DOCUMENT RESUME

ED 199 358 UD 021 314

AUTHOR Slaughter, Diana T.
TITLE Longitudinal Assessment of the Intelligence of Black Infants, Ages 22 to 41 Months. Revised.
SPONS AGENCY National Inst. of Child Health and Human Development (NIH), Bethesda, Md.
PUB DATE [79]
GRANT NICHD-5-R01-HD08533
NOTE 49p.: Revised version of ED 185 137.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Black Mothers: *Blacks: *Cognitive Development: Disadvantaged: *Intelligence Quotient: Language Acquisition: Language Skills: *Lower Class Parents: Parent Background: Parent Child Relationship: *Parent Education: Verbal Ability: *Young Children
IDENTIFIERS Cattell Infant Intelligence Scale; McCarthy Scales of Childrens Abilities: Peabody Picture Vocabulary Test

ABSTRACT

A study was conducted using a longitudinal, multivariate, multimethod approach to determine the effects of parent education programs on black mothers and children from lower socioeconomic status backgrounds. Black infants from lower status communities, ages 22-41 months, were evaluated during the early intervention study. Instruments used included the Peabody, Cattell, and McCarthy psychometric intelligence tests. Over time, all groups declined significantly in average IQ score, but program groups, by comparison to the control group, did much less so on the Cattell-McCarthy sequence. At the third and final testing, Peabody and McCarthy mean IQ scores differed an average of 25 points across the three groups. Program children, generally, were superior to controls on the McCarthy measure of expressive language. The findings suggest the effectiveness of some types of intervention programs, and also suggest the different socialization processes to which black and white children are exposed. (Author/APM)

# Longitudinal Assessment of the Intelligence of Black Infants, Ages 22 to 41 Months

Diana T. Slaughter

School of Education

Northwestern University

Revised

Running Head: Assessment of Black Infant Intelligence

# Abstract

Black infants from lower status communities (N=57), ages 22-41 months, were evaluated during a longitudinal early intervention study using the Peabody, Cattell, and McCarthy psychometric intelligence tests. Over time, all groups declined significantly in average IQ score, but program groups, by comparison to the control group, did much less so on the Cattell-McCarthy sequence. At the third and final testing, Peabody and McCarthy mean IQ scores differed an average of 25 points across the three groups. Program children, generally, were superior to Controls on the McCarthy measure of expressive language.

# Longitudinal Assessment of the Intelligence of Black Infants, Ages 22 to 41 Months

## Introduction

Since the early 60's, infant and preschool tests have been widely used to evaluate the impact of early intervention programs. For nearly as long a time, there has been considerable criticism of the use of cognitive performance tests as criteria for the success or failure of these programs. There are major measurement problems associated with assuring internal and external validity when such psychometric tests are designed for, and administered to, immature developing organisms. Children's rates and level of conceptual development, for example, may show considerable intraindividual variability across different functions and abilities. These difficulties may be exacerbated when the study participants are black, or otherwise different from the children on whom the tests are typically standardized. Because these tests, in the latter months of early childhood, become predictive of future school-related achievement and intelligence test performance, they continue to be widely used. In fact, their use is so pervasive, that an individual investigator risks being unable to enter scientific discourse without first establishing population comparability through such test results. The results become, in effect, indicators of who, demographically, study participants are, rather than of what they can, skill or ability-wise, do.

Although we were cognizant of the limitations of infant tests as indicators or predictors of level of cognitive development, we found it useful to first demonstrate replication of the results of other infant intervention studies in our own research before continuing to a more intensive study of other, we believed, more critical process variables which could be more naturally and directly linked to the short-term program goals identified by the parent education models used. However, because our results were somewhat unique, they merit further discussion. Thus, this report presents the infant testing results of a study which used a longitudinal, multivariate, multimethod approach[1] to the study of Black mothers and children from lower socioeconomic status backgrounds who participated in parent education programs designed to enhance early childhood development.

## Background

It is known that social class differences in infant test performance emerge somewhere between the second and third years of life, and it is also believed that this is because of the increased importance of the differentiated role of language in the tests themselves during this period (Golden and Birns, 1976). Broman, Nichols, and Kennedy (1975) have recently reported the results of a longitudinal study of 26,760 black and white mothers and their children which extended from the prenatal period through age 7. The researchers found that prenatal ratings of socioeconomic status and mother's education correlated better than the Bayley (8 months) with white children's Binet IQ at age 4. Though Bayley scores predicted Binets equally well in the two groups, socioeconomic status was a poorer predictor of Binet scores among the blacks. Only a neurological measure of delayed development obtained at the end of the young child's first year of life added to this list of the most consistently useful predictors of the Binet IQ at age 4 in a study which initially emphasized 169 indices of maternal and child biological and physical health, as well as other aspects of familial psychosocial functioning.

Although social class differences in infant test behaviors are typically not observed before age two, differences in how mothers and infants from different social class strata use language as a communicative act can be observed well <u>before</u> age two. Middle class infants are more responsive to maternal vocalizations, quieting during such times, and then vocalizing when mother does not talk. As early as three months (Lewis and Freedle, 1973) there is evidence of a process of "turn-taking" between white middle class mothers and their infants which is not evident between lower class mothers and their infants. Middle class mothers of three- and ten-month-olds have also been observed to be more verbally responsive (Lewis and Wilson, 1972; Tulkin and Kagan, 1972) with their babies. They are more likely to respond to a vocalization with another one. There is, however, no clear evidence that middle class mothers talk more overall to their infants than do lower class mothers.

While existing infant tests vary across the sensorimotor period in their emphasis on language functions, it appears that from birth onward the early precursors of the communicative act can be observed between child and mother, and that individual differences in styles of communication are evolving which are related to the social class background of the families. Recent research (Bloom, 1973; Nelson, 1973) has made this even more apparent since individual differences are noted in the first fifty words used by young children, differences which have been linked in a preliminary way, by at least one author (Nelson, 1973), to differences in maternal parenting style. While we cannot posit a "critical period" for development of the language function at this time, we can state that the first three years of life, particularly the second and third years, constitute a highly sensitive period and that parents differ by social class ~~positioning~~ in their responsitivity to this period and to this aspect of their infant's development.

## Hypothesis

One major hypothesis, dictated by earlier literature in this field, related directly to infant test outcomes:

> Children of participating mothers will demonstrate a greater overall intellectual gain over a two year period than children of non-participating mothers; the gain will be primarily a function of greater encoding skills, relative to verbal expression of thoughts and general word fluency (Stott and Ball, 1965).

Essentially, this hypothesis predicts that program children will be more advanced by comparison to control children at the end of year two of the programs in overall IQ score, and that this superiority will be due to a superiority in verbal production in contrast to verbal comprehension.

## Method and Procedures

### Programs

The research emphasized three conditions: (1) Levenstein's Toy Demonstration program, (2) Mothers Discussion group program, and (3) a Control group. The Discussion group was initially conceived by us to be a conservative control to the home-based Toy

Demonstration program.

The Toy Demonstration program utilizes a female "Toy Demonstrator" who, in effect, each week <u>models</u> play with the mother's child in the mother's presence using preselected toys which presumably have high stimulus value for young children. Two half hour visits per week over a school calendar year are made with each dyad, the first for introduction of a new toy which the child is allowed to keep, and the second for review. Both mother and child are participants in this program, whose primary purposes are to enable the mother to observe the value of interactive play with an adult for a developing child, and how this 'play' can be stimulated, extended, and enriched. It is assumed that mothers will privately begin to initiate and/or sustain more of such experiences with the children in the Demonstrator's absence. Rationalized in terms of its long range applicability to children's successful school achievement, this program has been depicted by Bronfenbrenner (1975) as one of the leading parent education programs.

