SC2(I+ 4) + SC2(I+12)
000675        XX(5) = XX(5) + SC2(I+16 )+ SC2(I+20)
000700        XX(6) = XX(6) + SC2(I+24) + SC2(I+28)
```

```
000703           XX(7) = XX(7) + SC2(I+16) + SC2(I+24)
000706           XX(8) = XX(8) + SC2(I+20) + SC2(I+28)
000711       815 CONTINUE
000713           J = 0
000714           DO 820 I = 2,5
000715           X(I) = XX(I-1+J) + XX(I+J)
000722           J = J+1
000723       820 CONTINUE
000725           DO 816 I = 1, NOCP2T
000726           X(8) = X(8) + SSC2(I)
000730           X(9) = X(9) + SCS2(I)
000732           X(10) = X(10)+ SSC2(I+NOCP2T)
000735           X(11) = X(11)+ SCS2(I+NOCP2T)
000740       816 CONTINUE
000741           K = 0
000742           DO 817 J = 1,4
000743           I = J +K
000745           X1(J) =  (XX(I)*(XX(I)) +  XX(I+1)*(XX(I+1)))/((NOCP2T/2)*M2)
000756           K = K+1
000760       817 CONTINUE
000762           DO 818 K = 3,4
000763           I = K-2
000765           FSE123(K,IPLACE)       = ((X1(I) - (((X(I+1))*X(I+1))/(NOCP2T*M2)))
                1*(NOCP2T*(M2-1)))/(X(8)-(X(9)/M2))
001006           L = K + 2
001010           J = L-2
001011           FSE123(L,IPLACE)       = ((X1(J) - (((X(J+1))*X(J+1))/(NOCP2T*M2)))
                1*(NOCP2T*(M2-1)))/(X(10)-(X(11)/M2))
001032           IF(MDIFF .NE. 5) GO TO 818
001034           IF(FSE123(K,5) .GT. FCRIT(K))  GO TO 909
001040           XNPFT(K) = XNPFT(K) + 1.0
001042       909 IF(FSE123(L,5) .GT. FCRIT(L))  GO TO 818
001046           XNPFT(L) = XNPFT(L) + 1.0
001050       818 CONTINUE
            C    PHASE III CREATION ROUTINE.
001052           DO 912 I = 1, NOCU3E
001054           MMOVCT (I) = 0
001055           MNTMCT(I) = 0
001056           DO 915 J = 1, MMAX2
001060           AAIII(J,I)  = 0.0
001063           AAIIIM(J,I)    = 0.0
001066           AAIIIN(J,I)    = 0.0
001071       915 CONTINUE
001073       912 CONTINUE
001076           AAIIM =XMEAN+FM*SDEV+SSIM*SDEV*ASSIDF(MDIFF)+EM*SDEV*AEDIFF(MDIFF)
001107           DO 911 J = 1, NOCU3E
001110           DO 911 I = 1,M2
001111           ADIF3E = AAII(I,J,MDIFF) - AAIIM
001117           RINCR = RANF(0)
001122           AAIIH        =(SQRT(ABS(AKP3ES))*ADIF3E) + ((15.5*RINCR)-7.75)
                1 + (AAE(J)*SDEV*AEDIFF(MDIFF))+AAII(I,J,MDIFF)
001145           IF(MDIFF .EQ. 3 .OR. MDIFF .EQ. 4)  GO TO 913
001154           IF(MDIFF .EQ. 7 .OR. MDIFF .EQ. 8)  GO TO 913
001162           PM = RANF(0)
001164           IF(PM .GT. PROBM(1))  GO TO 918
001167           IF(MMOVCT(J) .EQ. (M2-1))  GO TO  918
001172       920 MMOVCT(J)  =  MMOVCT(J) + 1
001175           KH = MMOVCT(J)
```

```
001176        AAIIIM(KH      ,J) = AAII(I,J,MDIFF)
001205        GO TO 911
001206    918 IF(MNTMCT(J) .EQ. (M2-1))  GO TO 920
001212        MNTMCT(J) = MNTMCT(J) + 1
001213        KH = MNTMCT(J)
001215        AAIII(KH,J) = (AAIIH - AAII(I,J ,MDIFF))
001225        AAIIIN(KH,J) = AAII(I,J,MDIFF)
001232        GO TO 911
       C      PHASE III DEVELOPMENTAL SELF-SELECTION ROUTINE.
