Educating Intelligence: Determinants of School Behavior of Disadvantaged Children

Abstract: This report discusses the effects of a two year intervention on preschool disadvantaged children, as found in a one year follow up. Research and experimental variables of a more or less controllable nature tend to obscure any clear cut conclusion as to the extent and quality of change in these children. It is certainly questionable whether existing standardized tests are appropriate with regard to either the intervention or the socioeconomic status of the children involved.

Aspects of the very broad, complex, and significant problem of the relationship between social class background and intellectual and academic growth were the foci of this study. More specifically, it was concerned with intervention in the preschool and early school lives of lower class children to reduce the likelihood that they would develop intellectual and academic deficits—i.e., mental retardation—so frequently found in children from such backgrounds.

Procedure
The original plan of this research was to locate a group of preschool children drawn from families designated as cultural familial mentally retarded and, during a two year intervention, to provide them with a variety of experiences calculated to engender and reinforce attitudes, motivations, and cognitive skills that are considered prerequisites for normal intellectual and academic growth. It was expected that, in comparison with an appropriate control group, the experimental group would display significantly fewer intellectual and academic deficiencies. The basic premise was that intelligence is educable and that if an appropriate intervention were provided for children “destined” to become mentally retarded, this retardation would be prevented or, at least, mitigated. As the research proceeded, its direction and focus deviated more and more from the original purpose.

A diagnosis of mental retardation for these preschool children was not a criterion for selection. This reflected the position that case finding difficulties of previous investigators were substantive rather than methodological, a position congruent with the authors’ previous experiences in a variety of educational and clinical settings. In addition, it seemed reasonable to assume that if one waited until such a diagnosis were possible, it would be more difficult, or perhaps impossible, to reverse the retardation.

The plan was to select from a population of preschool children where there was a strong likelihood of mild mental retardation (without accompanying central nervous system involvement) within the families of these children. For this purpose, it seemed reasonable to assume that if older retarded siblings were selected as a reference group, their younger siblings would be expected to develop in somewhat similar pat-
terns without outside intervention. In order to maximize the likelihood of mental retardation still further, it was also planned to select subjects who had at least one parent who was mentally retarded, as well as an older retarded sibling. In summary, the original criteria were that subjects: (a) come from a lower social class, (b) be of preschool age, (c) have at least one older retarded sibling, and (d) have at least one retarded parent. These criteria are similar to those for cultural familial retardation listed in the American Association on Mental Deficiency’s (AAMD) *A Manual on Terminology and Classification in Mental Retardation* (Heber, 1959).

It soon became apparent from evaluations of the pilot sample (N=14) and the main sample (N=60) that this method of selection was unworkable. All the families could be classified as culturally deprived, and there were inordinate amounts of school failure and lack of intellectual stimulation in the homes. Some of the families could be classified as cultural familial mental retardates, using the AAMD criteria, but the occurrence of documented mental retardation in the parent appeared to be relatively unrelated to its occurrence in the child. It was, therefore, not possible to obtain a sample of any size if the AAMD criteria for cultural familial mental retardation were to be met.

Therefore, the final criteria adopted for subject selection did not include having siblings in a special class or having a retarded parent, although, as it turned out, many subjects did meet these criteria. The most important criterion was residence in a deprived area characterized by high delinquency rates, a considerable proportion of school dropouts and school failures, low occupational status of parents, and run down homes. Other criteria were: (a) the family’s residence in the area was not temporary, (b) the level of parental education and occupation was usual for that area, (c) neurological examination of the child revealed no central nervous system pathology, and (d) the parents consented to the child’s inclusion in the project.

After a thorough search for eligible children, in which a variety of methods were used, 69 subjects were found who met the criteria. Five were ultimately dropped because their mothers never sent them to the program, two experimental children moved out of the area, and two nonexperimental children moved out of the state, thus reducing the sample to 60 children. Including the pilot sample, 74 cases met all criteria and were accepted as part of the project.

