LONGITUDINAL IQ OUTCOMES OF THE MOTHER-CHILD HOME PROGRAM, 1967-1973

John Madden, Phyllis Levenstein, Sidney Levenstein

Verbal Interaction Project
Family Service Association of Nassau County, Inc. and State University of New York at Stony Brook

5 Broadway
Freeport, New York 11520

The research described in this report was supported by the U. S. Dept. of HEW: Children's Bureau, National Institute of Mental Health, and Office of Education; and by the Carnegie Corporation of New York; the Foundation for Child Development; the Grant Foundation of New York; and the Rockefeller Brothers Fund.

November 1974

LONGITUDINAL IQ OUTCOMES OF THE MOTHER-CHILD HOME PROGRAM, 1967 - 1973

John Madden, Phyllis Levenstein, Sidney Levenstein

Verbal Interaction Project

Family Service Association of Nassau County and State University of New York at Stony Brook

The Mother-Child Home Program (MCHP) has combined child's play with mother-child dialogue to foster the cognitive and socioemotional development of low-income two to four year olds to prevent educational disadvantage in their later years. This paper will summarize one cognitive outcome, that reflected in general IQ scores on standardized tests for children entering this early childhood intervention program in 1967, 1968, and 1969, and for untreated groups compared with them.

The program consisted essentially of home visits by "Toy Demonstrators" to model for mother-child dyads the verbal interaction features of books and toys permanently assigned to the child. The Mother-Child Home Program was developed by the Verbal Interaction Project (VIP) from a pilot project in 1965-1966 (Levenstein and Sunley, 1968) and began its formal existence in 1967. Its rationale, method, and short-term results have been described in detail elsewhere (Levenstein, 1970; Levenstein, 1974; Levenstein, in press).

The program's major assumption, developed within a multi-disciplinary framework, was that the principal cognitive element missing from the early experience of many children vulnerable to educational disadvantage, was a sufficient amount of concept-building verbal interaction in the family, centered around perceptually rich and ordered stimuli, and embedded in the affective matrix of the child's most enduring relationships, especially that with his mother.

The 1967 cohort's results stimulated a number of questions related to the cognitive effect:

1. Will program "graduates" retain their average 17 point (approximate) General IQ gain into school years?

2. Will subsequent cohorts of the program make similar gains with Toy Demonstrators of lower educational and work skills than those of the social workers with Masters' degrees who pioneered this role in 1967-1968?

3. What will be the long-term effects of varying the intensity of the Program?

4. Will there be lasting differences in IQ among treated and untreated groups?

5. Which background and outcome variables will be found related to IQ score at follow-up?

Longitudinal research of these questions was pursued by the VIP in studies of short-term effects in 1968 and 1969, and in follow-up studies in 1969, 1970, 1971, and 1972, for the sample to be described in this paper. Cohorts enrolled in the MCHP in 1970, 1971, and 1972 were also studied (Levenstein, 1974) and found to have similar short-term cognitive results to those reported here. Other dependent variables studied longitudinally at the same time for outcome, and relation to General IQ, were: Verbal IQ; change scores in both general and Verbal IQ; academic achievement; socioemotional coping skills; associative and conceptual ability; and mother's parenting skills, including verbal interaction (short-term).

The time period selected for this report on General IQ as a cognitive outcome is the VIP's most recently completed follow-up study, Follow-up D, conducted from November 1972 to March 1973.

Method

Design

The research followed a "quasi-experimental" design (Campbell and Stanley, 1963). The basic plan consisted of pretesting with repeated measures following pretest. There was also one "after-only" comparison group which was not pretested. The specific intervention and testing schedules for each group are provided below in the group descriptions. At the time these groups entered the MCHP, social considerations and the intervention method precluded the possibility of randomly assigning Ss to groups. In 1967, when five groups were started, randomization was by location of three suburban housing projects, from which dyads were recruited. In 1968 and 1969, the majority of subjects were assigned to receive the same treatment so that there was no random sampling of groups for assignment to varying treatments. Group and S differences were controlled by their shared residence in low-income housing. Ss previously identified as Experimental (E) in Levenstein (1970) will be designated as Treated (T) in this report, and Ss receiving testing only and previously identified as Control (C) will be identified as Comparison (C).

Subjects

The follow-up Ss were 138 children (74 boys and 64 girls), five to eight years old at follow-up, ranging from kindergarten to fourth grade, and falling into seven 1967-1969 treated groups (N=83) and

three 1967-1972 comparison groups (N=55). The long-term cognitive and socioemotional effects of varying treatments in early years were evaluated in the 1972-73 follow-up study. All lived in three suburbs of New York City, and were from low-income families financially eligible for low-income housing projects. Almost all were American born, English speaking and socially defined as Black.

The 83 treated Ss had been recruited at age two and three years, with their mothers, by letter followed by door to door canvassing, from the Fairport and Green Cove low-income housing projects in 1967 and 1968. The project names are ficticious. In 1969, about half came from these sources and half from referrals by social agencies of dyads outside of the projects.

Of the 55 comparison Ss, 25 were recruited in the same way as the treated Ss from the Fairport and the Manton housing projects in 1967. The remaining 30 Comparison Ss (all found post-hoc to be Black) were recruited as an out-of-project, English speaking, American born "after-only" group on four low SES criteria from the first grade of the Fairport school system in 1972. The criteria were eligibility for low-income housing, residence in rented housing, and neither parent with an education above high school or with an occupation higher than skilled.

Procedure: Mother-Child Home Program (MCHP) and Cognitive Testing

The basic treatment (MCHP) in 1968 and 1969 was the same as in 1967 (Levenstein, 1970) except that children's entry age was restricted to two years instead of two and three; the duration of treatment was two years (from ages two to four) instead of one; the number of play materials was reduced from 28 to 23 for each year, and the minimum level

of education required for the Toy Demonstrators (TDs) was reduced from a Master's degree in social work to partial completion of high school.

Thus, in 1968 and 1969 the MCHP consisted of 46 semi-weekly visits to the dyad each year which followed the local ten month school calendar. The visits began with the child at age two (Program I) and continued throughout the following year (Program II). During these visits, the TDs demonstrated a dyadic verbal interaction with the child, and encouraged the mother to participate in the interaction. The interaction centered on permanently assigned play materials which the TD brought weekly, a total of 12 books and 11 toys each year. These Verbal Interaction Stimulus Materials (VISM) were selected on explicit criteria, foremost of which was their capacity to stimulate verbal interaction (Levenstein, 1974). The cost of the program, without research components, was estimated to be about $400 a year for each child.