In contrast, the Discussion group emphasizes the mother, not the child, as the primary initial participant. Specific toys are not considered as essential to the program. Rather, this program emphasizes mutual sharing of information and experiences of personhood, home and family life, of mothers of similar aged children. It is assumed that specific interest in those activities, such as play, which might enhance this development will develop secondarily, as a result of some of the mothers' own perceived personal needs being met through the group experience. At each of three housing sites, two groups of ten mothers each were initially formed and maintained where possible. The groups met weekly for about two hours; on occasion this time was used for outside field trips and consultation.[2]

The Control group received only the prescribed toys on a weekly basis, as developed in the Levenstein model.

Measures

The measures of infant intelligence used were the Cattell Infant Scale administered at times 1 and 2, the McCarthy Scales of Mental Abilities administered at time 3, and the Peabody Picture Vocabulary Test administered at times 1, 2, and 3. Both the Cattell and the Peabody had been used in previous intervention research. Levenstein has used the Cattell, followed by the Binet and, initially, we anticipated a similar testing program since the Cattell is frequently viewed as a downward extension of the Binet. In fact, Binet items are interspersed into the Cattell between 22 and 30 months. However, during the study we decided to use the McCarthy instead of the Binet at the third time point because of our increased sensitivity to the need for an assessment commensurate with the differential development of children's abilities at these ages.

Cattell Infant Intelligence Scale.

The first publication of this widely used test for infants between the ages of 3 to 30 months was made in 1940 (Cattell, 1940). Influenced by the Gesell tests, a major goal of its developers was to introduce items which could be objectively scored and which would be of pleasurable interest to all young children so as to better differentiate among them as to their abilities. In addition, consistent with the deterministic view of intelligence which prevailed during that era (Kamin, 1975), the developers wished to use items which were minimally, if at all, impacted by "home training."

The standardization sample children (N=274, ages 3-36 months) were white, of predominantly "North European stock" (a requirement of the developers), and of essentially middle class backgrounds. They participated in a Harvard clinic which charged $50 for enrollment. Most of their fathers had permanent jobs as policemen, clerks, storekeepers, etc.

Cattell reported split-half reliability coefficients between the ages of 18-30 months of .71 to .90. She also reported predictive validity correlations with the Binet at 36 months for similar earlier ages between .67 to .75. Harms and Spiker (1959) have reported test-retest correlations within 3-7 days of .82-.93 for 80 white children, ages 17-29 months. The lower correlation was obtained when children were evaluated twice by the same examiner.

Golden and Birns (1967) have reported the only normative data on black children currently available. Using a sample of 186 children of a New York Well Baby Clinic from three social class groupings at 12, 18, and 24 months of age, they reported mean Cattell IQ scores for each class/age subset available to them. Class A (families on welfare) had mean IQ scores of 111.3, 108.1, and 98.5, respectively, at these ages. Class B (both parents never employed at other than unskilled or semi-skilled occupations, and having no more than high school educations) had IQ scores of 112.1, 110.0, and 99.7. Finally, Class C (either mother or father had schooling beyond high school and/or employment at skilled or professional occupations) had IQ scores of 107.9, 109.4, and 99.7. Except for one group in Class B where the sample size was 26, sample sizes in each group were 20. No significant social class differences in test performance were found, but the lower status children were reported to be more difficult to engage in the testing session; mothers were present during the testing. Race of the examiner was not reported.

In a later article, Birns and Golden (1972) reported respective correlations of .13 and .60 between the 18 month and 24 month Cattell and the Binet at 36 months (N=89).

Peabody Picture Vocabulary Test

Somewhat less information is currently available on the Peabody Picture Vocabulary Test, an instrument used widely in early intervention research, and used by Levenstein in an evaluation of the Toy Demonstration program.

The Peabody seems, on face value, to be essentially a test of selected word knowledge, as these words may be applied progressively first to concrete phenomena represented in pictures, and later to more abstract concepts also represented by pictures. Words also become increasingly removed from everyday experience as the test progresses, since it is designed for use with children and adolescents ranging in age from two years, six months, to eighteen years of age. Children in our age range had to respond to very concrete descriptive words such as identifying objects (e.g., animals, persons, parts of persons, including clothing as well as body parts) or actions (e.g., body movements or typical adult activities which might be observed) named by the examiner in order to obtain an average or above average score. Because the child has only to listen to a word produced by the examiner, and then point to the one picture, of four presented simultaneously, which best represents that word, the PPVT is often identified as a measure of receptive language (verbal comprehension). No verbal production is required for a child to perform well on this instrument.

The PPVT was normed on 4,012 white children and adolescents from Nashville between the ages of 30 months and 18 years in 1958. At ages 2-6, 3-0, 3-6, and 4-0 years sample sizes were 92, 92, 119, and 110 respectively. Alternate form reliabilities reported for these ages are .72, .77, .81, and .75. Validity correlations with the Binet, using raw mental age scores, are reported to be between .60 and .87, with a median r of .71. However, the author of the 1965 Expanded Manual states:

> Here, as in all the statistical validity data, the lowest correlations are found when IQ scores of a restricted group of subjects on the intellectual continuum were used, and highest values were obtained when . . . scores were used for children over a wide age range, across the full spectrum on intellect. What is needed are data on correlations of PPVT and Binet IQ scores by age levels for subjects falling across the full intelligence range. (p.33)

At the 01 testing with this instrument, we could not always set basals on our children; the average age of 22 months was considerably below 30 months (2-6 years). Instead of setting a basal (8 consecutive correct responses), we discontinued the test whenever the child reached a ceiling (6 errors in eight consecutive responses) and used the resultant raw score (ceiling less errors) as the basis for further calculations.

McCarthy Scales of Mental Abilities.

This test provides a General Cognitive Index (GCI) or IQ score, as well as standardized information on the differential abilities of children, ages 2-1/2 to 8-1/2. The five subscales of the test are: Verbal, Quantitative, Perceptual-Performance, Memory, and Motor Development.[3]

The first publication of this test was in 1972. Thus, by comparison to the Cattell and PPVT, it is relatively recently normed. The standardization sample of about 100 children at each age level, 2-1/2 to 8-1/2 years, with equivalents for the GCI extending from ages 16 months to 12 years, 6 months, included white and nonwhite (about 16%) children (154 of the 170 nonwhites were black) from varied social strata (father's occupation as primary criterion) and regional backgrounds.

Split-half reliabilities range from .91 to .96 in the ages from 2-1/2 to 4-1/2 years on the GCI, and comparable reliabilities are reported for each of the five subscales at each level (Ns range from 100-104 across age levels).

As test distributor, the Educational Testing Service lists several publications that positively review this instrument, often with black populations. However, to our knowledge, with the exceptions to be reported below, this age group of black infants (mean=41 months) is among the youngest to which the McCarthy Scales have been administered.

The important exceptions appear to be a series of studies by Kaufman and Kaufman (1973, 1975a, 1975b). They report studies with black and white children using data from the full age ranges of the original standardization sample. On the McCarthy, no significant

IQ differences are reported between blacks and whites between ages 2-1/2 to 5-1/2 years. Separate factor analyses performed on black (N=124) and white (N=688) children, ages 3.0 to 7.6 years, yielded similar factor structures for both groups.

