An Analysis of the Performances of Six-Month-Old Infants on the Cattell Infant Intelligence Scale

Patricia Bledsoe
Loyola University Chicago

Follow this and additional works at: https://ecommons.luc.edu/luc_theses

Recommended Citation
Bledsoe, Patricia, "An Analysis of the Performances of Six-Month-Old Infants on the Cattell Infant Intelligence Scale" (1954). Master's Theses. 914.
https://ecommons.luc.edu/luc_theses/914

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License.
Copyright © 1954 Patricia Bledsoe
AN ANALYSIS OF THE PERFORMANCES OF SIX-MONTH-OLD INFANTS ON THE CATTELL INFANT INTELLIGENCE SCALE

by

Patricia Miesse

A Thesis Submitted to the Faculty of the Graduate School of Loyola University in Partial Fulfillment of the Requirements for the Degree of Master of Arts
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>General position of infant testing—Indications from studies of the Cattell Scale—Primary purpose of present research and related investigation—Secondary purpose.</td>
<td></td>
</tr>
<tr>
<td>II. REVIEW OF THE LITERATURE</td>
<td>7</td>
</tr>
<tr>
<td>Factors influencing infant test construction—History of infant tests—Summary of research on the Cattell Scale—Integrated summary of the literature.</td>
<td></td>
</tr>
<tr>
<td>III. DESIGN OF THE RESEARCH</td>
<td>35</td>
</tr>
<tr>
<td>Discussion of the Cattell Infant Intelligence Scale—Subjects of Present study—Test items and materials—Statistical procedures used in present comparative study.</td>
<td></td>
</tr>
<tr>
<td>IV. ANALYSIS OF RESULTS</td>
<td>44</td>
</tr>
<tr>
<td>Comparisons of performances of infant groups on total scale—Comparative analyses of performances of infant groups on individual items.</td>
<td></td>
</tr>
<tr>
<td>V. SUMMARY AND CONCLUSIONS</td>
<td>53</td>
</tr>
<tr>
<td>APPENDIX I</td>
<td>56</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>57</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table

I  MEANS, STANDARD DEVIATIONS AND RELIABILITY FOR THE FOUR GROUPS OF ADOPTIVE INFANTS ........................................... 44

II  COMPARISON OF PERCENT OF ADOPTIVE AND CATTELL SIX-MONTH-OLD INFANTS PASSING SIX AND SEVEN MONTH CATTELL ITEMS ........ 47

III COMPARISON OF PERCENT OF SIX AND SEVEN-MONTH-OLD ADOPTIVE INFANTS PASSING SIX AND SEVEN MONTH CATTELL ITEMS ........ 49
CHAPTER I

INTRODUCTION

Investigation in the area of infant intelligence, particularly that conducted with a view to devising and perfecting instruments of measurement, is a challenging proposition. The activities usually considered to be manifestations of general intelligence - the abilities to retain and to recall, to abstract and to relate - can be presumed to exist in some potential form in the very young infant, but they are as yet unrealized due to immaturity of physical development and lack of experience. As a result, the sub-tests of infant intelligence scales have been geared largely to an appraisal of observable behavior which, during the period of early infancy, is probably largely sensorimotor in character and, as available studies indicate, of questionable relation to intelligent behavior at later ages.

The behavioral adjustments upon which the psychologist depends when testing the school age child and the adult are absent as such in the infant; the child under one year is unable to sustain his attention for directed activities, to engage in oral communication or to manipulate the test materials with any precise degree of motor coordination. Therefore, he can only be evaluated on the basis of behavior elicited naturally upon presentation of the appropriate stimuli.
Finally, in the actual testing situation, external factors quite unrelated to actual mental ability are nevertheless extremely important in evaluating the test results. For example, the quality of rapport and cooperation obtained from the very young child is difficult to determine. It is often impossible to decide whether the infant is unable or merely unwilling for some reason to produce the expected behavior. The infant responds immediately and integrally to discomforting stimuli such as hunger, fatigue, or strangeness of environment; he can be fretful, moody, withdrawn or generally unresponsive without apparent cause and his reactions, because of his lack of understanding, are necessarily outside of the motivating influences of competition, and the wish to please and to perform at his best.