After 1967 the TDs were unpaid women volunteers (usually with four years of college) and paid former mother-participants (of no more than high school education). All were trained together in an initial eight-session Training Workshop, in weekly group conferences, and by individual supervision throughout the program year, learning the rationale, the structured cognitive curriculum, and the less formal "affective curriculum", built around the increasingly complex sequence of books and toys presented to the child. They were taught to model for the mother (rather than teach) verbal interaction techniques focused around the toys and books and spelled out in VISM Guide Sheets contained in a "Toy Demonstrator's VISIT Handbook". The TD involved the mother early in the Home Session with the aim of fading into the background of the session

while the mother took over the main responsibility for the verbal interaction, utilizing as much of the TD's modeled behavior as she wished.

The overall aim of the cognitive curriculum contained in the "Toy Demonstrator's VISIT Handbook", and demonstrated to the mother, was to assist the child in building concepts through "instrumental conceptualism", Bruner's phrase applied to the child's conceptual development through his interchange of language with his mother around meaningful experiences in his environment (Bruner et al., 1966). The general goal of the affective curriculum was to sensitize the TDs to the emotional needs of the dyad, to promote, without counseling or teaching, the socioemotional development of the child, and to promote child-rearing behavior in the mother functional to the child's learning and the well-being of both. The two curricula have been described in detail elsewhere (Levenstein, in press).

Treated Groups: Variations of the MCHP

Seventy of the 83 treated Ss in the 1972-1973 follow-up study were graduates of one of four one or two year variations of the MCHP, having been pretested and enrolled in the Program in September of 1967, 1968, and 1969.

Treated Groups (1967)

T67-I+Short II. - This group, from the Fairport Housing Project, entered in September 1967 at age two years, and received Program I plus a shortened version of Program II during the following year (nine home sessions and seven VISM) accepted as an option by the mothers. TDs for this group were social workers. The group was posttested five times (5/68, 5/69, 12/69, 12/71 and 1/73).

T67-I+Short II VISM. - This group, from the Fairport Housing Project, entered in September 1967 at age two years, and received Program I plus seven Program II VISM, accepted as an option by the mothers, during the second year. TDs for this group were social workers. The group was posttested five times (5/68, 5/69, 12/69, 12/71 and 1/73).

T67-I. - This group, from the Fairport Housing Project, entered in September 1967 at age three years, received only Program I, with social workers as TDs, and was posttested four times (5/68, 12/69, 12/70 and 1/73).

T67-C_1+I. - This group, age two and three, from the Green Cove Housing Project, entered and was pretested in September 1967, and received one year of placebo treatment (home visits plus non-VISM gifts) as a comparison group to control for the Hawthorne effect in the first year of short-term research. The placebo treatment had almost no cognitive effect (Levenstein, 1970). The MCHP was offered the following year to all eligible families in this group's housing project. For ethical and human relations reasons, this group was offered, and accepted, Program I with non-social worker TDs, following its year of placebo treatment with a social worker. Group T67-C_1+I was posttested five times (5/68, 5/69, 11/69, 12/71 and 1/73).

Treated Group (1968)

T68-I+II. - This group, from the Fairport and Green Cove Housing Projects, entered in September 1968 at age two years, received Program I and Program II (the first group to receive the full MCHP) with non-social worker TDs. The group was posttested five times (5/69, 5/70, 12/70, 12/71 and 1/73).

Treated Groups (1969)

T69-I+II. - This group, from the Fairport and Green Cove Housing Projects and from non-project residence, entered in September 1969 at age two years, received Programs I and II, with non-social worker TDs, the same treatment received by the T68-I+II Group. The group was posttested three times (5/70, 5/71 and 1/73). By 1969 this treatment was considered to be the stable format for the model program.

T69-VISM Only. - This group, from Fairport non-project residence, entered and was pretested in September 1969 at age two years, received only the VISM for Program I and II delivered on a regular schedule over two years. The group was posttested three times (5/70, 5/71. and 1/73).

Comparison Groups (Untreated)

The 55 Comparison (test only) Ss in the 1972-1973 follow-up study entered the research project as three different groups in two years, 1967 and 1972.

C2-67. - This group, from the Manton Housing Project, entered in September 1967 at age two and three years. The group was pretested in September 1967 and posttested together three times (5/68, 12/69 and 1/73). The group was also tested as two subgroups, the subgroup of Ss age two at entry in 12/71, and the subgroup age three at entry in 12/70, making a total of four posttests for each subgroup.

C4-67. - This group, from the Fairport Housing Project, entered in September 1967 at age four years and was pretested in 1967 (having been recruited for a version of the program for four-year-olds which was cancelled when it was discovered that all four-year-olds were going into the then new Headstart Program). This group was posttested twice (12/71 and 1/73).

C_5-72. - This group from Fairport non-housing-project residence, entry age six years, was tested once in 1/73, having been recruited on low-income criteria previously described. Group C_5-72 entered the 1972-73 follow-up study as an "after-only" group with no previous project contact.

Data Collection

IQ measures were the Cattell Developmental and Intelligence Scale (Cattell) for pretest, the Stanford-Binet Intelligence Scale (S-B) after Program I and in all subsequent tests through age 5, and the Wechsler Intelligence Scale for Children (WISC) for children of age six and older. General IQ at Follow-up D (referred to henceforward as IQ) was distinguished from a "Verbal" IQ obtained from the Peabody Picture Vocabulary Test (PPVT).

The old editions of the S-B (1960) and the WISC (1949) were used to minimize differences between tests, to insure comparability with past VIP results and with recent survey data which also indicate the mean population increase in scores that is shown by the test restandardizations (Roberts, 1971). Because the S-B test items have not essentially changed, it is possible to convert old into new scores. To give a rough idea of the effect of such a conversion on the data reported here, at age 66 months, the score of an S who received the same score as the mean of the untreated C_2-67 Group would be reduced from 92.5 points to approximately 85 or 86, and the score of an S who received the same score as the mean score of the treated T68-I+II Group at the same age would be reduced from 107 to about 99 (Terman & Merrill, 1973). It should be noted that the changes in the scale were not constant across age and raw score levels. No provision has been made for similar

rescaling of WISC to WISC-R scores. Some of the WISC items have been changed so that equivalency of raw scores may not be assumed (Wechsler, 1974).

The validity of the Cattell, standardized in 1940, rests primarily on its relation to 1937 S-B test scores obtained from a relatively small and probably atypical sample. Thus none of the test scores reported here may be taken as representing deviations from a population mean of 100 with standard deviations of 15 or 16.

The problems of generalizing from one test to another across different ages is still more complex due to item differences, differences in test standardization, and due to the probability that young children apply different skills to solve the same problems at different ages.

Despite these problems, IQ tests constituted the best measures available to assess early intellectual functioning, and we do have some reason to believe that they are useful. The correlation between pretest IQ obtained on the Cattell and posttest 1 IQ on the S-B was .60 both for those Ss in the present sample for whom these measures were available (N=91) and for Ss in the 1971-72 VIP short-term study (N=71). Considering the moderate restriction of the range of scores of these groups and the eight month period between pretest and posttest, this figure indicates a fair level of agreement for tests at this age level. Wechsler (1974) reported a correlation of .82 for six-year-olds between the new revision of the S-B and the WISC-R, obtained from a stratified sample (N=33), and it would be expected that a similar but somewhat smaller relation would exist between the older editions of each scale for the present groups.