001233    913 RAN = RANF(0)
       C    CHECKING ROUTINE FOR S.S. TYPE2 MAGNITUDE  PHASE III.
001236        IF(NOC .GE. 17  .AND. NOC .LE. 20)  GO TO 921
001247        IF(NOC .GE. 21  .AND. NOC .LE. 28)  GO TO 922
001256        IF(NOC .GE. 29  .AND. NOC .LE. 32)  GO TO 923
001265    922 K = 2  $  GO TO 916
001267    923 K = 3  $  GO TO 916
001271    921 K = 1
001272    916 WEIGHT = ABS(ROE(K))* (1.0 - ABS(ROE(K)))
001276        P = PROBM(K) + ((ROE(K)*SORT(WEIGHT)*(AAIIH-AAIIM))/SDEV)
001310        IF(RAN .GT. P)  GO TO 918
001313        IF(MMOVCT(J) .EQ. (M2-1))  GO TO 918
001316        GO TO 920
001317    911 CONTINUE
       C    PHASE III TESTING ROUTINE
       C    PHASE III S.S. TYPE2  TESTING ROUTINE
001324        DO 490 I = 1, NOCU3S
001326        SC3S2(I) = 0.0
001327        SCS3S2(I) = 0.0
001330        SSC3S2(I) = 0.0
001331    490 CONTINUE
001333        ICMAX = NOCU3S/2
001334        DO 491 IJ = 1,ICMAX
001336        NHM = MMOVCT(IJ)
001340        NHNTM = MNTMCT(IJ)
001341        IF(NHNTM .EQ. M2-NHM)  GO TO 444
001344        PRINT 555, MMOVCT(IJ), MNTMCT(IJ),IJ
001355    555 FORMAT (1X*ERROR*3I6)
001355    444 DO 492 IT = 1, NHM
001357        SC3S2(IJ) = SC3S2(IJ) + AAIIIM(IT,IJ)
001364        SSC3S2(IJ) = SSC3S2(IJ) + AAIIIM(IT,IJ)*AAIIIM(IT,IJ)
001370    492 CONTINUE
001372        SC3S20(IJ) = SC3S2(IJ)  / MMOVCT(IJ)
001375        SCS3S2(IJ) = (SC3S2(IJ)*SC3S2(IJ))/MMOVCT(IJ)
001400        DO 493 IT = 1, NHNTM
001402        SC3S2(IJ+16) =  SC3S2(IJ+16) + AAIIIN(IT,IJ)
001407        SSC3S2(IJ+16) = SSC3S2(IJ+16) + AAIIIN(IT,IJ)*AAIIIN(IT,IJ)
001413    493 CONTINUE
001415        SC3S20(IJ+16) = SC3S2(IJ+16) / MNTMCT(IJ)
001420        SCS3S2(IJ+16) = (SC3S2(IJ+16)*SC3S2(IJ+16))/MNTMCT(IJ)
001423    491 CONTINUE
001426        DO 494 I = 1, 5
001427        XX(I) = 0.0
001430        X(I) = 0.0
001431    494 CONTINUE
001433        DO 495 I = 1,ICMAX
001434        X(1) = X(1) + (SC3S2D(I))/16.0
001440        XX(1) = XX(1) + (1.0/MMOVCT(I))/ 256.0
001444        X(2) = X(2) + (SC3S2D(I+16))/16.0
```

```
001450          XX(2) = XX(2) + (1.0/MNTMCT(I))/ 256.0
001454      495 CONTINUE
001457          XX(1) = 1.0/XX(1)  $  XX(2) = 1.0/XX(2)
001461          X(3) = (((XX(1)*X(1))+(XX(2)*X(2)))**2)/(XX(1)+XX(2))
001466          DO 496 I = 1, NOCU3S
001470          X(4) = X(4) + SSC3S2(I)
001472          X(5) = X(5) + SCS3S2(I)
001474      496 CONTINUE
001476          SIGMAS = (X(4) - X(5))/((M2*NOCU3E)- NOCU3S)
001505          X1(1) = (XX(1)*X(1)*X(1)) + (XX(2)*X(2)*X(2))
001511          FSE123(7,IPLACE) = (X1(1)-X(3))/SIGMAS
001515          IF(MDIFF .NE. 5)  GO TO 478