If one is to judge from the amount of published research, the field of infant intelligence testing has been a neglected one. Of the numerous tests devised for the measurement of intelligence in the school age child and the adult, at least two of these, the Revised Stanford-Binet Scales and the Wechsler-Bellevue Intelligence Scale, have been greatly augmented for clinical usage by the number of formal studies on various phases of their application. The clinician can now use these tests with reasonable confidence in their value and awareness of their limitations. The situation has been quite the reverse in the field of infant testing. Several separate infant scales have been published in the past few decades, but there has been little follow-up work on their validity and reliability. To all appearances, the detection of defects in the existing scales has been the cue for the construction of a new
scale rather than for systematic investigation and refinement of the tests on
hand. Since the authors have borrowed liberally from one another, this could
mean that the weaknesses in the infant tests have been transplanted from
scale to scale, instead of being corrected by a wider application and apprais-
al of one of the more promising tests.

In spite of their limitations, we know from indirect references in
psychological and social work literature that certain of the infant tests, in
particular the Cattell Infant Intelligence Scale, are in common use in the
clinic. Articles by Carter and Bowles, Escalona, Fischer, Klatskin, and
Gallagher - to mention the only studies reporting on the Cattell Scale -
refer to a regular clinical use of the test, usually for the purpose of deter-
mining the suitability of young infants for adoption. Another likely area in

1 Psyche Cattell, The Measurement of Intelligence of Infants and
Young Children, New York, 1940.

of Psychological Examining," Journal of Clinical Psychology, IV, 1948, 109-
150.

3 Sibylle Escalona, "The Use of Infant Tests for Predictive

4 Liselotte K. Fischer, "Hospitalism in Six-Month-Old Infants,"

5 Ethelyn H. Klatskin, "Intelligence Test Performance at One Year
Among Infants Raised with Flexible Methodology," Journal of Clinical

6 James J. Gallagher, "Clinical Judgment and the Cattell Infant
which the scale is being used is the diagnosis of mental deficiency, since the need for parental guidance and plans for training the retarded child make detection desirable at as early an age as possible.

The Cattell Scale, covering the age range of two to thirty months, was so constructed as to form a downward extension of the Revised Stanford-Binet, Form L. It is convenient to administer and its findings are expressed in the common M.A. and I.Q. units. For these reasons it lends itself well to use in the clinical setting. However, there is some evidence in the above cited studies that one of the most crucial demands of intelligence test construction - that of suitable item placement - was not adequately met in the standardization of the scale. According to these investigations, when the scale has been used with other fairly comparable infant populations, results which are significantly higher than those reported by Cattell have been found. Carter and Bowles found this to be true of the mean total scores obtained with two groups of infants, two and three months of age. Klatskin found significantly higher scores on the total scale and a significantly greater number of successes on most individual items when she tested a large group of twelve and thirteen month old infants.

It is true that several factors could cooperate to produce these discrepancies. For example, standards for evaluating success and failure on the individual sub-tests may differ among examiners using the scale. Again, the infant populations may not be entirely alike. But one likely explanation which can not be overlooked is that some of the test items, arranged by
Cattell on the basis of percent passing, are misplaced and are too easy for the age groups for whom they are intended. In order to test the validity of this assumption, many more reports on wider applications of the scale by different examiners are needed. Thus, it has been considered to be of some value to extend the type of comparative item analysis done by Klatskin and investigate the performances of another sample of infants on other age levels of the test.

The present study will concern itself with the performance of two groups of adoptive infants, six and seven months of age, on the six and seven month age levels of the Cattell Infant Intelligence Scale. The primary purpose will be to determine whether any significant differences exist between the performances of the six-month-old adoptive infants, both with regard to total scores and successes on the individual items placed at those age levels, and the infants of the same age comprising Cattell's standardization group. As a related investigation of the suitability of placement of these items, an inter-group comparison will be made of the performances of the six and seven month adoptive infants on the six month items and on the seven month items. As Cattell did not test her infants at seven months, a more extensive comparison of the performances of the seven month infants is not possible.

As a secondary purpose, the findings of two investigations on sex differences in infant test performance will be explored further. Klatskin found no significant differences beyond the chance level in the performances on the Cattell Scale of the boys and girls in her study group. Nelson and
Richards, in an analysis of the six month items of the Gesell Developmental Schedules, also failed to find any significant differences in performance produced by the sex of the infant. Accordingly, the test performances of the six-month-old boys and the six-month-old girls of the present study group will be analyzed separately and compared, in order to determine whether the sex difference produced any significant differences in performance on the total scale or on individual items.

CHAPTER II

REVIEW OF THE LITERATURE

The period of infancy is variously considered as extending up to eighteen months or to twenty-four months. Beyond that age interval is regarded as the preschool period. There has, of course, been overlapping of these age levels by intelligence scales. The Cattell Scale extends well up into the preschool period. Two of the well known preschool tests, the Merrill-Palmer Scale\(^1\) and the Minnesota Pre-school Scale,\(^2\) presents items for as low as eighteen months. But since the present research is concerned with age levels within the first year of life, only those studies and tests which have specific reference to that period will be considered in the following resume of the literature.