In addition to the correlations between tests, the mean IQ scores of the comparison groups did not vary substantially from one age or test to the next, although the standard deviations did increase. Thus, although the use of these tests creates serious problems, they appear to be useful for two purposes. The first is to indicate the relative standing of Ss and groups within the present sample, and the second is for comparisons of the performance of the present groups to normative studies, using the same editions of the same tests used here.

Demographic Data

Demographic data were systematically collected in two interviews. The first occurred just before pretesting in a home interview with the child's mother, conducted by the program supervisor responsible for supervising the dyad's Toy Demonstrator throughout the program. The second was conducted at follow-up.

Data Analysis

Frequencies, means and standard deviations of demographic and IQ data as well as t tests for IQ "gain" are reported for each group. The significance of group and treatment differences at follow-up was tested by a mixed design analysis of variance. A mixed design analysis of covariance was used to account for the relation between pretest scores and follow-up differences between Ss. Correlations among 62 background, process and outcome variables were also calculated within broad treatment classifications.

Results

Demographic Characteristics of Sample

Table 1 presents demographic data for the 10 treated and control groups in areas usually considered associated with low-income status of

the family, education of both parents, occupation of father, amount of the father's presence in the home, family size, and proportion of mothers receiving welfare aid.

Insert Table 1 about here

Most of the group means and proportions of Table 1 items go in the expected directions for a low-income sample. The average level of father's education was below high school graduation, as anticipated, but higher than that sometimes reported for poverty groups (e.g., Klaus and Gray, 1968). The range is narrow among group means, from mid-ninth to mid-eleventh grades. Most fathers in the sample were in low status occupations: unemployed, unskilled, or semi-skilled. This was the pattern for all groups, but it was less true for the T68-I+II and T69-I+II Groups than for the others, raising a question of group equivalence on this SES variable. Two-thirds of the fathers in the total sample were living in the home, but there was a wide variation on this variable among the groups, with two C groups showing the greatest extremes, from 23% present for the C_5-72 Group to 93% present for the C_2-67 Group.

The average level of mothers' education was a little higher than that of fathers', but the range among the group means was about the same as that for fathers, from ninth grade (C_5-72) to eleventh grade. Most groups were at the tenth or eleventh grade level. About one-third of the mothers in the total sample were receiving welfare aid, but there was great variation among the groups, from 6.7% mothers on welfare for the C_2-67 Group to 82% on welfare for the C_5-72 Group. The proportion of mothers receiving welfare in every group was roughly similar to the proportion of fathers absent from home.

Family size (including the index child) varied among the groups, from means of 4.3 (T69-VISM Only) to 6.9 (T67-I+Short II), with an average of 5.4 for the total sample.

Retention Rates

The number of Ss in each group at program entry, completing designated programs and available for testing are reported in Table 2.

Insert Table 2 about here

These rates are of concern not only as they affect the data but as an indicator of the feasibility of the program. The overall rate of retention between pretest and follow-up was 73%, not including the C5-72 Group which was first contacted in 1972. Retention rates were similar for all groups, except the T69-VISM Only Group. Of all treated Ss, 94% completed the first year, and 77% completed the second year of intended treatment. The low rates of attrition appeared due to the high degree of maternal acceptance of the treatment programs, as noted by the reliability of T68-I+II and T69-I+II mothers in keeping about 85% of Home Session appointments; and by enthusiasm expressed in final interviews and on voluntary, anonymous evaluations mailed by treated mothers to the VIP at follow-up. The rate of return of the anonymous questionnaires was 46%, almost all of them containing opinions highly favorable to the program. In untreated groups, acceptance of the testing seemed equally high. All mothers appeared to view it as a service from which they received information about impending developmental disabilities, and they gave good cooperation. Of course, all such evaluations by "happy consumers" must be treated with caution.

IQ Outcomes

General IQ test scores are summarized in Tables 3 and 4 for all ten groups, from pretest in 1967 to the last posttest in follow-up, 1972-1973.

Treated Groups

Table 3 gives the IQ data for treated groups. Differing Ns between pretest and other test periods resulted from the fact that Ss not available during one test period were recovered in later test periods.

Insert Table 3 about here

IQ gains in treated groups. - IQ gains as reported here refer only to actual test scores. The meaning of these gains is severely qualified by the problems discussed above and by the fallibility of the measures which is not constant across tests.

Pretest scores for the six treated groups ranged from 82.6 (T67-I+Short II) to 90.4 (T68-I+II). Every treated group gained in IQ in its first year of the program, except for the T67-C_1+I Group, which received placebo treatment in the first year and made a gain of 0.9 points. When this group received the full treatment of Program I in its second year in the Project, its gain increased to 11.5 points.

The five treated groups posttested eight months after termination of the program (28 months after pretest) continued to maintain a significant gain over pretest scores, at the .05 or .01 level. Four treated groups posttested 40 months after pretest, when three of the groups were mainly in kindergarten (T68-I+II, T69-I+II, and T69-VISM Only) and one was in first grade (T67-I), continued to maintain a significant gain over pretest, all at the .01 level. Significant gains continued into

first grade for three of the groups tested (T68-I+II, T67-I+Short II, and T67-I+Short II VISM), and into first and second grade for one of the four groups tested 52 months after pretest (T67-C_1+I). Sixty-four months after pretest, significant gains continued to be maintained by three groups (T67-I+Short II, T67-C_1+I, and T67-I) into second and third grades. The only non-significant gain in treated groups in any follow-up occurred in the T67-I+Short II VISM Group 64 months after pretest, although a gain of 12.9 points over pretest score was retained. Thus, the answer to the first question regarding gain retention may be answered affirmatively.

IQ status of treated groups at follow-up. - The mean follow-up IQs of five of the seven treated groups were above the national norm of 97.7 for 6-11 year-old children with fathers completing 9-11 years of schooling (Roberts, 1971, p. 55).

The mean IQs of the only groups with two full years of the Program (T68-I+II and T69-I+II) were well above the same national norm. The mean age of the T68-I+II Group, with an IQ of 105.4 was 6 years, and more than half of the group was in first grade. The mean age of the T69-I+II Group, with an IQ of 113.3, was 5 years, and more than half of the group was in kindergarten.