Social status differences are significant and in favor of children of higher status families ($p < .01$). In the study (1975b) which contrasted social status and race, the authors reported differences between working class blacks and whites, favoring the latter. Conversely, no differences were found between middle class blacks and whites. Social status groupings and IQs for blacks and whites were, respectively, (a) Professional and Technical occupation of father: blacks 105.4 (N=11), whites 109 (N=145); (b) Managerial and Clerical Sales: blacks 99.9 (29), whites 102.8 (232); (c) Skilled workers: blacks 97.8 (25), whites 99.8 (213); (d) Semi-skilled workers: blacks 88.2 (62), whites 98.2 (234); (e) Unskilled workers: blacks 89.9 (27), whites 94.7 (38). Categories (a)-(c) were defined as middle class, while categories (d)-(e) were defined as working class. This same pattern of differences was also found with the five subscales. However, no data on the exact ages of the children in each subsample were provided.

Sample Selection

Participants were black, and from lower income communities in the Chicago area. Sample selection followed a similar format at each housing project site. At each site, nine target areas were identified according to degree of proximity to the base offices of United Charities' Family Service Bureau, and the internal organizational character of the building units in that site. About one-third of the participants in each of the three programs were recruited by representatives of each of the three treatment conditions from their individual target areas. These target areas were randomly chosen. We had assurances from the Chicago Housing Authority that the buildings did not contain marked differences among residents.

Each mother was offered the opportunity to participate in only one group. We believed that it would be less confusing to parents, and less threatening if we randomized by target area, rather than by subject. Children were between ages 18-24 months at the onset of the study in October, 1974.

The first child test assessment was completed about three months after programs began: 2/15/75; the second was conducted between 7/23/75-9/18/75; the third between 5/6/76-6/30/76. Programs officially ended 6/1/76.

At the first testing, children at all sites in all groups ranged in age from 20.3 to 24.2 months (M= 22 months). At the second, they ranged from 30.2 to 33.0 months (M=32 months). At the third, they ranged from 40.4 to 42.5 months (M=41 months).

We deliberately chose the ages of the children in our study. It is known that social class differences in intellectual performance emerge around ages 2 1/2 to 3 years, a time in which the performance tests appear to also become increasingly educationally relevant. At this time, earlier observed racial differences which favor black children seem to disappear (Freedman, 1974).

Assessment Procedures

The psychometricians were black females who had a least a B.A. degree in psychology or education. All had prior experience working in day care centers or nurseries as teachers. Their ages ranged from 23 to 37 years. All participated in an organized training program developed and conducted by supervisory staff. Several pilot testings were part of the four-week program.

Testing of study children was conducted at their housing site in a specially prepared room for this purpose. We hoped to maximize the child's and mother's familiarity and comfort in a new situation by evaluating in a room and location similar to that in which they lived. Special (round, wooden, low height) child testing tables were constructed by project staff for use by child and examiner. The mother of the child, but no other person, was to be present at each testing. At the beginning, background information for both test booklets was obtained from the mother. Then she was instructed to sit to the side, slightly to the rear of the child and observe, but not actively participate, during the session. Each mother was told to expect that, because each test covers a wide age range, there would be some items her young child could not do. We believed the proximity of their mothers would be especially important for the security of children in our age ranges, but we also wanted to anticipate and assuage some of the mothers' anxieties about how their children would perform.

During the first and second time periods, the Cattell was administered first, followed by the PPVT (Form A at 01 and 03 assessments; Form B at the 02 assessment). We attempted to emphasize the "game-like" quality of the tests, but each test was strictly administered according to guidelines developed in the test manual. During the third time period, the PPVT was administered first because of the length of the McCarthy. We wished to overcome the impact of fatigue on PPVT performance, and hoped that children were, by that time, more comfortable with the PPVT's "question-answer" format.

At the second and third assessments, testers completed one to two paragraph behavior profiles immediately following each child's session. These profiles were to include:

1. a brief statement of how responsive the child seemed to be to the examiner at the start of the testing, how willingly he entered into the testing room, and so forth.

2. a brief statement as to how the child performed on each test, as well as how he behaved over time during the course of the examination, noting strengths and weaknesses.

3. a brief statement about the mother's behavior during the examination, as well as any pertinent questions or comments about the testing which she addressed to the examiner.

4. a brief statement, using (3.) above as a guide, for the child.

Testing in the first and second sessions took about 30-40 minutes per child, but up to 60 minutes at the third session.

## Results

Initially, there were 41, 53, and 38 participants in the Toy Demonstration, Discussion, and Control Groups respectively, for a total N=132. Of 49 attrited subjects, 26 left in 8/75 following the conclusion of year one; 13 of the 26 were Discussion group participants. The total number of attrited cases for the Toy Demonstration, Discussion, and Control groups were, respectively, 15, 27, and 5 (=39.7 percent). The final base sample is 83, 26 in the Toy Demonstration Group, 26 in the Discussion Group, and 31 in the Control Group. The differential rates of attrition, 36.6, 50.9, and 13.0 percent, respectively, most likely reflect program demands. Discussion group pairs had the most to do in order to participate in that program since those mothers constituted the only group of mothers to have to leave home to have services rendered. Control mothers had the least to do.

Table 1 presents some important background data on the participants in the groups. Groups did not differ statistically on average child age at testing, maternal years of education, or maternal age. Groups also did not differ statistically in children's average birth order or age in months at the onset of the programs. No study group varied significantly from others in the amount of current or past maternal work experience, reported family income, maternal health, or length of time separated from the typical child during infancy. No group had a significantly greater number of other adults or children in the household. However, there are some differences between groups in children's access to adult males. Mothers of control children, in particular, are more likely than either of the other groups to report that the significant male in the child's life is not the child's natural father, and does not live in the child's home ($X^2(2) = 8.851$, $p < .02$). Nevertheless, the majority of mothers across all groups report that the children do have at least one significant adult male available to them.[5]

Insert Table 1 about here

Table 2 presents comparative test data between the main study children and those who attrited after the first assessment. The children who remained in the program do not, as a group, differ significantly from those who left, on either the Cattell Mental Age or the Cattell IQ.

Insert Table 2 about here

Table 2 also indicates an eleven point discrepancy between Cattell IQ and PPVT IQ scores. This discrepancy could not, we found, be attributed to either the recency of norms generated by the PPVT or the shift in sample sizes. For example, at a later age, the children performed better on the McCarthy, than the PPVT, though the McCarthy is the most recently normed of all three tests. Rather, the two tests probably measure diverging child abilities at these ages.

Table 3 presents means and standard deviations for all children tested (ATSs) and for children with complete data (CDSs) in either of the two test sequences (PPVT 01-03; Cattell 01, 02, McCarthy 03). Analyses of variance were conducted only on samples with complete data since, generally, results were comparable for the two groups.

---

Insert Table 3 about here

---

Significant mean differences, favoring the program groups over controls, were obtained ($p<.05$) on the McCarthy IQ at the end of the study. However, two-way analyses of variance, using group and housing site as independent variables ($N$=56) and the McCarthy IQ as the dependent variable resulted in nonsignificant main effects and interactions. When the McCarthy Verbal Subscale was used as the dependent variable in a similar analysis, we obtained a significant main effect for group ($F$=3.452, $p$=.04), but not for site ($F$=0.905, $p$=.41) and interactions were nonsignificant ($F$=1.560, $p$=.20). Clearly, the initial superiority of program to control groups on the McCarthy IQ is predicated upon this one subscale.