There are several points of difference between the development and behavior of the young infant and the preschool child which justify a separate consideration of their testing problems. The young infant must be tested either while lying in a crib or placed in a well supported sitting position. The normal preschool child can walk and move independently, and hence can


assume a position at a table with an examiner. The infant does not engage in verbal communication. The preschool child uses speech as a tool and can participate in activities demanding verbal response. The infant reacts primarily to the testing equipment which must be especially designed to elicit the desired behavior. Only secondarily does he usually relate to the examiner, who thus acts in the testing situation as an observer. In contrast, the preschool child is capable of entering into a direct interpersonal relationship with the examiner and the quality of rapport assumes a much greater importance.

But most particularly for the purposes of test construction, the young infant's developmental rate is much more rapid and the nature of the test items themselves is quite different. As will be pointed out in more detail, the test items in the first six months are largely sensorimotor in character, but this component has been found to decrease gradually in importance after that age level. In contrast, the test activities designed for the preschool child are more clearly of a problem solving nature.

Several writers have indicated characteristics of the infancy period which bear upon the general problem of evaluating infant mental development. One of the more important of these is the transient nature of infant behavior. Gesell, who is probably the leading American investigator in the area of infant development, describes the period of infancy as one of changing and fugitive behavior, exceeding all other age intervals in the wealth of phenomena displayed. Referring to the difficulties involved in selecting suitable norms for infant behavior, he states that "even with simplification we must
recognize in the first year of life at least three developmental intervals and devote attention to the stages of maturity presented at four months, six months, and nine months.Important developmental changes occur which mark off these short age intervals one from another, and which have significance as levels at which to evaluate developmental or growth rate. However, Gesell has concluded as a result of his extensive observations, that infant growth, although rapid, is manifested in patterns of behavior which are governed by deep-seated or ontogenetic laws of developmental sequence. Therefore, graded tests of behavior can be utilized to determine the rate of maturity of the growing system.

Thompson, a co-worker of Gesell in the Yale Clinic of Child Development, concluded as a result of daily observations of infant behavior under well controlled conditions that behavior growth proceeds fully as rapidly as physical growth. These growth increments may occur in different functions on successive days or they may occur in more than one function on the same day. Further, the behavior growth increment may manifest itself in one of the four following ways: (1) The greater frequency of one item of behavior; (2) the improved performance of an activity; (3) the appearance of a new

---


activity, and (4) the integration of previous activities. A summary presented by Gesell of the progress of prehensory development during the first year of life - in which growth proceeds from the tight reflex grasp in the neonate to the precise finger-thumb opposition at twelve months - provides a clear illustration of the span of development in one type of behavior.6

The fluctuations occurring in infant growth have some important implications for infant intelligence test construction. For example, in her longitudinal study regarded as one of the outstanding contributions to infant testing, Bayley reported some inconsistencies in reliability coefficients obtained for the infant levels of her scale, notably at one to three months, when the same children were tested at different ages.7 In reviewing this situation, Cronbach concluded that her neonate tests were unreliable because at a level where a new activity is just emerging "the pattern is diffuse, varied, and inconsistent from time to time; measurement of such functions is therefore unstable."8 He further stated that a scale showing a satisfactory over-all reliability may be unreliable at certain levels or for certain groups. Cattell likewise found some wide irregularities in the individual

6 Arnold Gesell and Catherine Armatruda, Developmental Diagnosis, New York, 1941, 191.

7 Nancy Bayley, "Mental Growth during the First Three Years of Life: A Developmental Study of Sixty One Children By Repeated Tests," Genetic Psychology Monographs, XIV, 1933.

curves for some of her children who were tested at several different ages. In explanation, she proposed that some of the variations noted in the test-retest scores resulted from changes in the tempo of development rather than from inadequacies in the tests themselves.²

Anderson has indicated that the changing nature of infant growth has an important effect on the customary criterion used in standardizing infant scales. He stated that of the four traditional criteria used in validating intelligence tests, only that of increase in score with chronological age has been employed in constructing infant scales. But "since (infant) development is a timed series of reactions or sequences, there are for many functions periods below which only a small portion of the function can be measured and above which a progressively larger portion can be measured. Hence the possibilities of prediction are limited and progression with age is not an infallible indicator of the value of a measurement."³

Because of the use of the criterion of progression with age, infant scales have always included many motor items. Motor behavior is readily observable, and as Gesell's studies have revealed, there is an orderly progression from gross to fine activity in relation to age maturation. However, motor behavior has not been found to correlate well with intelligence at later ages. Bayley found some community of function between mental and motor