The small sizes of the other treated groups make it difficult to draw conclusions about the effect of the treatment received by any particular group. The T67-I, T67-I+Short II, and T67-I+Short II VISM Groups received roughly comparable treatments. Their mean IQ scores at follow-up were roughly comparable to each other and near the national norm for children with fathers with 9-11 years of schooling. The mean

IQ of the T67-C_1+I Group was comparable to that of the full two-year treated groups. Although this result is inconclusive, the pattern of IQ scores of this group over the years compared to other groups may suggest a preparatory effect for the placebo treatment received in the first year of intervention. The mean IQ of the T69-VISM Only Group was also above the expected norm, but this result is tempered by the small N of the group and the 50% rate of retention, a level well below that of other groups. Because of the obvious economy of such a program and the high follow-up scores of the group, the treatment will be investigated in a controlled experimental study which is currently under way.

Comparison Groups

Table 4 presents a consistent picture of follow-up IQ gains and status for the comparison groups.

Insert Table 4 about here

The two comparison groups which entered in 1967 started off with almost identical IQ scores, about 91.

The C_4-67 Group, now nine years old and in third and fourth grades, rose six points in mean IQ from pretest to follow-up, close to the norm for the population with fathers of 9-11 grade education.

The C_2-67 Group obtained a mean follow-up IQ of 95.5, appearing quite similar in performance to the C_4-67 Group. These Ss are now ages seven and eight and in second and third grades.

The "after-only" comparison group, C_5-72, was recruited for the 1972-1973 follow-up study and achieved a mean follow-up IQ of 91. The mean score of this six year old, first grade group was well below the

national norm and was similar to that of the two other groups at pretest. The actual pretest status of this group is, of course, unknown. Although matched on key background variables when recruited, this group was sampled by a different operation than the other groups, which may have introduced a sampling bias.

The descriptive results provide a preliminary set of answers to the questions raised above. As far as the data go, it appears that the groups who participated in two full years of the program (T68-I+II and T69-I+II) performed at a satisfactory level on IQ tests after program intervention by TDs who had no formal social work training.

With the exception of the T67-C_1+I and T69-VISM Only Groups, groups receiving less than two years of treatment obtained somewhat lower scores at follow-up. The C groups consistently obtained scores at or below their expected level based on norms for children with fathers of 9-11 years of education.

Treatment Comparisons

The long-term effects of varying intensities of the program were examined for nine of the ten groups. The T69-VISM Only Group was excluded from these analyses because the treatment received by this group is not qualitatively comparable to other treatments. On the basis of program duration, this group received two years of treatment. On the basis of kind of intervention, the group ranks somewhere between no treatment and two years of treatment. Because of the IQ scores obtained by those T69-VISM Only Ss who were available for testing at follow-up and because of the obvious economy of such a program, this program variation will be examined in an experimental study currently under way.

The analyses comparing the remaining nine groups required several assumptions which qualify the interpretations of the results.

First, the groups considered in this report were not all sampled by the same operation, and there were differences in how they were assigned to treatments. In 1967, when a variety of program variations were attempted in initial formative research, the residentially constituted groups were randomly assigned to different treatments. In 1968 and 1969, all Ss, with the exception of the T69-VISM Only Group, received the same treatment. The C_5-72 Group was obtained by applying low SES criteria to the first grade of a local school system.

These differences are to some extent confounded with amount of treatment, in that the T69-I+II and T69-VISM Only Groups were sampled both in and outside of the housing projects. They were, however, sampled from the same general population as the other groups. The untreated C_5-72 Group was also differently obtained as described above.

Although Table 1 demonstrated some demographic differences among the groups within the low-income criteria, which may have influenced the IQ differences appearing in Tables 3 and 4, detailed examinations of the data ruled out many other possible sources of IQ differences, such as self-selection, differing attitudes toward acceptance of treatment, tester bias, and selective attrition. The facts remain, however, that individual Ss were not randomly assigned to experimental and control groups, and not all groups were randomly assigned to treatment categories.

Second, the groups considered here were not all of the same age at program entry, and these age differences were partially confounded with treatment differences (See Tables 3 and 4). It was thus not possible to

correct for age differences when examining treatment differences between groups.

Third, to avoid missing data due to a testing schedule which staggered groups over year of follow-up test, these analyses considered follow-up times to be equivalent for the different groups. This is an unfortunate compromise, but it appears to be the best that the data offer. The follow-up scores were 40 months after program entry for the T69-I+II Group and 52 months after program entry for the T68-I+II Group. The C5-72 Group (excluded from the analysis of covariance but included for the IQ score analysis below) was recruited from local first grade classes as an after-only control. The group's mean age at testing was equivalent to that of most other Ss 52 months after program entry. The other groups in the analyses were tested at 64 months after program entry.

These three differences between groups which were not consequences of treatments were all confounded to some extent with differences in treatments which each group received. The differences were all finally eliminated from ongoing research when, in 1973, it was possible to assign Ss randomly to treated and untreated groups. However, we did not wish to withhold this report pending the future receipt of long term results for groups currently in the program.

The long-term effects of varying intensities of treatment are indicated by the correlation between amount of treatment received by each group and the mean follow-up IQ for each group. Spearman's rho for these variables was .66, which illustrates a positive linear relation between follow-up IQ and amount of treatment received.

The long-term differences in IQ between groups receiving varying amounts of treatment were tested by an analysis of variance of follow-up scores. The hierarchical analysis of variance which was performed partitioned variance into sources of subject sampling, group sampling within treatments, and treatments (Kirk, 1968, p. 232). The analysis allowed for random sampling differences between groups within treatments as a possible cause of what could otherwise appear to be an effect of treatments. This model was necessary because of the non-random assignment of Ss to groups.

The treatment classifications were two years of treatment, one year of treatment and no treatment. For the assignment of groups to treatments, a year of treatment was defined as more than half of a 46 session program year. According to this criterion, the T68-I+II and T69-I+II Groups were classified within the two-year treated category; the T67-I, T67-I+Short II, T67-I+Short II VISM and T67-C_1+I Groups were classified within the one-year treated category, and the C_2-67, C_4-67 and C_5-72 Groups were classified within the untreated category. An approximate method suggested by Snedecor (1956, p.271) was used to correct for the unequal numbers of Ss per group.

The test for IQ differences between groups receiving the same kind of treatment showed an insignificant effect, but one sufficient to prevent the pooling of subject with group variance, $F(6,123) = 1.51, .1 < p < .25$. Without pooling variance, the effects of treatments were significant, $F(2,6) = 10.96, p < .01$.

These results indicate first that there were significant differences in follow-up IQ scores between differently treated groups and,

second, that these differences were not due to group differences which occurred within treatment categories. The results are consistent with the hypothesis of a linear relation between amount of intervention and follow-up IQ score, but these results are qualified by the assumptions listed above.