No differences between groups were observed on any other subscale. Figure 1 graphically presents these data. No differences were observed at any other assessment period or with any other instrument except the first administration of the PPVT. Program groups were initially superior to controls on this measure at the first assessment ($p<.05$); over time this difference was not sustained.[4]

Insert Figure 1 about here

We found evidence of a significant linear decline in both test sequences, the alternate form administrations of the PPVT, and the Cattell-McCarthy sequence. When we observed the effects of program, site, and time on the PPVT, using a repeated measures analysis, we found a highly significant effect for time ($F=83.96$, $p<.001$), and a moderately significant effect for site x time ($F=2.942$, $p<.025$). We did not perform similar analyses on the Cattell-McCarthy sequence. We could not argue strict comparability between the Cattell and the McCarthy and some writers (eg. Achenbach, 1978) question whether a repeated measures design should be applied when there are only two data points (i.e. Cattell 01 and Cattell 02). However, as noted by Figure 1, the results appear similar. Specifically, with advancing chronological age many of the children, particularly control children, did not advance according to existing mental age norms on these instruments. Somewhere between 22 and 41 months of age, these black children lost ground by comparison to existing norms.

Data in Table 3 indicate that the absolute differences in means between the two test sequences also increased over time. Differences between means at time 1 range from 8-16 points, at time 2 from 12-19 points, and at time 3 from 21-28 points. Harms and Spiker (1959) tested 80 white children average ages 17, 21, 25, and 29 months using the Kuhlmann and the Cattell. They found average performance on the Kuhlmann to be 16 points lower than the Cattell in all four age groups (N=20 per group). Our results are comparable with the Cattell and PPVT at ages 22 and 31 months, since average discrepancies favoring the Cattell at these ages were 12 and 17 points. However, by 41 months, when we shifted to the McCarthy, this discrepancy had increased to an average of 25 points.

Harms and Spiker (1959) also report average intercorrelations between their tests for each age group. These range from .71-.84. Ours are considerably lower. The correlation between the Cattell and the PPVT at time 1 is .50 ($p<.01$); at time 2 it is .39 ($p<.01$). The correlation between the McCarthy GCI and the PPVT at time 3 is .41 ($p<.01$). By comparison to the tests in the Harms and Spiker sample, there is considerably less evidence of concurrent validity between the two tests used in this sample.

Table 4 presents the correlations for all measures used on all children (ATSs).

Insert Table 4 about here

Our stability correlations (Cattell 01 and 02, PPVT 01-03) do compare favorably with others reported within related age ranges. Bayley (1933) has reported correlations ranging from .04 to .09 between scores at 0-3 months and those at 18-36 months. More recently, McCall (1972) has reported similar instability with Fels data, and

Lewis (1972) reports stability correlations between -.25 and .54 on the Bayley Mental Scale (median r=.26) between 3 to 24 months of age (N=20). He found the intercorrelations of the Escalona-Corman scales to range between -.13 to .48 over the same period of 3=24 months, with a median r of .07. In this sample, stability correlations range between .30 and .40. However, test scores at 31 months are more highly related to scores at 41 months, than scores at 22 months are to scores at 31 months. Children between ages 2 1/2 to 3 1/2 years perform more similarly on these tests than children between ages 1 1/2 to 2 1/2 years.

Both Lewis (1972) and Harms and Spiker (1959) report finding a significant correlation between spontaneous "meaningful" speech production and infant IQ tests around 21-24 months. Possibly, our greater obtained stabilities between 32 to 41 months, by comparison to 22 to 32 months, reflect a similar underlying process. As the children matured, speech production became more salient to the overall continuity of their performances.

Two longitudinal studies have examined the predictive validity of the Bayley with black infants. Goeffney (1971), using black and white children (N=229 Black, 397 White), found modest correlations between the Bayley at 8 months and the WISC at age 7 years (.13 ($p \leq .05$) and .23 ($p \leq .01$), for blacks and whites respectively). Broman et al. (1975), using black and white children (N=14,550 Black, 12,210 White) found correlations of .24 and .38 between both Bayley scales and the Binet at age 4. The present examination of an important intervening period between ages 8 months and 4-7 years suggests that, while there is greater predictive validity than previous reports relative to black infants from predominantly lower socioeconomic status communities indicate, the general pattern of the data is most similar to that reported for similar aged peers, regardless of socioeconomic and racial background.

Most striking in Table 4, however, is the tendency for measures given at any one assessment session to be either more or as highly correlated with each other as the same measure given at a subsequent session. For example, Cattell 01 correlates .50 with PPVT 01, but .30 with Cattell 02, while PPVT 01 correlates .30 with PPVT 02. PPVT 02 correlates .39 with Cattell 02, and .40 with PPVT 03. PPVT 03 correlates .41 with McCarthy IQ and .44 with the McCarthy Verbal subscale.

Birns and Golden (1972) have reported that ratings of Cattell test taking behaviors at 24 months related significantly to Binet scores at 36 months (N=89 black infants from varying SES strata), even when second year Cattell scores were partialled out. The correlations for rated "Pleasure in Task" and "Cooperation" were respectively .43 ($p<.005$) and .26 ($p<.05$). They argue that "...while there may be discontinuity between perceptual-motor development and later problem-solving ability on the verbal level...there may be continuity in terms of certain personality traits related to both preverbal and verbal intelligence" (p. 57). In their view this apparent "pleasure" precedes, rather than follows, successful problem-solving; it is a view also held by Levenstein (1970) and Yarrow (1976), and is linked by all authors to the intrinsic motivation for adaptive behavior characteristic of developing organisms.

There is an important exception to the general rule of higher within session correlations across measures than between session within measures. Cattell 01 correlates .53 with the McCarthy Motor subscale, but is not related as highly to PPVT 01 scores. This finding is certainly consistent with the view of Cattell 01 as primarily measuring perceptual-motor abilities, in contrast to the PPVT, Cattell 02, or other McCarthy subscales. These other measures appear to index aspects of emergent language abilities, whether they focus on verbal comprehension or verbal production. It is to be noted that Cattell wished to measure aspects of fine, rather than gross, perceptual-motor skills; McCarthy has emphasized measurement of gross, rather than fine, motor skills in describing the Motor subscale.

Certainly it is not surprising that the various subscales of the McCarthy are significantly intercorrelated with each other, and with McCarthy IQ, but it is important that the highest correlation (.92) is achieved by the Verbal subscale. These results are consistent with earlier interpretations of our data. Considered together, the findings suggest that the infant's verbal responsivity to examiner requests is the most important factor associated with relative success on the PPVT, the Cattell 02, and the McCarthy. Individual variability in this responsivity can be observed and measured among infants between ages 22 to 41 months, not individual differences in competence per se. We cannot, from these data, deduce that some children possess more linguistic competence than others as such; we can state that in this situation some children behaved as if they were more responsive and competent with an unfamiliar adult than others.

Samples from our 03 behavior profiles illustrate these points. Presented below are abbreviated profiles of four children: M, C, L, and S. The first two children performed better than the group average on the McCarthy, but while the first child also performed better than the group average on the PPVT, the second child did not. The last two children performed below average on the McCarthy, but while the third child performed above average on the PPVT, the fourth child did not.

M's Profile: McCarthy IQ=127 (High); PPVT IQ=96 (High)

Upon meeting the examiner, M was slightly shy, but warmed up rather quickly. Although he appeared to be a bit nervous initially, he became increasingly more comfortable as the testing session proceeded. He responded to questions asked of him without hesitation and was rather spontaneous in communicating with the examiner.