---

² Cattell, Infant Intelligence, 60.

scores during the first fifteen months—correlations were in the vicinity of .5—but the relationship dropped markedly after that age level. Bayley sought to overcome the influences of motor items by arranging tests of this nature into a separate motor scale. Cattell also sought to eliminate items which appeared to be related chiefly to motor ability. Gesell separated infant behavior into several categories, among them motor development, so that separate evaluations could be made. Similarly Buhler separated her test items and designated certain sub-tests as involving "bodily control." However, the difficulty in making a clear-cut distinction between motor and other types of behavior during the first year of life, and hence of excluding entirely the influence of motor development, has been pointed out by Gesell in the following comment:

Motor and adaptive behavior are intimately combined in early life, because under the pressure of growth, a normal infant feels impelled to put each newly attained motor ability to repetitive use, and to exercise it with experimental variations. For example, an eight week old infant can not reach for a rattle but will briefly retain a rattle placed in his hand—a slight bit of adaptive behavior which is not altogether pure reflex. At twelve weeks he will hold the rattle actively and even glance in its direction. At sixteen weeks he regards it immediately and intently. He also deploys his eyes in a roving manner to "contact" his surroundings. In the next two months

11 Bayley, "Mental Growth during the First Three Years," Genetic Psychology Monographs, XIV, 1933.


he reaches out to contact, to grasp and to hold. Thus by subtle
growth stages which begin very early the infant's visual and manual
behavior takes on voluntary and adaptive characteristics.\textsuperscript{14}

An inspection of the existing infant scales reveals a similarity
in the kinds of test items included, especially for the age levels under
doll's visual and manual
twelve months. In the first place, the items are limited by the small range
behavior that can be elicited from the infant. Further, most of the
of behavior that can be elicited from the infant. Further, most of the
recent test authors have borrowed heavily from Gesell's normative items,
although they have frequently placed them at other than the originally des-
ignated age levels in accord with their own findings. The following items
although they have frequently placed them at other than the originally des-
ignated age levels in accord with their own findings. The following items
from Bayley's scale,\textsuperscript{15} together with the exact age placement for each item,
are presented here as fairly typical of the activities expected in testing
the six-month-old infant:

<table>
<thead>
<tr>
<th>Activity</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reaches persistently</td>
<td>6.05</td>
</tr>
<tr>
<td>Turns after spoon</td>
<td>6.1</td>
</tr>
<tr>
<td>Mirror image, approach</td>
<td>6.1</td>
</tr>
<tr>
<td>Picks cube deftly</td>
<td>6.1</td>
</tr>
<tr>
<td>Several syllables</td>
<td>6.3</td>
</tr>
<tr>
<td>Bangs in play</td>
<td>6.35</td>
</tr>
<tr>
<td>Sustained inspection of a ring</td>
<td>6.4</td>
</tr>
<tr>
<td>Unilateral reach</td>
<td>6.45</td>
</tr>
<tr>
<td>Vocalizes satisfaction</td>
<td>6.5</td>
</tr>
<tr>
<td>Lifts cup by the handle</td>
<td>6.6</td>
</tr>
<tr>
<td>Exploits string play</td>
<td>6.7</td>
</tr>
<tr>
<td>Rotates wrist</td>
<td>6.7</td>
</tr>
<tr>
<td>Scoops pellet</td>
<td>6.8</td>
</tr>
</tbody>
</table>

\textsuperscript{14} Gesell, \textit{Infant Development}, 58.

\textsuperscript{15} Nancy Bayley, \textit{The California First Year Mental Scale},
Berkeley, 1933.
Most of the test authors have avoided any logical explanation of the nature of the behavior underlying their test items. An exception to this is Gesell, who classified items into one of the four following categories: (1) motor development; (2) language development; (3) adaptive behavior, and (4) personal-social behavior. Duhler also attempted some classification, by labeling the individual items in accord with the behavior they were intended to evaluate. Her system included four general lines of development - bodily control, mental ability, manipulation of objects and social development.

Bayley attempted several classifications of her test items but found them to be unsatisfactory. "In many cases an adequate response to a test situation requires abilities of more than one kind, so that items may equally well be assigned to two or more classes." As has already been pointed out, Gesell found this to be true but he attempted to select the type of behavior chiefly involved in order to provide a means of evaluating progress in certain general areas of growth. An added difficulty noted by Bayley in classifying tests into sub-groups was that no two areas of infant development show parallel development. She made a broad classification of test items into two categories - sensorimotor and adaptive behaviors (the first group involving sensory acuity and fine motor adjustments, the second

16 Gesell and Armatruda, Developmental Diagnosis, 5.
17 Bayley, "Mental Growth during the First Three Years," Genetic Psychology Monographs, XIV, 61.
demanding learning and problem solving) - and studied their influences on test performances during the first year. She discovered that development during the first six or eight months was largely sensorimotor in character, whereas the more truly adaptive behavior is measured by the tests only after that period.  