An analysis of covariance was performed to account for differences in pretest scores. This analysis followed the same hierarchical design as the analysis of follow-up scores except that the C_5-72 Group, which was not pretested, was not included. The T68-I+II and T69-I+II Groups constituted the two-year treated category. The T67-I, T67-I+Short II, T67-I+Short II VISM, and T67-C_1+I Groups were classified within the one-year treated category. The C_2-67 and C_4-67 Groups were classified within the untreated category. As in the preceding analysis, Snedecor's (1956, p. 271) method was used to provide an approximate correction for the unequal number of Ss per group.

Cronbach and Furby (1970) recommended that, in such analyses, pretest scores be regressed to the means of the treatment groups before entering them in the analysis. There were several estimates of test-retest reliability of the Cattell from which to choose the regression coefficient. As a lower limit, VIP data have twice yielded a correlation of .60 between pretest Cattell and posttest I 1960 S-B IQ scores with Ns of 71 and 91. As an upper limit, Cattell (1940, p. 49) reported a correlation of .83 between a 30-month Cattell and a 36-month 1937 S-B. Somewhat arbitrarily, Cattell's (1940, p. 49) correlation of .71 between the 24-month Cattell and the 36-month S-B was used as an estimate of test-retest reliability. Using this estimate, pretest

scores were regressed to the mean of each group rather than to the mean of each treatment category because groups were the unit of sampling.

The results of this analysis were equivocal concerning a treatment effect on adjusted IQ scores. The degree to which groups receiving the same kinds of treatments obtained different adjusted scores (groups within treatment effect) was too large to allow pooling subject and group variances F (5,93) = 2.69, $p < .05$. Thus, the degree to which groups receiving different kinds of treatments obtained different adjusted scores (treatment effect) had to be tested for significance against the group error term rather than the error term for subjects. This resulted in a substantial reduction in the degrees of freedom for the test and a consequent loss of power to detect differences. The effect for treatments was not significant, F (2,5) = 3.07, $.1 < p < .25$. Thus, in this analysis, differences between treatments could not be discerned from differences between groups.

These results are generally consistent with, but not conclusive of, a hypothesis of increased benefit with increased program intensity up to two years of treatment beginning at age two. One analysis supported the hypothesis, and one failed to. Given the small number of degrees of freedom at which the treatment effect was tested, even a large effect is difficult to detect. The results are also qualified by initial assumptions that they were not produced by group sampling differences which may have been confounded with treatment differences, that age differences at program entry did not affect the results, and that obtaining follow-up IQ scores at times varying between 40 and 64 months after program entry did not systematically alter the outcomes

between different treatments. Unfortunately, these same factors could also have contributed to differences between groups. Only additional research can resolve these issues.

The data give support to a positive answer to Question 2 regarding similarity of gains regardless of educational level of TDs. Questions 3 (varied intensity) and 4 (long-term differences among T and C groups) have not, to use a legal model, received support "beyond a reasonable doubt" despite the possibility of acceptance based on a "preponderance of evidence".

Correlation Between IQs and Other Variables

The remaining question concerned the relation of follow-up IQ to other background and outcome variables. To trace the sources of IQ variability, and to identify the concomitants of IQ, the follow-up IQs of the Ss were correlated with ten other outcome and 51 background variables, a full list of which is available from the Verbal Interaction Project. The 51 background variables included demographic attributes of parents, grandparents, and family (e.g., education, occupation, health, family size), other characteristics of parents (e.g., father's employment and mother's style of dealing with home physical environment), other characteristics of index children (e.g., psychosocial problems), and program-related variables (e.g., number of VISM still usable).

The ten other outcome variables were Verbal IQ (PPVT), four gain scores for General and Verbal IQ, reading and arithmetic achievement (Wide Range Achievement Test, standard score), associative ability (WISC Digit Span Subtest), conceptual ability (WISC Block Design

Subtest), and socioemotional coping skills in the classroom as measured by the Child's Behavior Traits (CBT), a 20-item Likert-scale, VIP-developed instrument.

Table 5 lists only those Pearson r's, point-biserial coefficients, and Spearman rho's (corrected for ties in rank) significant at $p < .05$ or less for the two combinations of groups considered to be most relevant to the question of program effectiveness. Correlations not attaining this level are not reported. The combinations were composed of a two-year treated category which included all children who entered the program at age two and who received any treatment beyond one year of the MCHP (T68-I+II, T69-I+II, T67-I+Short II, and T67-I+Short II VISM) and of all comparison groups combined (C_2-67, C_4-67, and C_5-72). (Variations among Ns in Table 5 result from differing age-eligibility among measures, or from differences in the availability of data.)

Insert Table 5 about here

Eight of the fifteen non-chance correlations with follow-up General IQ were roughly similar between the two groups, those with Verbal IQ, General and Verbal IQ gain, reading and arithmetic achievement, and with conceptual ability. The correlations ranged from moderate to rather high and were in an expected direction for both groups; that is, IQ related positively to almost all other cognitive measures for the untreated as well as for the treated group, with one exception.

The one exception to the general tendency of cognitive measures to relate to follow-up IQ was the .16 correlation with associative ability

for the Two-Year Treated Group, in contrast to a .65 correlation of this variable with general IQ for the Untreated Group. This lack of relationship, especially when compared with the correlation between conceptual ability and IQ for this group indicates that full-scale IQ is determined more by Level II than by Level I intelligence (Jensen, 1968) for the Two-Year Treated Group. This observation must await support from replication in future years with a larger N and different sampling procedures.

The correlation between pretest IQ and follow-up IQ was larger for the treated than for the untreated groups (.54 vs .35). This probably does not of itself indicate that children who entered the program with the highest IQ scores derived the most benefit from the program. First, if the assumptions necessary for making a direct comparison between the correlations had been met, the difference would not be significant at the $p < .05$ level. Second, the comparison groups were, on the average, older than the treated children, hence a lower correlation with pretest IQ should be expected for them. Third, the conclusion is not supported by the correlations between pretest IQ and IQ gain which are -.22 (N=57) for the two-year treated group and -.11 (N=25) for the comparison groups (the negative values of these correlations may be attributed to the effects of regression to the mean). Thus, the difference between groups in correlations of follow-up with pretest IQ does not of itself support the conclusion that the MCHP differentially benefited those children with the highest pretest IQ scores.

The differing correlations of IQ with follow-up age for the two groups are still more difficult to interpret. The inverse relationship

significant for the Two-Year Treated Group probably resulted from an artifact; the groups which received two full years of the program and which had the highest IQ scores at follow-up entered the program later than the other groups and therefore contained the youngest Ss.

The remaining three correlation differences between the Two-Year Treated and the Untreated Groups raised some unanticipated questions about the Program's effects on the socioemotional development of its graduates, the cognitive relevance for treated children of preschool experience after the MCHP, and of the level of the mother's management of the physical home environment, (PED).

The child's socioemotional coping skills were based on blind teacher ratings on the CBT. There is little difference in the correlation of CBT with follow-up IQ between the treated and untreated conditions. Table 6 indicates that program graduates have obtained higher average scores on the CBT than have the untreated groups. It is possible that there were independent program effects on both CBT and IQ.