M was eager to begin playing with the toys and games, and entered the testing session with great anticipation. He appeared to be genuinely interested and very much involved in the tasks. From the beginning, he seemed to have no difficulty concentrating...and was attentive when instructions were given him. It appeared that he put his every effort into the tasks at hand, but did not become exceptionally discouraged when unable to perform a particular task. While soft-spoken, he showed special adeptness on the verbal tasks, but was fairly relaxed and quite confident on all the tasks throughout the testing session. He smiled frequently and responded favorably to the praise and support given...

C's Profile: McCarthy IQ=113 (High); PPVT IQ=68 (Low)

C entered the testing situation very friendly. He was eager to start the test.

C did well on the numerical memory and also on the motor coordination items....He got off the subject quite frequently. He was verbal, but only on the subjects he wanted to talk about, not (always) pertaining to the test...

L's Profile: McCarthy IQ=75 (Low); PPVT IQ=84 (High)

L was a child who was hard to "figure out." She came willingly to the test...and sat down and said she was ready to begin. She tired easily at many of the items, especially those above her abilities--more so than other children.

She always said she understood, but it was clear from her responses...that she did not know what was expected of her. Her abilities were scattered; she could name pictures well, but could not state concepts and she had perceptual and memory difficulties. Her motor coordination was good, but scattered. Her quantitative abilities were minimal.

L was pleasant and seemed willing, but was also fidgety and babbled incoherently periodically....

S's Profile: McCarthy IQ=68 (Low); PPVT IQ=60 (Low)

Upon meeting the examiner, S was friendly and, without hesitation, accompanied her to the testing area. She came willingly to the testing table and appeared to be eager to begin playing with the toys and games.

S was always so anxious to see what was to come next that she had difficulty concentrating on the tasks presently at hand...was quite distractible and responded actively to environmental stimuli not related to the testing session. Although she appeared to have been listening to directions, often her nonresponsiveness was indicative of her attention being elsewhere. This was particularly true on verbal tasks, while on tasks of motor skills and those requiring the manipulation of objects, her attention span increased somewhat. On many occasions, however, S did not wait for instructions, but attempted to do the tasks without them. When she could not or would not perform a task, she became silent and put her head down on the table.

Thoughout the testing session, S smiled frequently and verbalized spontaneously. She looked for reassurance and praise...and smiled when it was given...

While these four children are in no sense "typical" relative to the wide diversity of styles of interaction noted by our examiners, it is clear that different children respond differently to their testing experiences, both with test items and psychometrician.

## Discussion

We have indicated two factors which appear to be related to the testing performances of our children in the observable test situation: both the initial language skills displayed by them with an unfamiliar adult, and the social relationship, in general, between the child and the psychometrician as the test progressed.

Lewis and others (1976) have recently discussed how infant intelligence appears to be neither highly stable nor unitary. Lewis himself (1969) has found that a measure of infant attention (i.e., rate of decrement of visual attentiveness when consistently presented with a familiar stimulus, and rate of return when reintroduced to a novel stimulus) has higher predictability from age 12 months to subsequent Binet test performance at 3-1/2 years than do any of our existing infant tests. It appears that infant tests may be of some diagnostic value at any given age, but that both child clinical and, possibly, neurological assessments would be essential to increasing the validity of present predictions (Goeffney, 1971; Broman et al., 1975).

Existing infant measures of sensorimotor intelligence which emphasize a Piagetian developmental framework show similar patterns of instability and frequently do not relate to psychometric measures at any age. Lewis (1972), for example, found the Escalona-Corman scales to be significantly correlated with the Bayley Mental Index only at six months (.60). The scales did not, as did the 18 and 24 month Bayley Mental Scale, significantly correlate with an independent assessment at 24 months of language comprehension and production. Golden and Birns (1967) have reported similar results using the Cattell and the Escalona-Corman scale, obtaining a significant correlation of .24 at 12 months of age, but not at 18 or 24 months. Finally, Uzgiris (1976) reports that, of seven Piagetian-based scales, only one, The Development of Schemes scale (focused on a general assessment of the differentiation and coordination of schemes) correlated at 14, 18 and 24 months with the Bayley Mental Scale (cf. King and Seegmiller, 1973). Uzgiris concluded:

> Overall these studies suggest, first that any significant correlations are more likely to be obtained in the first year of life than in the second, possibly because all tests during the first year rely on the infant's fairly limited repertoire of actions for

> test items, even if the performances are given different interpretations. Second, object concept development, of central importance theoretically for sensorimotor development, hardly relates to performances measured by psychometric tests. Third, Bayley's Mental Scale correlates more highly than the Psychomotor Scale with the level in sensorimotor functioning during infancy. There are hardly any reports on the relationship between measures of sensorimotor functioning and psychometric tests over some interval of time. Wachs...results indicate that at each successive age level a greater number of the scales (i.e., Uzgiris-Hunt scales) are significantly related to the Stanford-Binet scores at 31 months. However, until additional studies are carried out to test the predictability of intellectual functioning in early childhood, no definite implications can be drawn for the relationship between the rate of coordination of sensorimotor functioning and intellectual functioning in later periods. (p. 148)

These results suggest that the organization of sensorimotor intelligence is complex and rapidly changing in the first two years of life. During the second, and presumably continuing into the third year, it is also particularly difficult to assess what with the emergence of many important new discrete functions (e.g., language) whose relationships to earlier ones are presently empirically unstable and multidimensional, rather than stable and unitary.

We do not know precisely <u>how</u> the social relationship with the examiner affects early infant performance. It seems reasonable to assume that such a situation is subject to many of the rules uiding all positive adult-child interactive contacts. Certainly, familiarity and appropriate adult responsivity are likely to increase the children's pleasure with the overall task and, considering the findings of Birns and Golden (1972), optimize their performances.

A similar, but important, perspective is offered by Haviland (1976), who believes that infant intelligence tests often demand clinical judgments of affect by the examiner. Related points in regard to the situational aspects of infant assessment were developed by Stott and Ball (1965) and Escalona (1950) in classical writings on the subject. However, Haviland expands these points by indicating that affect may be a good predictor of later intelligence <u>because</u> it must be used by psychometricians to render judgments of item success or failure in the first instance. She argues that we constantly use infants' affective cues to make judgments about their ongoing adaptability and implies that, realistically, we adults have little choice in the matter when we attempt to assess the intelligent behavior of an essentially preverbal child. However, these specific and immediate factors have to be considered from the vantage points of development, early intervention, and the sociocultural context in which both occur.

There are three major findings of this study which contradict previous research and thinking. First, at all assessment points, our children were superior to existing norms for black infants on the Cattell-McCarthy sequence.

Golden and Birns (1967) found a Cattell mean IQ of 98.5 for black children of welfare families at age 24 months, with the two higher classes of black children having IQs of 99.7. While the background of our sample, generally, is most similar to the Golden and Birns black welfare sample, we have found a mean Cattell IQ 111.6 at 22 months, and 105.2 at 32 months. Similarly, while Kaufman and DiCuio report a mean McCarthy IQ of 88.7 (N=87) for children of black unskilled workers using the McCarthy standardization sample, our average McCarthy IQ is 102.9 (N=57), roughly equivalent to that reported by these authors (1975b) for the black middle class standardization sample (IQ=101.0, N=65). Unfortunately, Kaufman and Kaufman do not present their data according to the ages of children, though we do know that at ages 2-1/2 to 3-1/2 blacks, as a group, average 97.4. Similarly, at ages 4 to 5-1/2 they average 94.6.