The pilot study, organized one year prior to the principal study, provided the project staff with an exploratory group, permitting study of selection, testing, and curricula procedures before the formal project began. This gave the senior staff of the project the time to train teachers for work with children in a specially designed teaching situation: the Responsive Environment (Moore, 1963). In this method, an electric typewriter was used to enable children to learn through their own discoveries.

The division of the principal sample into two experimental groups and a nonexperimental group was done by stratified random assignment, using Stanford-Binet IQ score, chronological age, and sex in the stratification. This stratification assured maximum efficiency and group equivalence.

Beginning in May, 1962—the case finding year—and ending in May, 1965—the first year of public and parochial school followup of the children—the following types of testings were accomplished yearly:

1. **Cognitive (Aptitude, Achievement, Language)**
   - Stanford-Binet, L-M, 1960
   - Peabody Picture Vocabulary Test
   - Illinois Test of Psycholinguistic Abilities
   - Lee Clark Reading Readiness Test
   - Gesell-Ilg Norms for Musical Ability
   - Typewriter Test (only taken by subjects in Responsive Environment)
   - School Achievement of Study Child (rated by teacher)
   - Metropolitan Reading Readiness Test

2. **Noncognitive (Personality, Social)**
   - Rorschach Inkblot Test (overall rating of differentiation and form level)
   - Vineland Social Maturity Scale
   - Test Taking Behavior (as assessed by psychological examiner)
   - Sociogram Score (sociogram developed by teachers)
   - Anxiety Scales for Children
3. Environmental

Warner Index of Status Characteristics
Absences from Preschool or School
Family Evaluation (quantitative assessment
of interview protocols)

Several characteristics of this research differ-
entiate it from prior and current related investiga-
tions:

1. The subject population was systematically
randomized into experimental and control
groups. In view of the probability that con-
trol children received special treatment and
educational opportunities, they were later
designated as a nonexperimental group. On
the other hand, a fair number of experi-
mental children were not afforded the fullest
opportunities to participate in the
experimental program. However, although
execution was imperfect, an experimental
design was maintained.

2. In the formal evaluations of subjects,
"blinds" were rigorously developed and up-
held. Tests of "blinds," using examinations
of the psychologists who administered the
various protocols, disclosed a lack of knowl-
dge on their part as to which children were
in the experimental group and which were
not.

3. Subject attrition was insignificant. Of the
74 children originally in the project, all com-
pleted participation in the formal interven-
tion years. Of the 60 children from the main
sample, only one child was lost during the
one year followup of children in public
school.

4. A high degree of parental cooperation was
maintained throughout the course of the study.
Two days prior to the completion of
the project, 41 parents attended a social
gathering with the total project staff. The
families of both the nonexperimental and
experimental groups were included.

5. Our hypothesis was that in studying cul-
tural deprivation we were also studying fac-
tors that give rise to mental retardation. Sev-
eral recent studies have made distinctions
between "true" mental retardation and
"pseudoretardation"; i.e., subnormality is in-
dicative of mental retardation only if it is
diagnosed as irremedial. It was our conten-
tion that all children whose performances
show mental retardation are, in fact, men-
tally retarded. As Binet pointed out long
ago, mental retardation is a state of current
subnormal intellectual functioning. Al-
though a child may have an intact central
nervous system and may be categorized as
culturally deprived or cultural familial men-
tally retarded, if he behaves in an intellectu-
ally subnormal manner, he is as mentally re-
tarded as another child with demonstrable
brain damage who is functioning at approxi-
mately the same intellectual level.

6. In testing the hypothesis that intelligence is
educable and that, in some instances, low
intelligence is a manifestation of a deprived
cultural experience, there are certain de-
sign problems. The study of deprivation
and its relationship to social and school per-
fomance, by its very nature, must be either
partially or wholly retrospective. It is not
possible to assign children to experimental
and control groups, and then systematically
deprive the experimental children of certain
experiences in order to observe the effects
of that deprivation. Legal and moral codes
demand that studies be conducted within ex-
sting cultural educational situations, which
give indirect insights into the effects of de-
privation.