Insert Table 6 about here

It should also be noted that treated Ss have received scores which are clustered near the top of the CBT scale, thus restricting the range of CBT scores and, consequently, it may be expected, the correlation of CBT with other variables for treated Ss.

The lack of relation between attendance at preschool and follow-up IQ may also be an artifact. In all groups considered here except the C5-72 Group, virtually all Ss had some form of preschool attendance. This adds an important qualifier to any statement of program

effectiveness. If the program was effective, it must be added that the program plus some form of center-based preschool attendance was effective. That preschool attendance was not of itself effective is indicated by the performances of the C_2-67 and C_4-67 Groups who attended preschool but did not receive the program.

The difference between the correlations of IQ with PED Score for treated and untreated groups was the third unexpected divergency between the two groups and perhaps holds somewhat clearer meaning. The project-developed Physical Environment Description (PED) is considered to reflect maternal styles of home management. The ten Likert-scale items in the description include ratings for interview room features (e.g., spatial arrangement of furniture) and yield a summative score from 10 to 40. The split-half reliability coefficient for the PED, with Spearman-Brown correction, was .94. In score summaries, there was little difference among the mean scores for treated and untreated groups (64.5 and 61.3 for treated and untreated groups respectively). Thus it appears that for MCHP graduates the children's physical environment was irrelevant to their cognitive development as reflected in IQ, a lack of relationship not seen for the intellectual development of untreated children.

Feasibility of MCHP Variations

The summative data suggested that all two-year variations of the MCHP, conducted by paid or volunteer interveners with a wide range of education and prior skills, were more effective than the one year (T67-I) version, and that the abbreviated (T67-I+Short II and T67-I+Short II VISM) two-year versions may have been almost as effective

as the full two-year program (T68-I+II and T69-I+II). The 1968 changes from a shortened second year to a full two-year program were feasible and less costly than the 1967 MCHP, and, contrary to expectation, the full two-year program was easier to administer. The shortened two-year variations took as much staff time and effort, and caused more staff frustration than did the full program. When dyads were seen less often in home sessions in the second year, the mothers tended to forget appointments and withdraw from their involvment, requiring an unusual number of repeat visits and expenditure of ingenuity by the TDs and their supervisors. Since personnel and administrative time absorbed the main cost of the program, the full two-year MCHP (T68-I+II and T69-I+II) seemed the most feasible of all the two-year variations, with the possible exception of the T69-VISM Only treatment.

Discussion and Conclusions

This report of findings from a 1972-73 follow-up study of 138 VIP subjects began with five questions about the longitudinal cognitive effects of the VIP's early childhood intervention program, the Mother-Child Home Program. The one-year program had been followed by large short-term effects in 1967-68 after one year, when conducted with social workers as interveners. Essentially, the questions concerned IQ gain stability, the feasibility and effectiveness of utilizing volunteer and non-professional interveners, the amount of intervention necessary for maximum effect, the relation of these effects to other outcomes or events in the program graduate's life, and whether the program did indeed have a significant effect on the IQ's of treated children as compared with untreated subjects.

The data provided answers to the five questions supporting the first year's promise of the MCHP's effectiveness, albeit with some qualifications.

The IQ gains of graduates relative to controls were retained at least into first grade when the program was expanded to two years instead of the original one year, and especially if the second program year was as intensive as the first year. Otherwise, the short-term IQ gains diminished. The results for the full two-year program have thus been demonstrated to be stable over time.

At the same time that the program was expanded to its present two-year format, non-social worker TD's were introduced as interveners. Because the most stable results were obtained under these conditions, we have concluded as a practical matter that interveners (Toy Demonstrators) with a range of education from less than high school completion through college were found to be at least as effective as graduate social workers in producing both short-term and long-term effects. This finding greatly increased the feasibility of the program for application in other settings outside of the research project, a feasibility supported by the estimated annual unit cost of $400.

The MCHP graduates appeared to retain their short-term gains along a continuum of amount of exposure to the program, with groups in the full two-year version superior to other treated groups and to untreated groups, an observation first made by Bronfenbrenner (1974). Teachers' ratings of the children's school behavior also provide some indication that these differences are not completely test-specific. The full two-

year program was also found to be more feasible than abbreviated versions with the exception of the VISM Only variation. Since methods and results for the latter treatment could not be clearly interpreted, this treatment will be repeated, with a larger, randomized sample, starting in September 1974.

The results have thus far been encouraging, and, insofar as IQ scores may be taken as an index of level of cognitive functioning, the children who have received two full years of the MCHP are not laboring under the cognitive disadvantage usually associated with the demographic attributes which determined their acceptance into the program.

The conclusive attribution of this result to the effects of the program must await the long-term results of the groups currently enrolled in the program in which extraneous variables have been experimentally controlled. The data which are currently available, however, generally indicate either that the program is responsible for the effect or that no other easily identifiable factor is likely to have been responsible. The analysis of variance of follow-up IQ scores supported the conclusion of program effectiveness apart from the contributions of other factors. The analagous analysis of covariance with adjustment for pretest scores was equivocal regarding program effectiveness apart from differences which may have existed between groups. Due to the quasi-experimental design and the possibility of group differences, the tests for significance were performed at only 2,5 degrees of freedom, making an effect solely attributable to the program difficult to detect. Both of these analyses were also

qualified by initial assumptions regarding variables which may have been confounded with treatment. Several specific factors such as selective attrition from treated groups, tester bias and differential acceptance of the MCHP by differently treated and untreated groups have been eliminated as plausible causes of the difference in IQ between differently treated groups.

Correlations have not identified any factor which might plausibly have produced observed group differences. The follow-up IQ status of the two-year MCHP graduates appeared to depend on a very few antecedents beyond the program itself. Their eventual IQ status at follow-up seemed to be unrelated to the home physical environment within the range of environments observed. It is notable that no single SES-related background variable (such as low parents' education, father's absence, large family size, or receipt of welfare aid) was strongly related to IQ for either treated or untreated Ss. Thus within this low SES sample, variables usually clearly associated with IQ, when middle-income and low-income children are compared with each other after about age three, did not distinguish between high and low intellectual status. This lack of association may be due to the proportionate increase of error in measuring SES which accompanied the restriction of range of that variable. Because the data contain only one measure of SES (Hollingshead Index of Social Position) collected in an interview, its reliability cannot be directly estimated.

Correlations with other outcomes for both treated and untreated Ss support the observations of most investigators that, whatever IQ actually

measures, it predicts academic achievement. There are also provocative hints that the program fosters socioemotional coping skills separately from IQ, and that it may build a foundation for Level II (conceptual) intelligence.