At both the second and third assessment points only our controls appear similar to the findings of other researchers using these two instruments. In short, our program children, at ages 32 and 41 months, are superior to existing norms for their social class/racial group. In addition, they are either equivalent or superior to recent norms offered by Kaufman and Kaufman on the McCarthy for two different white socioeconomic groups (Middle class=103.9, N=590; Working class=96.4, N=272).

We attribute this evident superiority to the efforts we made to evaluate the children in environmental surroundings which are familiar to them with examiners who, though personally unknown, were of the same racial background. Clearly also, the impact of programs on the human environments of the children was positive: program children are the most different from established norms for similar populations.

Secondly, it seems that, for those who are interested in such school-related criteria as these performance tests, our data do not support notions of deficit which require therapeutic interventions. We predicted IQ gains, as do most early intervention researchers in this field, but this prediction did not accurately reflect existing developmental data. Like Golden and Birns (1967) who observed a similar decline on the Cattell between ages 12 and 24 months, our psychometric data were in the direction of decrement rather than increment.

Since it is known that racially black infants are more advanced in earlier infant assessments than white infants (Freedman, 1974), the function of a developmentally-based intervention model would be to ensure that children continue to thrive and develop normally, such that class-related cultural-familial retardation (Heber, 1975) does not occur. At the conclusion of our study, this was certainly generally true of program children using our Cattell-McCarthy infant tests as evaluative criteria.

However, it is also clear that all infant tests of sensorimotor intelligence are markedly discontinuous from one age period to the next, both in items used, and in regard to theoretically sound linkages to the broader concept of sensorimotor intelligence (Lewis, 1976). The essentially negative findings with the PPVT, if these findings are contrasted with the Cattell-McCarthy findings, certainly support notions of diverging emergent intellectual abilities in infancy and in early childhood. Therefore, the message is clear. Any early intervention program must seek multiple confirming, or, better still, substitutive criteria of effectiveness. Structured efforts to enhance and support development simply cannot otherwise be evaluated fairly or thoroughly.

Thirdly, despite the absolute levels of performance at any age, roughly between ages two and three and one-half years, the children declined significantly in relative performance on the standardized psychometric measures of intelligence. This finding seemingly occurred on both test sequences. It suggests that children in black homes from these communities are being socialized differently during this period from their norm-setting peers in other communities. For example, in other analyses of observational data on the children's play styles, we found the Demonstration, Discussion, and Control children to use language differently during play; it was not as fruitful to characterize the three groups as having more or less of a particular attribute. Rather, the attributes combined differently. We need more such descriptive observational data on black and white children from diverse social status backgrounds from at least ages one through three and one-half years to confirm or disconfirm this hypothesis, data which developmentalists are only just beginning to accumulate. From our study, we do know that children whose mothers assimilate the posture and beliefs of contemporary child development specialists are, on average, least likely to decline as much in their performances.

At these ages, it is not the specific content of test items that is so unique for the child; rather, it seems that transactions between the child and others may generate different behavioral repertoires. In short, how the sociocultural environment influences which intellectual competencies are used and when, how they develop, and how they are integrated during this important formative period may well be one of the most critical issues to be faced in the future by those who would consider the generation of adaptive intelligence among socially and culturally different groups of children. In our view, these issues constitute the backdrop against which behaviors with psychometricians and other outsiders vary from what is typically predictable in the examining situation, and what, ultimately, will be designated as successful performance.

# References


Achenbach, T. Research in developmental psychology: concepts, strategies, methods. New York: Free Press, 1978.

Auerbach, A. Parents learn through discussion: principles and practices of parent group education. New York: Wiley and Sons, 1968.

Badger, E. Mother's guide to early learning. Paoli, Penn.: McGraw-Hill, 1973.

Bayley, N. Mental growth during the first three years. A developmental study of sixty-one children by repeated tests. Genetic Psychological Monographs, 1933, 14, 1-92.

Bayley, N. Development of mental abilities. In P. Mussen, (Ed.), Carmichael's Manual of Child Psychology (Vol. 1). New York: Wiley, 1970.

Birns, B. & Golden, M. Prediction of intellectual performance at three years on the basis of infant tests and personality measures. Merrill-Palmer Quarterly, 1972, 18, 53-58.

Bloom, L. One word at a time. The Hague: Mouton Press, 1973.

Broman, S., Nichols, P., & Kennedy, W. Preschool IQ-prenatal and early developmental correlates. Hillsdale, N.J.: L. Erlbaum Assoc., 1975.

Bronfenbrenner, U. Is early intervention effective: In U. Bronfenbrenner (Ed.), Influences on human development (2nd ed.). Hinsdale, Illinois: Dryden Press, 1975.

Cattell, P. The measurement of intelligence of infants and young children (Rev. ed.). New York: Johnson Reprint Corporation, 1960. (Originally published, 1940)

Dunn, L. Expanded manual for the Peabody picture vocabulary test. Minneapolis: American Service, 1965.

Escalona, S. The use of infant tests for predictive purposes. Bulletin Menninger Clinic, 1950, 14, 117-128.

Freedman, D. An ethological perspective on human infancy. New York: Erlbaum, 1974.

Goeffney, B., Henderson, N. & Butler, B. Negro-white, male-female eight-month developmental scores compared with seven-year WISC and Bender Test scores. Child Development, 1971, 42, 595-604.

Golden, M. & Birns, B. Social class and cognitive development in infancy. Merrill-Palmer Quarterly, 1967, 14, 139-149.

Golden, M., & Birns, B. Social class and infant intelligence. In M. Lewis (Ed.), Origins of intelligence. New York: Plenum Press, 1976.

Harms, I. & Spiker, C. Factors associated with the performance of young children on intelligence scales and tests of speech development. Journal of Genetic Psychology, 1959, 94, 3-22.

Haviland, J. Looking smart: the relationship between affect and intelligence in infancy. In M. Lewis (Ed.), Origins of intelligence. New York: Plenum Press, 1976.

Heber, R., & Garber, H. The Milwaukee project: a study of the use of family intervention to prevent cultural-familial mental retardation. In B. Friedlander, et al. (Eds.), Exceptional infant: assessment and intervention (Vol. 3). New York: Brunner-Mazel, 1975.

Kamin, L. The science and politics of IQ. New York: Wiley, 1975.

Kaufman, A., & Kaufman, N. Black-white differences at ages 2-1/2 to 8-1/2 on the McCarthy scales of children's abilities. Journal of School Psychology, 1973, 11, 196-206.

Kaufman, A., & Kaufman, N. Social class differences of the McCarthy scales for black and white children. Perceptual and Motor Skills, 1975a, 41, 205-206.

Kaufman, A., & DiCuio, R. Separate factor analyses of the McCarthy Scales for groups of black and white children. Journal of School Psychology, 1975b, 13, 10-18.

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

Levenstein, P. The mother-child home program. In M. Day and R. Parker (Eds.), The preschool in action. Boston: Allyn and Bacon, 1977.

Lewis, M. A developmental study of information processing within the first three years of life. Monographs, Society for Research in Child Development, 1969, 34, (Serial No. 133).

Lewis, M., & McGurk, H. The evaluation of infant intelligence: infant intelligence scores--true of false? Research Bulletin. Princeton, New Jersey: Educational Testing Service, 1972, 72-82.

Lewis M., & Wilson, C. Infant development in lower class American families. Human Development, 1972, 15, 112.

Lewis, M., & Freedle, R. The mother-infant dyad. In P. Pliner, et al. (Eds.), Communication and affect: language and thought. New York: Academic Press, 1973.