Ideally, the proper study of deprivation
would examine the null hypothesis that cer-
tain kinds of social and intellectual depriva-
tion will not cause differences between
children who are exposed and those who
are not exposed. Instead, we are forced to
study a less satisfactory null hypothesis: that
children from a deprived living situation
will not benefit from a stimulating school
curriculum. This hypothesis puts the bur-
den of proof on the curriculum rather than
on the deprivation which is the focus of the
study. Deprivation cannot be systematically
controlled and, therefore, cannot be consid-
ered a true experimental main effect; the
main effect is thus the presence or absence
of a preschool program with children de-
scribed as deprived.
The Curriculum

The objective of our preschool curriculum was to provide an optimal nursery school environment. This was attempted in three principal ways: (a) by helping children learn how to function socially in a group instruction situation so as to be maximally receptive to that instruction; (b) by providing a concentration of experiences designed to arouse curiosity and promote inquisitiveness and positive attitudes toward learning; and (c) by attempting to provide training in certain psychological functions generally considered to be fundamental to acquisition of academic skills in the primary grades.

Certain categories of experiences or activities, or areas of concentration, were commonly used by all teachers: language development, visual discrimination, quantitative thinking, auditory discrimination, auditory memory, speech training, motor coordination, visual memory, speech training, motor coordination, visual memory, and creative and imaginative thinking. The program is described best as experimental, emergent, child centered, and adhering to the basic principles of any sound preschool program—but over and above this, focused on the intensified development of preacademic skills. Intervention consisted of an ongoing, changing, complex social psychological setting having personal, interpersonal, educational, and cultural components.

We do not wish to convey the impression that our curriculum was a separate variable, i.e., that it had an existence and effect independent of the social psychological setting and the person employing it. Such independence never exists. On paper it is possible to compare curricula independent of setting and teacher; in practice it is impossible. To attribute consequence to a curriculum, therefore, is to do more than oversimplify. It is to misrepresent the external reality. For this reason, two major concentrations in this research dealt with the curriculum: one centered on the psychological climate of the intervention and the other on the formal curriculum, its contents and goals, materials used, tasks assigned, skills learned, and overall daily objectives.

Moore (1963) pioneered in the development of theoretical as well as technical aspects of the learning of preschool children, and his contribution to the field fits neatly into our efforts at environmental stimulation. Moore described how children aged two to five can learn to type, read, and write. These skills are acquired through an enjoyable experience, derived from what has been labeled a Responsive Environment. An environment is responsive if it satisfies the following conditions: (a) it is attuned to children's exploratory activities; (b) it informs children immediately about the consequences of their own actions; (c) it permits children to make extensive use of their capacities for discovering relations; and (d) it is so arranged that children are likely to make a series of interconnected discoveries about some aspect of the physical, cultural, and social world.

The setting of this project was not one which made description and manipulation of the component variables easy. In planning and organizing this study, we realized that in the event differences emerged between the experimental and nonexperimental children it would be impossible to say which aspects of the intervention were more or less influential. However, we did assume that a setting (the intervention) could be developed for the experimental children which would contain elements obviously not found in the daily lives of the nonexperimental children. Put in another way: we assumed that we could develop and describe an intervention which would clearly indicate that the two groups of children were experiencing such different things that predicted findings, contrary findings, or no findings would be of significance.

Results

The research hypothesis of this study, that a two year intervention with preschool lower class children will enhance their demonstrated educability, was rejected. This hypothesis was tested "with a variety of measurements over a three year period and included testing of cognitive, noncognitive, and environmental factors. Analyses of the data led to the unequivocal inference that there was no more difference between the groups at the conclusion of the study than there had been at the beginning.

This conclusion can be viewed as: (a) failure...
of the intervention as an effective force in the lives of the experimental children; (b) failure of the measuring instruments to register differential changes in functioning over a three year period; or (c) evidence of our inability to maintain a true experimental design. The implications of these alternative explanations are explored with the explicit goal of setting the stage for future research in this area.