The results described in this report illustrate rather dramatically a research dilemma almost inherent in social intervention evaluation: should the researcher, for valid social reasons, abandon experimental research according to the laboratory model, and thus leave the findings' internal validity in doubt? Or should the investigator somehow hew to the requirements of a true experiment, infinitely more difficult to follow in the field than in the laboratory, so as to extract (for himself and for society) the most reliable implications from the intervention study?

The present study initially presented the choice of obtaining data in a quasi-experimental design, with less than optimal control over some variables, or of not obtaining data at all. Now, because of the community groundwork laid by several years of being compelled to follow the first course of action, the VIP was recently fortunate enough to be able to start on the second. It is currently studying the effects of the MCHP in a true experimental design with the 1973-74 cohort, a sample of 50 Ss randomly assigned to E and C groups.

Other VIP studies either currently under way or about to be launched concern the effectiveness and feasibility of out-of-project replications (now 25 throughout the country), of graduates' school adjustment and related cognitive skills, of effects on and of mother-

child interaction, and of comparisons of Ss with siblings along behavorial and academic dimensions.

The findings from the 1972-73 follow-up study appear to provide enough support for the hypothesis that the MCHP prevents educational disadvantage to justify continuing the rigorous research necessary for social policy decisions preliminary to possible national implementation.

References

Bronfenbrenner, U. Is early intervention effective? A report on longitudinal evaluations of preschool programs. (Vol. 2) Washington, D.C.: DHEW Publication No. (OHD) 74-25, 1974.

Bruner, J. S., Olver, R. R., Greenfield, P. M. et al. Studies in cognitive growth. New York: Wiley, 1966.

Campbell, D. T. and Stanley, J. C. Experimental and quasi-experimental designs for research. New York: Rand McNally, 1963.

Cattell, P. The measurement of intelligence of infants and young children. New York: The Psychological Corporation, 1940.

Cronbach, L. J. and Furby, L. How we should measure "change" - or should we? Psychological Bulletin, 1970, 74, 68-80.

Jensen, A. R. A theory of primary and secondary familial mental retardation. In N. R. Ellis (Ed.), International review of research in mental retardation. (Vol. 4) New York: Academic Press, 1969.

Kirk, R. E. Experimental design: Procedures for the behavioral sciences. Belmont, Calif.: Brooks/Cole, 1968.

Klaus, R. A. and Gray, S. W. The early training project for disadvantaged children: A report after five years. Monographs of The Society for Research in Child Development, 1968, 33 (Serial No. 120, No. 4).

Levenstein, P. and Sunley, R. Stimulation of verbal interaction between disadvantaged mothers and children. American Journal of Orthopsychiatry, 1968, 38, 116-121.

Levenstein, P. Cognitive growth in preschoolers through verbal interaction with mothers. American Journal of Orthopsychiatry, 1970, 40, 426-432.

Levenstein, P. A message from home: A home-based intervention method for low-income preschoolers. Paper presented at conference sponsored by NICHD and the Rose F. Kennedy Center, Albert Einstein College of Medicine on "The Mentally Retarded and Society: A Social Science Perspective", Niles, Michigan, April, 1974.

Levenstein, P. The mother-child home program. In R.K. Parker (Ed.), The preschool in action. (2nd ed.) Boston: Allyn and Bacon, in press.

Roberts, J. Intellectual development of children by demographic and socioeconomic factors. Vital and health statistics data from the national health survey, series 11, No. 10. U.S. Department of Health, Education and Welfare, 1971, DHEW Publication No. (HSM) 72-1012.

Snedecor, G. W. Statistical methods applied to experiments in agriculture and biology. Ames, Iowa: Iowa State University Press, 1956.

Terman, L. M. and Merrill, M. A. Stanford-Binet Intelligence Scale 1972 norms edition. Boston: Houghton Mifflin, 1973.

Wechsler, D. Manual for the Wechsler Intelligence Scale for Children: revised. New York: The Psychological Corporation, 1974.

Table 1

Demographic Characteristics of Treated and Comparison Groups at Entry (Pretest)

Subject Group[a]		Child			Father			Mother		Family		
Entry Year	Program Designation	N	Entry Age	% Male	Years of School	% Occ. 6 or 7[b]	% Always Present	Years of School	% Rec. Welfare	Size	Residence Proj.[c]	Residence Non-pr.
Treated												
1967	T67-I	13	3	38.5	9.7	84.6	61.5	10.8	23.1	4.8	x	
1967	T67-I+Short II	7	2	42.9	9.9	100.0	85.7	11.4	14.3	6.9	x	
1967	T67-I+Short II VISM	6	2	83.3	10.3	100.0	83.3	11.2	33.3	5.8	x	
1967	T67-C_1+I	7	2&3	85.7	9.8	85.7	71.4	10.4	14.3	5.6	x	
1968	T68-I+II	21	2	33.3	10.4	61.1	57.1	10.3	42.9	5.8	x	
1969	T69-I+II	23	2	56.5	11.7	50.0	68.2	10.6	34.8	4.8	x	x
1969	T69-VISM Only	6	2	50.0	10.5	100.0	33.3	10.5	66.7	4.3		x
Comparison												
1967	C_2-67	15	2&3	80.0	10.5	73.3	93.3	11.2	6.7	5.2	x	
1967	C_4-67	10	4	50.0	10.3	71.4	30.0	10.6	60.0	5.9	x	
1972	C_5-72[d]	30	6	50.0	9.6	100.0	23.3	9.8	82.1	5.8		x

[a] "Subjects" defined by inclusion in most recent follow-up test.

[b] "Occupation 6 or 7": Hollingshead Scales 6 or 7 (Unemployed, unskilled, semi-skilled).

[c] "Proj": residence in low-income housing project. "Non-pr.": residence out of low-income project.

[d] Entry at Follow-up D, 1972-1973 as "after-only" group.

Table 2

Number of Ss and Percentage of Original Group Retained in Program and Follow-up

Subject Group	Time Period: Entry N	After First Program Year N	After First Program Year %	After Second Program Year N	After Second Program Year %	Follow-up N	Follow-up %	Follow-up: Nos. of Mos. Post Entry
T67-I	18	18	100	-	-	13	72	64
T67-I+Short II	9	8	89	8	89	7	78	64
T67-I+Short II VISM	6	6	100	6	100	6	100	64
T67-C_1+I	10	9	90	-	-	7	70	64
T68-I+II	30	27	90	21	70	21	70	52
T69-I+II	33	31	94	25	76	23	70	40
T69-VISM Only	12	12	100	9	75	6	50	40
C_2-67	19	-	-	-	-	15	79	64
C_4-67	10	-	-	-	-	10	100	64
C_5-72	-	-	-	-	-	30	-	0

Table 3

Pretest and Posttest IQs for Treated Groups, 1967-1973

(Cattell, Stanford-Binet, WISC)