Lewis, M. (Ed.). Origins of intelligence. New York: Plenum Press, 1976.

McCall, R., et al. Transitions in infant sensorimotor development and the prediction of childhood IQ. American Psychologist, 1972, 27, 728.

McCarthy, D. Manual: McCarthy scales of children's abilities. New York: Psychological Corporation, 1972.

Nelson, K. Structure and strategy in learning to talk. Monographs, Society for Research in Child Development, 1973, 38 (1-2, Serial No. 149).

Slaughter, D. Modernization through education of mother-child dyads. Final Report I to National Institute of Child Health and Human Development. December, 1978.

Stodolsky, S. Defining treatment and outcome in early childhood education. In H. Walberg (Ed.), Rethinking urban education. Phi Delta Kappa, 1972.

Stott, L., & Ball, R. Infant and preschool mental tests: review and evaluation. Monographs, Society for Research in Child Development, 1965, 30 (3, Serial No. 101).

Tulkin, S., & Kagan, J. Mother-child interaction in the first year of life. Child Development, 1972, 43, 31.

Uzgiris, I. Organization of sensorimotor intelligence. In M. Lewis (Ed.), Origins of intelligence. New York: Plenum Press, 1976.

Valentine, C. Culture and poverty. Chicago: University of Chicago Press, 1968.

Wiener, G., Andrews, S. & Blumenthal, J. The New Orleans Parent Child Development Center. Paper presented at the Society for Research in Child Development, Denver, Colorado, April, 1975.

Yarrow, L., & Pedersen, F. The interplay between cognition and motivation in infancy. In M. Lewis, (Ed.), Origins of intelligence. New York: Plenum Press, 1976.

Footnotes

A version of this paper was first presented at the Fourth Conference on Empirical Research in Black Psychology, The Third College, University of California, San Diego, January 7-9, 1979. The author is on the faculty of the School of Education, Northwestern University, 2003 Sheridan Rd., Evanston, Ill. The research was supported by a grant from the National Institute of Child Health and Human Development (5 R01 HD08533).

[1]Most early childhood studies of this type have emphasized only child performance outcomes. Though not reported here, this study also attended to outcomes related to perceptions of childrearing and values, observational study of parent-child interaction, and parental development. We wished to determine whether early intervention in the form of parent education for childrearing would influence these outcomes; whether changes in children's behavioral outcomes would be paralleled by changes in mothers; and what would be the process by which and through which these changes occurred. At its onset, only one study had experimentally demonstrated changes in parents as children change (the PCDC Center in New Orleans, reported in a paper at the 1975 biennial meetings of the Society for Research in Child Development), and no study has lined such changes to the personal adult development of the parents. Our work was unique in these latter respects.

[2]The Discussion group model combined principles and procedures from the approaches to parent education of Aline Auerbach and Earladeen Badger. See those references and, especially the unpublished paper of Charlie Simpson, Coordinator of the groups,

as well as the December, 1978 report by Slaughter for more detail.

[3]The Verbal Scale is depicted in the Manual as measuring "...the child's ability to express himself verbally, and also... the maturity of his verbal concepts" (p. 3). The Quantitative Scale measures the child's aptitude for numbers while the Perceptual-Performance Scale measures his ability to reason (non-verbally) with materials. The Memory Scale emphasizes the child's short-term memory, while the Motor Scale emphasizes gross motor coordination. For each of the five subscales as well as the GCI, the child's raw score is converted to a scaled score or Index in accordance with his chronological age. The GCI has a mean set at 100 with a standard deviation of 16, while each of the sub-scales have means set at 50 with a standard deviation of 10.

[4]Since (a) the first Cattells did not differ, and (b) early PPVT differences were not sustained later on, analyses between PPVT scores and other social background variables to identify possible sources of the average initial PPVT differences were not conducted.

[5]In all but four reported instances, the natural mother of the child participated in the programs. In these special instances, the "mothers" were actually grandmothers.

COMPARATIVE SOCIAL BACKGROUND DATA ON MOTHERS AND CHILDREN BY PROGRAM GROUP (PERCENTAGES)[a]

| Variable | Program Group: Toy Demonstration (n = 26-25) | Mothers Discussion (n = 26-22) | Controls (n = 29-31) |
|---|---|---|---|
| 1. No Current Work Experience | 92.0 | 100.0 | 96.8 |
| 2. No Past Work Experience prior to Child's Birth | 20.0 | 40.0 | 43.3 |
| 3. Previous Full-Time Work Experience | 36.0 | 37.5 | 24.1 |
| 4. Primary Family Income=AFDC | 88.5 | 96.0 | 93.3 |
| 5. No Secondary Income Reported | 80.8 | 96.0 | 83.3 |
| 6. Maternal Health Perceived Below Average | 24.0 | 18.2 | 20.0 |
| 7. Never Apart from Child More Than One Week | 65.4 | 69.6 | 73.3 |
| 8. Contact with Child's Natural Father is Maintained | 80.8 | 58.3 | 32.3 |
| 9. No Adult Males in Household (over age 17) | 26.9 | 48.0 | 38.7 |
| 10. Only One Adult Female in Household (over age 17) | 50.0 | 64.0 | 35.5 |

[a]The average ages of mothers at the time of maternal interviews (7/1/75) were, repectively 24.8±8 (Toy Demonstration), 26.9±9 (Discussion group), and 23.7±6 (Controls).

The average years of maternal education were, respectively 11.0±2 (Toy Demonstration), 10.6±1 (Discussion group), and 10.5±2 (Controls).

Table 2. Comparison of Program Stays on Drops on 01 Measures of Child Intellectual Functioning[a]

| TEST | Program Status | N | Mean | Standard Deviation | Sum of Squares | Df | F | t |
|---|---|---|---|---|---|---|---|---|
| Cattell Mental Age (01 test) | Stayed in | 83 | 24.89 | 3.47 | B: 12.633 | 1,127 | 1.1741 | 1.0836 |
| | Dropped out | 46 | 24.24 | 2.91 | W: 1,366.430 | | | |
| Cattell I.Q. | Stayed in | 83 | 110.98 | 12.23 | B: 30.875 | 1,127 | 0.2352 | 0.4850 |
| | Dropped out | 46 | 109.96 | 9.90 | W:16,668.125 | | | |
| PPVT Mental Age (01 test) | Stayed in | 72 | 22.49 | 3.41 | B: 7.113 | 1,107 | 0.6772 | 0.8229 |
| | Dropped out | 37 | 21.95 | 2.88 | W: 1,123.906 | | | |
| PPVT I.Q. | Stayed in | 74 | 99.08 | 12.10 | B: 1.000 | 1,109 | 0.0065 | 0.0809 |
| | Dropped out | 37 | 99.22 | 12.90 | W:16,672.000 | | | |