Bayley's findings are probably applicable to all of the infant intelligence scales. Watson, for example, has indicated in a general way the chief abilities demanded by the Cattell sub-tests during the first twelve months. According to his designation, the tasks are largely perceptive in nature at the earliest levels - for example, attending to sounds or visually following a moving ring. Beginning at about five months, there is a gradual change to more manipulatory (adaptive) tasks. The first verbal type of test appears at nine months, involving adjustment to words - that is, performing an activity in response to a spoken request. The first speaking vocabulary item appears at eleven months. From that point on, more verbal tests are utilized although manipulatory tests still predominate.  

The few available reports on the predictive value of infant intelligence tests have been discouraging. Cattell found the validity coefficients between the six and nine month levels of her scale and the Revised

18 Ibid., 63.

Stanford-Binet, Form I at thirty-six months to be only .34 and .18 respectively. In her study, Bayley found her correlations between her early tests and those at eighteen months approached zero and with later tests they reached a negative correlation as high as -.21. Furfey and Muhlenbein,20 in an investigation of the Linfert-Hierholzer Scale,21 found that scores made by children when tested between the ages of six and twelve months with that scale showed no relation to scores on the Stanford-Binet four years later.

As one of the main conclusions of her longitudinal study, Bayley sounded a pessimistic note on one of the most crucial of the issues involved in infant testing, when she questioned whether there is a clear-cut, unitary relationship between intelligence in the infant and intelligence in the adult. On the basis of her findings, she concluded that "the behavior growth of the early months of infant development has little predictive relation to the later development of intelligence - even though the later behavior may depend in large part on the previously matured, elementary connections or behavior patterns."22 In further explanation, Bayley suggested that there occurs a shifting of functions as the child matures and the most that the existing


22 Bayley, "Mental Growth during the First Three Years of Life," Genetic Psychology Monographs, XIV, 74.
tests can do is to measure these functions, or groups of functions, at successive age levels and offer an appraisal of development only for the particular age level tested, rather than to measure a unit function of intelligence that extends from birth throughout life. If proved by further longitudinal study to be true, this is a grave restriction, since infant tests are used largely to predict intelligence at later ages.

In view of the weaknesses of infant tests, their clinical usage demands a much more limited and qualified application than is customary with tests for older age groups. After their survey of the current status of infant testing, Carter and Bowles\(^\text{23}\) warned that the task of interpretation is essentially that of a rough screening job in which classification into one of three groups is attempted; those infants whose developmental progress is subnormal or in any way doubtful, those who appear to be making average or satisfactory progress, and those who are found to have made unusual or accelerated progress. However, even this rough application makes the infant intelligence scale a valuable clinical tool, since it can aid in discerning wide deviations from developmental norms, and this is a substantial contribution to adoption practices, to cite one example. In a recent discussion of her infant scale, Cattell makes the following comment on this issue: "The younger the child, especially below the age of fifteen months, the less the

predictive value of the tests, but even for a child as young as three months they have been found to be of value in indicating extreme variations from the norm in the direction of both feeble-mindedness and superiority."24

**Historical Development of Infant Intelligence Tests**

A survey of the history of infant intelligence test development reveals an early interest in the field. Over the years a large number of test items for evaluating different aspects of infant behavior have been devised and presented either as supplements at the lower end of tests for more advanced age levels or as groups constituting separate infant scales.

As early as 1904, when he published his scale for measuring intelligence in the school child, Binet25 presented some four or five items which were suitable for evaluating reactions of infants under one year. However, he intended these items to be used in differentiating mental defectives who were too retarded to perform on the regular scale; in consequence, Binet did not assign any exact age placement to these tests in terms of their applicability to infant behavior, but it is interesting to note that all of these items - namely, reaction to light and sound,prehension after tactile excitation and after visual perception, and imitation of movements and execution


of simple orders in response to word or gesture - appear in present day infant scales in some form.

In his 1922 revision of the Binet Scale, Kuhlmann\(^{26}\) extended the test at the lower level from three years down to three months, presenting five items for scoring at each of the age levels three, six, nine, twelve, eighteen, and twenty-four months. However, he standardized the items on very small groups - the three month tests were given to twenty children, the six month tests to forty-six children - and the scoring of the tests was indefinite since many of them were dependent upon the report of the child's parents.