The Measuring Instruments

The measurements used were comprehensive both with respect to substance and technique. Data were obtained in many ways: tests directly administered to children; rating scales where the information was supplied by parent, teacher, or psychologist; measurements which covered specified testing periods and measurements which represented ratings of a child, or a series of direct measurements over an extended period of time; tests of specific abilities and global abilities; measurements concerned with school behavior and behavior in a testing situation; single measurements obtained in any one of the four testing periods and repeated measurements obtained two, three, or four times on each child; and measurements from the domains classified as cognitive, noncognitive, and environmental. Thus, the testing program did not depend upon either one kind of test or one kind of administration of a test.

Given the particular sample of children, the curricula, and the variety of measurements over an extended period of time, our inability to demonstrate significant differences between experimental and nonexperimental groups caused us to view the rejection of the research hypothesis as an internally valid inference. We have no evidence that all curricula would fail to produce changes in all kinds of children from lower class homes or other kinds of environments (problems of external validity), but we do feel confident that the study throws considerable light on generalized problems of external validity.

Although forced to reject the research hypothesis as specifically applied to the sample studied and the curricula used, we do not necessarily reject the generalized hypothesis about the educability of lower class children. The problem of how intelligence manifests itself—the extent to which it is affected by behavior, on one hand, and the extent to which it unfolds, on the other—is a spectre which continually faces us. However, there is an implicit contradiction between the supposition that children can change in response to specific teaching techniques and the probability that these changes will be reflected by tests such as the Stanford-Binet Intelligence Scale. Such tests are constructed in line with an operational principle of stability. Items that show relative variability over time are rejected in favor of items that are more stable. The result is an apparent stability of global measures which is an indication of test specific stability, but not necessarily an indication that children do not change. When a child’s score is based on a comparison with the scores of his age peers, the chances are not very great that he will change with respect to these same age peers over a period of years. Obviously he is changing and, during preschool years, changing very rapidly. An apparent lack of change is not a reflection of his own growth, but rather of his position in a frequency distribution.

The importance of global measures of scholastic aptitude, or of “intelligence” as this is commonly called, has to do with efficiency in predicting future academic performance. More specific measures of ability do not have a comparable level of efficiency, even when the prediction is to be made within the specific area tested. For example, the Stanford-Binet is reputed to be a good predictor of future mathematical or verbal performance; it does not necessarily follow that it measures a more innate quality, although many investigators have thought so. However, the Stanford-Binet was developed to be stable, and this is reflected in high predictive efficiency with respect to other kinds of test or nontest behavior which relate to intellectual functioning. Therefore, it follows that, for an instrument to be highly predictable of future behavior, the instrument must be concerned with behavior that is extremely stable over time. Just the opposite is necessary in tests designed to be sensitive to changes in individuals. The latter tests tend to be relatively unstable over time and have minimal ability to predict future performance. Tests which are measuring nonstable factors are not necessarily
unreliable, although this presents problems for the test constructor, who must use psychometric and statistical techniques of measuring reliability which are more or less independent of the time factor.

The problem alluded to is the obvious dilemma of any curriculum which attempts to provide specific developmental and, if necessary, remedial activities for children. These activities attend very carefully to fragments of the reading process, the quantitative process, or whatever intellectual discipline is being taught, as opposed to those procedures which are more global in their design and execution and which treat larger units of behavior. This does not mean that the more global approach to teaching does not involve attention to pieces of learning behavior, but these are not the main focus. Furthermore, although it is important that some pieces be attended to, it does not make too much difference which pieces they are.

**Problems of Validity**

The question is whether the results obtained in this study (which, at this point, are assumed to be internally valid) have external validity—i.e., whether they have general application in the area of the preschool education of lower class children. To find some resolution to this rather crucial problem, we see four rather distinct areas for discussion. The first three are concerned with problems of curriculum, timing, and sampling of children. The fourth concerns problems in measuring changes in children over a period of time with instruments that have been designed to measure relatively stable factors.