001517          IF(FSE123(7,5).GT. FCRIT(7))  GO TO 478
001523          XNPFT(7) = XNPFT(7) + 1.0
        C   PHASEIII  FFFECT TESTING ROUTINE
001525      478 DO 480 I = 1, NOCU3E
001527          SC3E(I) = 0.0
001530          SSC3E(I) = 0.0
001531      480 CONTINUE
001533          DO 481 IJ = 1, NOCU3E
001535          NH = MNTMCT(IJ)
001537          DO 482 IT = 1, NH
001540          SC3E(IJ) = SC3E(IJ) + AAIII(IT,IJ)
001545          SSC3E(IJ) = SSC3E(IJ) + AAIII(IT,IJ)*AAIII(IT,IJ)
001551      482 CONTINUE
001553          SC3ED(IJ) = SC3E(IJ) / MNTMCT(IJ)
001556          SCS3E(IJ) = (SC3E(IJ)*SC3E(IJ))/MNTMCT(IJ)
001561      481 CONTINUE
001564          DO 483 I = 1,12
001565          XX(I) = 0.0
001566          X(I) = 0.0
001567      483 CONTINUE
001571          DO 484 I = 1,4
001572          X(1) = X(1) + (SC3ED(I   ) + SC3ED(I+ 4))     /8
001577          X(2) = X(2) + (SC3ED(I+ 8) + SC3ED(I+12))     /8
001604          X(3) = X(3) + (SC3ED(I   ) + SC3ED(I+ 8))     /8
001610          X(4) = X(4) + (SC3ED(I+ 4) + SC3ED(I+12))     /8
001615          XX(1) = XX(1) + (1.0/MNTMCT(I   ) + 1.0/MNTMCT(I+ 4))/64.0
001623          XX(2) = XX(2) + (1.0/MNTMCT(I+ 8) + 1.0/MNTMCT(I+12))/64.0
001631          XX(3) = XX(3) + (1.0/MNTMCT(I   ) + 1.0/MNTMCT(I+ 8))/64.0
001640          XX(4) = XX(4) + (1.0/MNTMCT(I+ 4) + 1.0/MNTMCT(I+12))/64.0
001646      484 CONTINUE
001650          DO 485 I = 1, 4
001652          XX(I) = 1.0/XX(I)
001654      485 CONTINUE
001656          J = 0
001657          DO 486 I = 7,8
001660          X(I)=(((XX(I-6+J)*X(I-6+J))+(XX(I-5+J)*X(I-5+J)))**2)/(XX(I-6+J)+
               1XX(I-5+J))
001674          J = J+1
001676      486 CONTINUE
001677          DO 487 I = 1, NOCU3E
001701          X(10) = X(10) + SSC3E(I)
001703          X(11) = X(11) + SCS3E(I)
001705      487 CONTINUE
001707          ISUMN = 0
001710          DO 479 I = 1, NOCU3E
001711          ISUMN = ISUMN + MNTMCT(I)
```

```
001713    479 CONTINUE
001715        SIGMAS =(X(10) - X(11)) /(ISUMN - NOCU3E)
001722        K = 0
001723        DO 488 J = 1, 2
001724        I = J+K
001726        XI(J)=(XX(I)*X(I)*X(I))+(XX(I+1)*X(I+1)*X(I+1))
001735        K = K+1
001736    488 CONTINUE
001737        DO 489 K = 8, 9
001741        I = K-7
001743        FSE123(K,IPLACE)      = (XI(I)-X(I+6))/SIGMAS
001751        IF(MDIFF  .NE.5) GO TO 489
001753        IF(FSE123(K,5) .GT. FCRIT(K)) GO TO 489
001757        XNPFT(K) = XNPFT(K) + 1.0
001761    489 CONTINUE
001763    850 CONTINUE
001766        MS1 = 1 + MMIN1 + (INTEV)*(MMM-1))
001773    700 CONTINUE
        C     ENDING OF MAIN TESTING LOOP.