Over the years since 1919, when he began his investigations in the Yale Clinic of Child Development, Gesell has developed many of the standard items used in the current infant tests. The authors of the recent infant scales - Bayley, Cattell and Gilliland - have all indicated that they drew heavily on Gesell items. In 1925, Gesell published his first schedule of developmental norms,\(^{27}\) extending from birth up to five years. Over five hundred children were examined at four, six, twelve, eighteen, twenty-four, thirty-six, forty-eight and sixty months of age and separate schedules, involving a total of one hundred and fifty normative items, were then arranged for each level of development. The chief objection raised against this first


\(^{27}\) Gesell, *The Mental Growth of the Preschool Child*. 
scale was that no precise ratings of an individual child's development could be established. Gesell described different grades of success for each item, so that a general idea of a child's developmental level could be obtained by examining the different degrees of success attained on the total scale but he did not indicate any exact at-age values for his items.

In 1947, Gesell published the most recent description of his revised and supplemented version - the Gesell Developmental Schedules - which presented over two hundred items representing behavior characteristics for the age levels between four weeks and forty-two months. As in the early scale, items are arranged to check behavior in the four separate areas of development - language, motor, personal-social, and adaptive. For example, at the six months (twenty-four weeks) level there are six motor items, six adaptive items, three language items and four personal social items. The norms in each area were derived from observations of infants and young children and placed with objective reference to the age at which the elicited behavior patterns are normally expected to appear. It is important to remember that Gesell's schedules comprise a normative scale, rather than an intelligence scale in the strict sense. His method of tapping behavior at its expected level of appearance differs from the empirical method used, for example, by Cattell and Gilliland, of assigning item placement on the basis of percentage of successes by the given age group. The Gesell scale therefore does not lend itself to an I.Q. rating, although the infant's total

28 Gesell and Armatruda, Developmental Diagnosis.
score on the four areas of behavior can be divided by the chronological age to give a developmental quotient, which indicates the proposition of normal development present at the time of the examination. There has been one study on the six month items of the Gesell scales which will be reported later in this chapter.

Another normative scale was published in 1928 by Hetzer and Wolf.29 On the basis of twenty-four hour observations of infants in the laboratories of the Psychological Institute of Vienna, these authors presented monthly norms through eleven months. In 1930, Buhler30 published another version of these tests, in which the scale was extended up to two years and revised along the general lines of the Binet scale. A series of ten items was presented for each month level up to twelve months, items being selected to evaluate the four general lines of development already described. The tests were drawn up after ten preliminary trials were made for each month and then given to thirty children at each month level. The final score is expressed in terms of a developmental age, obtained by adding credits to a basal score. Certain of the Buhler tests were used by Cattell and Gilliland in the construction of their scales. However, the Buhler scale has been criticized for clinical usage because it was standardized on institutional babies and because it involves many situations which are frustrating or frightening to

30 Buhler, The First Year of Life.
the child. Further, using Buhler's own classification, Cattell found less than half of the items to relate mainly to mental development.31

In 1928, Linfert and Hierholzer,32 graduate students at the Catholic University of America, published their point scale for the first twelve months, based mainly on the Gesell tests. It was claimed to be the first standardized scale with age norms published for that period of life. The scale was divided into two series, and included tests for one, two, four, six, nine, and twelve months. Tables indicating percentage of successes in the various tests were presented for calculating age norms, the final results to be expressed in an L-H quotient. According to the test authors, the total point scores showed a linear increase with age. In the previously mentioned study by Furfey and Muhlenbein, no significant relationship was found between the results of this scale administered during the first year of life and the 1916 Stanford-Binet administered four years later to the same children. No other studies have reported on this scale. The extent of its clinical usage after publication is not known.

As an outgrowth of her longitudinal study, which involved 1142 tests on sixty one children over a three year period, Bayley published the California First Year Mental Scale,33 covering the age range of one month to

31 Cattell, Infant Intelligence, 22.
33 Bayley, The California First Year Mental Scale, 1933.
eighteen months. Using a large number of the Gesell items, Bayley included tests of adaptibility or learning, and tests of sensory acuity and fine motor coordination. She placed her items on a continuous scale in order of difficulty by the Thurstone Method of Absolute Scaling and indicated exact at-age values for each item. Results are expressed in terms of a cumulative point score based on the number of the child's successes. Bayley's scale is considered to be fairly well standardized and to include a sufficient variety of items. However, as indicated earlier, the author found that the test had poor predictive value for her group of children.