**Curricula Intervention.** Many questions can be raised about the optimal strategies in a preschool program for lower class children. The evolving curriculum of this study was clearly teacher dominated, although there was extensive collaboration between the investigators and all of the teaching staff.

The Responsive Environment provided a methodology which, by design, was not teacher dominated and which depended upon detailed instructions aimed specifically at minimizing teacher variability. Of course, it is possible that in spite of the care taken to specify the procedures of the Responsive Environment, the teacher may still have dominated the learning situation.

It might be that optimal strategy calls for comprehensive programming throughout the schooling of lower class preschool children so that the curriculum is dominated by theoretical considerations which are independent of variations among teachers or within any one teacher over a period of time. The view of the principal investigators of this project was that the curriculum should not be preordained, but rather that it should be developed by teachers in response to individual children, the interpersonal relationships between teacher and children, and relationships within groups of children. A great deal of attention and energy was given to observing teachers and children and to holding seminars with the teachers on programs they were developing and methods they were using.

Alternatively, other investigators have discussed the use of tightly prescribed curricula dominated by considerations other than the personalized curricular development of particular teachers. For example, the prescription can go in the direction of specific operant techniques applied to particular learning sequences, or it can be concerned with a therapeutic climate in the classroom.

In light of the above discussion we must raise the question of whether our interventions of a preschool program and a Responsive Environment provided a sufficient test of the hypothesis of educability. Perhaps greater attention should be paid to extensive and systematic variations of interventions which either minimize or use teacher variation and which use measurements directly related to curricular procedures. These problems of measurement of the extent of teacher or methodology domination have always been a source of concern. A teacher dominated curriculum does not lend itself readily to direct measurement. The consequent dependence upon global measurement suggests a dilemma in designing studies that attempt to compare variations along the continuum of teacher domination.

A second criticism of the curriculum concerns the focus of the entire program on the preschool environment rather than specifically attending to the education and treatment of the entire
community and the families. It was not within the plans of this study to treat families or to deal extensively with evaluations of siblings and other children in the community. The curriculum was concerned with the Responsive Environment and the preschool program, with peripheral attention to families during occasional meetings and home visits by teachers. The failure of this program, as described, may have been due to the failure of the intervention to affect total family behavior. Since we found a relatively high correlation between a measure of family adequacy and the average school performance of all siblings, the inference follows that school failure is family linked and must, therefore, be family treated.

Timing and Duration. It is possible that the failure of the intervention to produce demonstrable results in the experimental children was due to the timing of the program in the lives of the children or to the duration of the program for individual children. Possibly children of different ages respond to different kinds of intervention in diverse ways. In any event, timing has variable effects when considered in conjunction with curricular strategies, sampling variation, and measurement problems.

Sampling of Children. There is some reason to believe that the children selected did not provide the most advantageous sample for testing the hypothesis of educability of lower class children. Some investigators have reported that transiency in some lower class schools is so high that the turnover in classrooms in a single school year sometimes reaches 100 percent. The families of the children in this study were unusual in that they tended to stay in the same geographic area and, for the most part, in the same houses—59 of the 60 principal subjects remained with the study over the entire three year period. These families, therefore, had the opportunity to receive different kinds of continuing services from private and public agencies, and the schools were able to maintain contact with families and children. Although the neighborhood was clearly lower lower class and the families in this neighborhood were classified as extremely impoverished and largely dependent on welfare assistance, the general nature of the community may have been such as to support the intellectual growth of the children. If this were true, one might expect that an experimental intervention, such as the one provided, would not produce any demonstrable results.