        C     FSE123 PRINTING AND OR STORAGE .
001776        WRITE(5 ,92)  ((FSE123(I,J), I = 1,9), J = 1, 8)
002013     92 FORMAT  ( 9F10.5/)
002013        WRITE(39,92)  ((FSE123(I,J), I = 1,9), J = 9,16)
002031        WRITE(85,92)  ((FSE123(I,J), I = 1,9), J = 17,24)
002047   1000 CONTINUE
002052        IT= ITIMER(0)
002054        PRINT 521, IT
002061        PRINT 95, XNPFT, NOFT
002071     95 FORMAT(10X//9F10.1,I10)
002071        CALL EXIT
002072        END
```

2. Optimization Program

```
          PROGRAM  MAXI(OUTPUT,TAPE5,TAPE39,TAPE95)
000003    COMMON  /AAA/  FSE123,  V,C,D,F,C1,C2,C3,NOFT
000003    COMMON/PPP/PG03E,PFAL3E,PFAL2E,PFAL1E,PG03NE,PFAL3N,PFAL2N,PFAL1N
000003    DIMENSION  FSE123( 9,8,500),FCRIT(6), WORK(6,2)
000003    EXTERNAL  PAYOFF
000003    DATA NOFT, N        /500,6/
000003    DATA V,C,D,F,C1,C2,C3/10.0E6,0.0E6,-10.E6,0.0E6,2.0E5,4.0E5,6.0E5/
000003    DATA NOMS  /3/
000003    FCRIT(1) = 20.4 $ FCRIT(2)=.313 $ FCRIT(3)=11.0 $ FCRIT(4)=0.56
000011    FCRIT(5)=9.0 $ FCRIT(6)=7.5
000014    DO 100 NR = 3,NOMS
000016    IF(NR .EQ. 1) GO TO 700
000020    IF(NR .EQ. 2) GO TO 800
000021    IF(NR .EQ. 3) GO TO 900
000023 700 READ(5,1)  FSE123
000031   1 FORMAT(9F10.5/)
000031    GO TO 600
000032 800 READ(39,1) FSE123
000040    GO TO 600
000041 900 READ(95,1) FSE123
000047 600 DELTA = .5 $  EPS = .1
000052    CALL FMIND(N,FCRIT,DELTA,EPS,IERR,PAYOFF,WORK)
000061    PRINT 40,  FCRIT
000067  40 FORMAT  (1X*FCRIT = *6F10.3)
000067    PRINT 501, PG03E,PFAL3E,PFAL2E,PFAL1E
000103    PRINT 501, PG03NE,PFAL3N,PFAL2N,PFAL1N
000117 501 FORMAT(50X,4F6.4)
000117    P = -PAYOFF(FCRIT)
000122    PRINT 50, P
000130  50 FORMAT  (1X,*PAYOFF = *F14.2)
000130 100 CONTINUE
000133    CALL EXIT
000134    END
```

```
                FUNCTION PAYOFF(FCRIT)
000003          COMMON /AAA/  FSE123,  V,C,D,F,C1,C2,C3,NOFT
000003          COMMON/PPP/PGO3E,PFAL3E,PFAL2E,PFAL1E,PGO3NE,PFAL3N,PFAL2N,PFAL1N
000003          DIMENSION  FCRIT(6), FSE123( 9,8,500)
000003          GO3E=0.0 $ GO3NE=0.0 $ FAL1E=0.0 $ FAL1NE=0.0 $ FAL2E=0.0
000010          FAL2NE=0.0 $ FAL3E=0.0 $ FAL3NE=0.0
000013          DO 1 0 K = 1, NOFT
000014          DO 1 1 J = 1, 8
000015          IF((FSE123(1,J,K)*FSE123(1,J,K)).GE.(FCRIT(1)*FCRIT(1)))GO TO 21