On the basis of examinations conducted on "several hundred" children at the Iowa Child Welfare Research Station, Fillmore\(^3\) published the Iowa Tests for Young Children in 1936. These tests, including forty-nine items, covered the age range of four months to two years, but they were never adopted extensively for infant testing in clinical practice. One possible explanation is the very small number of items presented for the first twelve months - only ten items were given for the period from five to nine months, two of them for the six month level. Originally, the test items were arranged according to the percentage of successes, and mental age credits were found for each item by dividing the age range covered by the number of items in the particular age range. This system was then discarded in favor of a point scale with the items ranged in the order of difficulty, according to

Thurstone's Method of Absolute Scaling. According to the author, the Iowa Tests measure some ability which increases with age. In agreement with other investigations of infant scales, Fillmore found that her tests failed to correlate highly with later Stanford-Binet I.Q.'s. In explanation, she suggested that infant tests are performance tests which measure something which is not highly related to those abilities expressed in the verbal responses of the tests at later ages. No studies, aside from the author's original presentation, have reported on this scale.

The most recently developed infant scale has been the Northwestern Infant Intelligence Scale, originally described in 1943 by Gilliland and Shotwell,35 and later presented in its revised form by Gilliland.36 The authors began their work on the scale at the request of a child care institution interested in determining the suitability for adoption of the very young infant. A large number of items from existing scales, mainly those of Gesell, Cattell and Buhler, as well as some new tests, were administered to approximately five hundred children, mostly institution babies. On the basis of tests for 276 babies whose records were complete, a final revision was made which consists of forty items arranged in two overlapping series, covering the age groups of four to fourteen weeks and thirteen to thirty six weeks. An I.Q. can be computed for any age, the raw score of the test being the number


of items passed. In placing his items, Gilliland used the method of increase in percentage of passes with chronological age. If seventy five per cent of the infants at a given age level could pass an item, it was considered to be correctly placed. Gilliland claimed evidences of high validity for his scale on the basis of later Stanford-Binet results, but he did not publish data to amplify this statement.

The Cattell Scale, published in 1940, will be discussed at length in the next chapter. Only five studies, apart from the original presentation of Cattell, have been published in which a report is given of some use of the test. Two of these studies investigated environmental influences on test performance, two investigated the relationship between test performance and satisfactoriness of the examination, and one article, actually a detailed report on psychological examining, presented some limited Cattell findings for comparison with the standardization group.

Fischer,37 conducted a follow-up study of later development of sixty-two infants who had been cared for from birth to a period beyond six months in a maternity home, and whose mean Cattell IQ at six months was 76.11. Thirty-six of these children who were then placed in adoptive homes were later found to have a mean Cattell IQ of 97.54. On the basis of an analysis of the six month test records and the behavior reported at the time of the examination, Fischer concluded that a definite "hospitalism" syndrome

occurs in a large number of institutionalized babies of six months which inhibits cooperation in developmental examinations but which is not yet in the nature of an irreversible pattern. As a part of her analysis, Fischer made a percentagewise comparison of the total successes incurred by her group of infants on the six and seven month items with those reported by Cattell for the six month infants in her standardization group. She found that her children performed significantly below Cattell's group on all of the items except Item 3 on the six months level (reaction to own mirror image, level I). Fischer emphasized the concurrence of her findings with those of other well known reports on the retarding effects of institutional care on young children.

Klatskin examined the Cattell test performances of 316 infants between the age of 11.5 and 13.5 months followed in the Yale Rooming-in project. She found the mean Cattell IQ for her group of children to be 112. Klatskin analyzed her records in terms of the percentage passing the items. When these percentages were compared with those of Cattell, significantly higher percentages for the Yale group were found on most items, with the constant exception of vocabulary. The author offered the fact of flexible methodology in rearing the Yale group as a partial explanation, but she emphasized the need for re-standardization of infant tests and caution in interpreting records. Klatskin also investigated but found no significant

---

relationship between test performance and sex; ordinal position in the family, whether the infant was breast fed, and satisfactoriness of the examination.

Carter and Bowles,39 reported that percentages of successes on the Cattell and Gesell scales tended to be consistently higher than those reported by the authors of the scales, when these tests were administered to two and three month infants in the Wichita Guidance Clinic. Sixty six two month old infants attained an average Cattell test age of 2.8 months and sixty two three month old infants had an average Cattell test age of 3.7 months. Carter and Bowles concluded that, to a considerable extent, these differences appeared to result from different examining procedures. They also offered two criticisms of the Cattell Scale in relation to their data: (1) The items placed at the two, three and four month levels are heavily weighted with visual tasks, often resulting in high scores for infants who have unusual visual alertness and responsiveness, but only average or even below average abilities in other areas, and (2) the failure of Cattell to make allowances for refusal of tasks decreases the value of the quantitative scores.