Subject Group: Entry Year	Program Designation	Variable	Number of Months After Pretest: Pretest	8 Post I	20 Post II	28	40	52	64
1967	T67-I	N[a]	13	13	-	13	13	-	13
		IQ	89.5	105.9	-	103.6	99.1	-	95.0
		SD	11.8	16.5	-	13.7	11.8	-	10.3
		Gain[b]		16.4*	-	14.1**	9.5**	-	5.5*
		Age(Grade)[c]	3	3½	-	5(K)	6(1)	-	8(3)
1967	T67-I+Short II	N	7	7	6	7	-	7	7
		IQ	82.6	101.1	105.0	103.6	-	100.6	100.3
		SD	6.6	10.5	13.7	11.8	-	10.4	11.6
		Gain		18.5	21.7	21.0**	-	18.0**	17.7**
		Age(Grade)	2	2½	3½	4	-	6(1)	7(2)
1967	T67-I+Short II VISM	N	6	5	5	6	-	6	6
		IQ	82.8	101.8	102.6	98.5	-	98.8	95.7
		SD	8.1	11.1	9.1	7.0	-	12.4	15.8
		Gain		21.2	22.0	15.7*	-	16.0*	12.9
		Age(Grade)	2	2½	3½	4	-	6(1)	7(2)
1967	T67-C_1+I	N	7	7	6	7	-	7	7
		IQ	88.0	88.9	101.0	105.0	-	104.3	107.0
		SD	10.5	8.2	11.3	8.8	-	5.9	6.7
		Gain		0.9	11.5	17.0*	-	16.3**	19.0**
		Age(Grade)	2,3	2½,3½	3½,4½	4,5(K)	-	6(1),7(2)	7(2),8(3)
1968	T68-I+II	N	21	21	19	21	21	21	-
		IQ	90.4	101.8	108.9	108.3	107.3	105.4	-
		SD	9.1	9.0	8.5	11.1	11.6	13.0	-
		Gain		11.4**	17.4**	17.9**	16.9**	15.0**	-
		Age(Grade)	2	2½	3½	4	5(K)	6(1)	-
1969	T69-I+II	N	23	22	23	-	23	-	-
		IQ	88.8	105.6	108.2	-	113.3	-	-
		SD	13.8	16.5	15.6	-	15.9	-	-
		Gain		17.4**	19.4**	-	24.5**	-	-
		Age(Grade)	2	2½	3½	-	5(K)	-	-
1969	T69-VISM Only	N	6	6	6	-	6	-	-
		IQ	87.0	98.0	96.7	-	103.2	-	-
		SD	7.4	8.3	6.5	-	10.1	-	-
		Gain		11.0	9.7	-	16.2**	-	-
		Age(Grade)	2	2½	3½	-	5(K)	-	-

* $p < .05$

** $p < .01$ (t test, two tailed).

[a] Ss in latest test data.

[b] Calculated from pretest IQ.

[c] School grade for 50% + of group.

Table 4

Pretest and Posttest IQs for Comparison Groups, 1967-1973

Subject Group			Number of Months after Pretest						
Entry Year	Program Designation	Variable	Pretest	8	20	28	40	52	64
1967	C_2-67	N[a]	15	10	-	15	7	8	15
		IQ	91.3	93.4	-	92.5	104.3	88.9	95.5
		SD	8.6	9.4	-	18.4	13.6	25.2	22.3
		Gain[b]	-	2.1	-	1.2	13.0**	-2.5	4.2
		Age(Grade)[c]	2,3	2½,3½	-	4,5(K)	6(1)	6(1)	7(2),8(3)
1967	C_4-67	N	10	-	-	-	-	10	10
		IQ	91.0	-	-	-	-	96.3	97.0
		SD	8.2	-	-	-	-	10.3	10.3
		Gain		-	-	-	-	5.3	6.0
		Age(Grade)	4	-	-	-	-	8(2,3)	9(3,4)
1972	C_5-72	N	-	-	-	-	-	30	-
		IQ	-	-	-	-	-	91.0	-
		SD	-	-	-	-	-	11.5	-
		Gain		-	-	-	-	-	-
		Age(Grade)	-	-	-	-	-	6(1)	-

* $p < .05$
** $p < .01$ (t test, two tailed).
a Ss in latest test data.
b Calculated from pretest IQ.
c School grade for 50% + of group.

Table 5

Correlations with General IQ at Follow-up; Full and Part Two-year Treated Groups, and Untreated Groups. (Pearson's and point-biserial r and Spearmans rho corrected for ties in rank.)[a]

Variable	Total Two-year Treated Group			Total Untreated Group		
	N	r	rho	N	r	rho
Pretest Cognitive Variables						
Pretest General IQ	57	.54**		25	.35	
Pretest Verbal IQ (PPVT)[b]	55	.27*		24	.35	
Background Variables						
Preschool Attendance (yes, no)	57	.04		55	.30*	
Home "Physical Environment Description" at Follow-up	54		.13	55		.34*
Age at Follow-up	57	-.42**		55	.21	
Follow-up Cognitive Variables						
Verbal IQ (PPVT)[b]	57	.58**		55	.65**	
General IQ Gain, Pretest to Follow-up	57	.70**		25	.89**	
Verbal IQ Gain, Pretest to Follow-up	56	.36**		24	.47*	
Reading Achievement (WRAT)	29	.45*		54	.41**	
Arithmetic Achievement (WRAT)	29	.64**		54	.57**	
Conceptual Ability (WISC Block Design)	22	.62**		54	.48**	
Associative Ability (WISC Digit Span)	23	.16		54	.65**	
General IQ Gain, Post 1 to Follow-up	56	.60**		10	.73*	
General IQ Gain, Post 2 to Follow-up	54	.56**		NA	NA	
Follow-up Teacher's Rating of Socio-emotional Competence						
School Coping Skills (CBT)	56		.20	55		.29*

* $p < .05$

** $p < .01$

[a] Variables are those of original list of 62 whose correlation with follow-up IQ is non-zero at the $p < .05$ level.

[b] Not standardized for two-year olds.

Table 6

Mean Teachers' Ratings of Socio-emotional Competence (CBT) at Follow-up

Treated and Comparison Groups

Variable	Treatment Group									
	Full Two-year Treated		Part Two-Year Treated		One-year Treated	Placebo + One Year	VISM Only	Comparison		
Group	T68-I+II	T69-I+II	T67-I+Short II	T67-I+ Short II VISM	T67-I	T67-C_1+I	T69-VISM Only	C_2-67	C_4-67	C_5-72
Entry Year	1968	1969	1967	1967	1967	1967	1969	1967	1967	1972
N	21	22	7	6	13	7	5	15	10	30
CBT	77.2	70.5	76.1	64.7	64.0	65.8	69.8	61.9	59.4	66.1
SD	16.7	12.4	9.5	23.5	19.5	16.3	16.3	21.4	19.6	15.9