000024          IF((FSE123(2,J,K)*FSE123(2,J,K)).LT.(FCRIT(2)*FCRIT(2)))GO TO 21
000035          DO 2 0 I = 3, 4
000036          IF((FSE123(I,J,K)* FSE123(I,J,K)).GE.(FCRIT(3)*FCRIT(3)))GO TO 200
000047          IF((FSE123(I+2,J,K)* FSE123(I+2,J,K))  .LT.  (FCRIT(4)*FCRIT(4)))
              1 GO TO 200
000060          GO TO 333
000060      200 CONTINUE
000062          GO TO 22
000063      333 IF((FSE123(7,J,K)*FSE123(7,J,K)).GE.(FCRIT(5)*FCRIT(5)))GO TO 23
000073          DO 300 I = 8, 9
000074          IF((FSE123(I,J,K)*FSE123(I,J,K)).LT.(FCRIT(6)*FCRIT(6)))GO TO 300
000105          GO TO 1000
000105      300 CONTINUE
000107          GO TO 23
000110     1000 IF(J .GE. 5)  GO TO 11
000113          GO3E = GO3E + 1.0
000115          GO TO 101
000115       11 GO3NE = GO3NE + 1.0
000117          GO TO 101
000120       21 IF(J.GE. 5)  GO TO 12
000123          FAL1E = FAL1E + 1.0
000125          GO TO 101
000125       12 FAL1NE = FAL1NE + 1.0
000127          GO TO 101
000130       22 IF(J .GE. 5)  GO TO 13
000133          FAL2E = FAL2E + 1.0
000135          GO TO 101
000135       13 FAL2NE = FAL2NE + 1.0
000137          GO TO 101
000140       23 IF(J .GE. 5)  GO TO 14
000143          FAL3E = FAL3E + 1.0
000145          GO TO 101
000145       14 FAL3NE = FAL3NE + 1.0
000147      101 CONTINUE
000151      100 CONTINUE
000154          PGO3E= GO3E/(NOFT*4) $ PGO3NE= GO3NE/(NOFT*4)
000162          PFAL1E= FAL1E/(NOFT*4) $ PFAL1N= FAL1NE/(NOFT*4)
000170          PFAL2E= FAL2E/(NOFT*4) $ PFAL2N =  FAL2NE/(NOFT*4)
000176          PFAL3E= FAL3E/(NOFT*4) $ PFAL3N= FAL3NE/(NOFT*4)
000205          PAYOFF=(PGO3E*(V-C3)+PFAL3E*(C-C3)+PFAL2E*(C-C2)+PFAL1E*(C-C1)+
              1PGO3NE*(D-C3)+PFAL3N*(F-C3)+PFAL2N*(F-C2) +PFAL1N*(F-C1))/2
000244          PRINT 500, PAYOFF, FCRIT
000255      500 FORMAT  (1X,*PAYOFF = *F14.2,5X,*FCRIT = *6F8.3)
000255          PAYOFF = -PAYOFF
000256          RETURN
000257          END
```

```
              SUBROUTINE FHIND(NX, X, DELTA, EPS, IERR, F, WORK)
000012        INTEGER I, IERR, IHALF, ISRCH, MXHALF, MXSRCH, NX
000012        REAL DELTA, EPS, F, FNEW, FOLD, FTEMP, WORK, X, XSAVE
000012        EXTERNAL F
000012        DIMENSION X(NX), WORK(NX, 2)
000012        DATA MXHALF/50/, MXSRCH/100/
        C
        C  CHECK FOR CORRECT DIMENSIONING AND FOR VALUE OF EPS.