[a]One way ANOVA, fixed effect model

Table 3. Descriptive Child Test Data by Program Group Using All Tested Ss (ATSs) and Ss with Complete Data (CDSs)

| Test (Time) | | Toy Demonstration | | Mothers Discussion | | Control | |
|---|---|---|---|---|---|---|---|
| | | ATSs | CDSs | ATSs | CDSs | ATSs | CDSs |
| Cattell (01) | $\bar{x}$ | 113.0 | 114.3 | 110.1 | 110.9 | 110.1 | 109.7 |
| | s.d. | 14.1 | 14.0 | 9.2 | 9.6 | 12.9 | 11. |
| | n | (26) | (21) | (26) | (15) | (31) | (21) |
| PPVT (01) | $\bar{x}$ | 101.8 | 102.0 | 101.7 | 101.9 | 94.4 | 94.4 |
| | s.d. | 12.7 | 12.6 | 9.8 | 9.7 | 12.3 | 11.8 |
| | n | (27) | (20) | (23) | (15) | (27) | (20) |
| Cattell (02) | $\bar{x}$ | 108.1 | 109.6 | 104.6 | 104.0 | 103.2 | 102.0 |
| | s.d. | | 10.2 | 10.3 | 10.6 | 10.5 | 11.0 |
| | n | (26) | (21) | (22) | (15) | (28) | (21) |
| PPVT (02) | $\bar{x}$ | 90.6 | 91.4 | 89.2 | 91.7 | 83.9 | 84.0 |
| | s.d. | 10.1 | 11.3 | 12.7 | 13.0 | 9.4 | 8.4 |
| | n | (26) | (20) | (21) | (15) | (30) | (20) |
| McCarthy (03) | $\bar{x}$ | 104.1 | 104.1 | 100.8 | 105.6 | 95.6 | 96.5 |
| | s.d. | 12.8 | 12.8 | 18.2 | 14.1 | 12.1 | 12.1 |
| | n | (21) | (21) | (19) | (15) | (24) | (21) |
| PPVT (03) | $\bar{x}$ | 77.5 | 76.5 | 75.5 | 77.2 | 73.7 | 74.6 |
| | s.d. | 10.5 | 10.1 | 8.0 | 7.8 | 9.6 | 8.8 |
| | n | (22) | (20) | (19) | (15) | (24) | (20) |

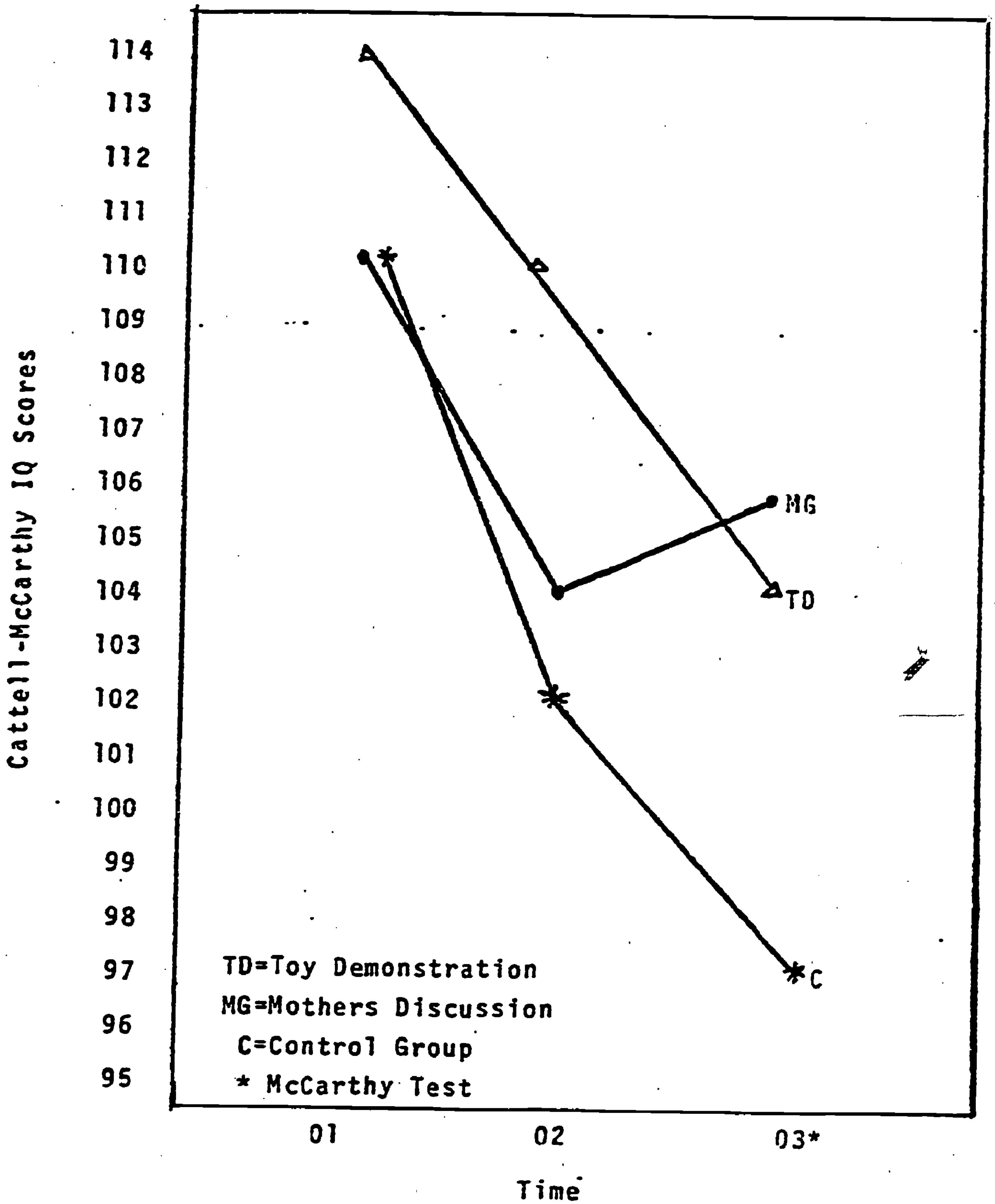
## Table 4. Stability Correlations of Child IQ Test Measures over Times 01-03

| | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|
| (1) Cattell 01 | .50** (74) | .37** (76) | .30** (77) | .27* (64) | .26* (64) | .14 (64) | .23 (64) | .36** (64) | .53** (64) | -.05 (65) |
| (2) PPVT 01 | | .21 (69) | .30* (71) | .10 (57) | .15 (57) | -.18 (57) | .06 (57) | .20 (57) | .24 (57) | -.05 (58) |
| (3) Cattell 02 | | | .39** (73) | .45** (57) | .38** (57) | .34** (57) | .30* (57) | .34** (57) | .08 (57) | .18 (58) |
| (4) PPVT 02 | | | | .29* (59) | .34** (59) | .12 (59) | .15 (59) | .30* (59) | .22 (59) | .40** (60) |
| (5) McCarthy 03 | | | | | .92** (64) | .69** (64) | .74** (64) | .78** (64) | .45** (64) | .41** (63) |
| (6) McCarthy Verbal 03[a] | | | | | | .48** (64) | .55** (64) | .84** (64) | .40** (64) | .44** (63) |
| (7) McCarthy Perceptual-Performance | | | | | | | .50** (64) | .35** (64) | .27* (64) | .38** (63) |
| (8) McCarthy Quantitative | | | | | | | | .51** (64) | .23* (64) | .22* (63) |
| (9) McCarthy Memory | | | | | | | | | .51** (64) | .28** (63) |
| (10) McCarthy Motor | | | | | | | | | | -.00 (63) |
| (11) PPVT 03 | | | | | | | | | | |

[a]Items (6)-(10) are Subtests of the McCarthy Scales of Mental Abilities.
**$p \leq .01$
*$p \leq .05$

# Figure Legend

Figure 1. Profile of Average Infant Test Scores (Cattell-McCarthy) by Program Group Across Time.


Cattell-McCarthy IQ Scores
114
113
112
111
110
109
108
107
106
105
104
103
102
101
100
99
98
97
96
95
MG
TD
C
TD=Toy Demonstration
MG=Mothers Discussion
C=Control Group
* McCarthy Test
01
02
03*
Time