Along the same line of reasoning, and again in spite of the fact that the neighborhood under consideration was rated as a lower lower class neighborhood, an important discrepancy appears with respect to sampling strategy. The area is not adjacent to other lower lower class neighborhoods; rather, it is a pocket within a city, surrounded by a variety of neighborhoods, universities, and business areas. None of the families were geographically distant from social agencies, hospitals, stores, or universities. Over the past 15 years, there had been city planning projects, university programs, religious group involvement, social work, two well established neighborhood house programs, and many other kinds of service activities.

It is difficult to speculate on the effects of such sampling variation, but we feel that, unwittingly, we did not obtain as educationally disadvantaged a sample of homes and children as desired. The more deprived a child is, the more likely that he will respond to an intervention. This reasoning follows directly from the rationale which maintains that negative deviations from normal functioning are likely to be associated with family and educational deprivations. We are not against the generality of the thesis of educability, but we think, for the time being, that this thesis can best be demonstrated with the most severely deprived children.

Measurement Problems. Measurement problems will always plague investigations such as this one. The day to day intervention which involves a variety of teachers and children is a different kind of substance than the very specific and highly reliable tests often used to measure the effects of interventions. When measurements are concerned with abilities that are closely tied to developmental factors, the apparent relationship between interventions and criteria is spuriously high. It is more parsimonious to conclude that the increased ability of individual children over periods of time is due more to their growth than to any intervention, whether it be school or a particular remedial or therapeutic sequence.
This is partially a question of precision as it applies to different kinds of measurements. A child's growth in mental age refers to his increasing ability to respond to items on a test as he grows older. The concept of chronological age, as it is used in psychometrics, refers to the average performance of children at any particular chronological age. The development of most tests of aptitude and ability hinges upon the changes that take place over time either directly or as can be inferred from the performance of a cross-sectional sample of children. Our study was focused upon the question of whether the slope of developmental growth can be affected by an intervention. The question raised was whether the developmental acceleration of particular children can be varied by systematically providing specific interventions. It is altogether possible that developmental levels can be changed, but that this change is hidden in the relatively small amount of variance that is left over after chronological age is literally partialed out. As a matter of fact, the residual variance that remains is not grossly different from that which must be attributed to error. Therefore, relatively little is left for the measurement of changes in children, particularly when these children have more or less normal developmental slopes.

**Conclusion**

Insofar as this study is concerned, we have neither significant nor convincing data to substantiate our research hypothesis that intelligence is educable, i.e., a function of practice and training. However, this study revealed that we still have a great deal to discover concerning the nature/nurture interaction, about the most efficient and sufficient period to begin interventions, and about the possible intervention models that may have the greatest desired effects. The information obtained in this study has encouraged us to continue the quest for processes and methodologies to educate intelligence and, for certain children, to prevent mental retardation.

**References**


**Burt Blatt is Professor and Chairman, Special Education Department, and Frank Garfunkel is Associate Professor and Director, Headstart Evaluation and Research Center, School of Education, Boston University, Massachusetts. This paper summarizes the author’s monograph, A Field Demonstration of the Effects of Nonautomated Responsive Environments on the Intellectual and Social Competence of Educable Mentally Retarded Children. Washington: US Office of Education, 1965.**

---

**CONTINENTAL’S**

**Preprinted Carbon Masters for Liquid Duplicating Provide...**

... the flexibility of usage needed in Special Education work

**SPECIAL EDUCATION PROGRAM**

| U.S. Money | levels 1, 2, and 3 |
| Measurement | levels 1, 2, and 3 |
| Time | levels 1, 2, and 3 |
| Useful Language | levels 1, 2, and 3 |
| Number Concepts | levels 1, 2, and 3 |

Also available to help you

**181 TITLES FOR REGULAR CLASSES.**

No grade level designed on the pages.

- Reading Readiness
- Phonics and Reading
- Health and Safety
- Mathematics
- English
- Social Studies
- Science

Write for FREE Demonstration Catalog showing contents of each book in Special Education Series and a Complete Catalog listing materials for regular classes.

**THE CONTINENTAL PRESS, INC.**

Elizabethtown, Pa. 17022

---

608 May • 1967