000012        IF (NX .GT. 0 .AND. EPS .GT. 0.) GO TO 10
000022        IERR = 3
000023        GO TO 99
        C
        C  INITIALIZE VARIOUS VARIABLES.
000023  10    IERR = 0
000024        DO 20 I = 1, NX
000026  20    WORK(I, 1) = 0.
000034        FOLD = F(X)
000045        FNEW = FOLD
000046        DO 90 IHALF = 1, MXHALF
000050        DO 60 ISRCH = 1, MXSRCH
        C
        C  GIVE DELTA SIGN OF LAST SUCCESSFUL CHANGE, IF ANY.
000051        DO 40 I = 1, NX
000052        IF (WORK(I, 1) .NE. 0.) DELTA = SIGN(DELTA, WORK(I, 1))
000063        XSAVE = X(I)
000066        X(I) = XSAVE + DELTA
000070        FTEMP = F(X)
000101        IF (FTEMP .LT. FNEW) GO TO 30
        C
        C  NO DECREASE, TRY OPPOSITE DIRECTION.
000104        X(I) = XSAVE - DELTA
000107        FTEMP = F(X)
000120        IF (FTEMP .LT. FNEW) GO TO 30
        C
        C  NO DECREASE, STAY PUT.
000123        X(I) = XSAVE
000125        GO TO 40
        C
        C  DECREASE, SO GO TO NEW POINT.
000126  30    FNEW = FTEMP
        C
        C  SAVE DIRECTION OF CHANGE, IF ANY, FOR NEXT ITERATION.
000130  40    WORK(I, 1) = X(I) - XSAVE
        C
        C  END OF LOOP ON I.  IF NO CHANGE(I.E., NO DECREASE) IN F, LEAVE
        C  ISRCH-LOOP.
000141        IF (FNEW .GE. FOLD) GO TO 70
        C
        C  OTHERWISE, TRY CHANGING ALL COMPONENTS OF X BY AMOUNT OF LAST
        C  CHANGE, IF ANY.
000143        DO 50 I = 1, NX
        C
        C  SAVE PRESENT COMPONENTS OF X.
000145        WORK(I, 2) = X(I)
000153  50    X(I) = X(I) + WORK(I, 1)
000162        FOLD = F(X)
000173        IF (FOLD .GE. FNEW) GO TO 57
        C
        C  F HAS DECREASED, KEEP VALUE AND TRY AGAIN.
000176        FNEW = FOLD
000176        GO TO 60
        C
        C  F HAS NOT DECREASED, RESTORE VALUES AND TRY AGAIN.
000177  57    DO 58 I = 1, NX
000201  58    X(I) = WORK(I, 2)
000213        FOLD = FNEW
000215  60    CONTINUE
        C
        C  END OF LOOP ON ISRCH.
000220        IERR = 1
000221        GO TO 99
        C
        C  RETURN IF DELTA IS SMALL ENOUGH.
000221  70    IF (ABS(DELTA) .LE. EPS) RETURN
        C
        C  OTHERWISE, HALVE DELTA AND TRY AGAIN.
000225  90    DELTA = 0.5 * DELTA
        C
        C  END OF LOOP ON IHALF.
000231        IERR = 2
000232  99    WRITE (6, 1001) IERR
000240        RETURN
000241  1001  FORMAT(* FHIND  ERROR NUMBER *, I1)
000241        END
```

223