Although Klatskin failed to find any relationship between test performance and the satisfactoriness of the examination, two studies have indicated positive findings in this area. Emphasizing the importance of the

infant's response to the test situation, Escalona\textsuperscript{40} reported an attempt to demonstrate an assumed positive relationship between "optimal functioning" of the infant during the administration of an intelligence examination and the predictive value of the examination. Seventy two children were tested in early infancy with the Cattell Scale and the Gesell Normative Schedules, and a judgment was made in each case as to whether such functioning had been elicited from the child. These children were later retested from one to six times. When the two groups of test-retest series were compared for predictive accuracy, it was found that predictive value was greater for that group of tests initially considered to have elicited optimal functioning. Of the non-optimal group, only nineteen per cent were found to remain in the same intelligence range upon retesting, fifty three per cent moved into the adjacent higher range, and twenty seven per cent were in ranges one step removed or more. In determining the quality of test functioning, the following aspects were recorded: (1) quality of the child's motility; (2) his fatigability; and capacity for muscular relaxation; (3) respiratory and circulatory phenomena; (4) quality of responsiveness to objects and persons, and (5) degree of differentiation shown in extra test behavior. As a general conclusion, Escalona urged the "Gestalt" view of psychological testing, in which the infant's test behavior is considered in conjunction with his actual performance, for more effective prediction of future developmental events.

\textsuperscript{40} Escalona, "The Use of Infant Tests for Predictive Purposes," \textit{Kenninger Clinic Bulletin}, XIV, 1950, 117-128.
A study by Gallagher on the question of infant responsivity in
the test situation reported findings in essential agreement with those of
Escalona. Forty-three infants ranging in age from 4.1 months to 24.1 months
were placed in two groups for retests - a Mandatory Retest group, including
all of those infants who were suspected of not doing their best on the origi-

nal test, and a Routine Retest group, in which no special reason for retesting
was noted. The Mandatory Retest group made a mean gain of 8.53 IQ points
on the retest, a difference significant at the one percent level of confi-
dence. The mean IQ on the first test was 88.05, on the second test, 96.58.
The Routine Retest group made no significant gains in scores. The mean IQ
for this group on the original test was 100.62, the mean IQ for the retest,
101.25. Changes in range placement of IQ were reported for both groups but
much less for the Routine Retest group.

A study similar to the present research was conducted on the six
month items of the Gesell schedules by Nelson and Richards, as a part of
the longitudinal studies conducted by the Fels Research Institute of Antioch
College. A percentagewise comparison of the successes of 123 six month old
infants on the forty-eight Gesell items placed at that age level (the early
version of the test was used) was made with Gesell's own findings, as well


as with findings on the same items by other researchers. The authors found the items to be passed by a significantly greater number of their infants. Further, since twelve of the forty-eight items were passed by ninety to one hundred per cent of the Fels infants, they concluded that there was an over-abundance of easy tests, and hence the distribution of items in terms of difficulty was neither normal nor even. The authors suggested that the significant differences might in part be explained by differences in the size of the groups compared (groups ranged from fifty in number to one hundred and thirty-nine), and by differences in techniques for evaluating success on individual items.

Nelson and Richards investigated the possibility of sex differences in performance but they found no reliable difference between the sexes in total test scores. In the matter of prediction, the authors found a validity coefficient of .46 between the total Gesell test score at six months and the Stanford-Binet at three years for a group of forty-eight children. This is the highest correlation reported in the literature between tests during the first year of life and those at later ages.

Summary of the Literature

One of the important characteristics of the infancy period which bears upon the general problem of evaluating infant mental development is the transient nature of infant behavior. Researchers at the Yale Clinic of Child Development have concluded as a result of their observations that behavior
growth proceeds fully as rapidly as physical growth and that even with simplification, at least three developmental intervals - four, six and nine months - must be recognized in evaluating rate of maturity during the first year. Normative tests of behavior can be utilized since behavior growth proceeds in accord with laws of orderly developmental sequence.

The fluctuations occurring in infant growth are considered to have important implications for intelligence test construction. Lower reliability at certain age levels and individual variations in IQ curves have been explained in terms of the diffuse nature of new behavior and changes in the tempo of development. Anderson has suggested that the changing nature of infant growth lowers the predictive value of infant scales since there necessarily occurs an irregularity in the amount of a given function that can be measured at different age levels.

Because of the validity criterion of progression with age, infant scales have always included many motor items. However, motor development has not been found to correlate well with intelligence at later ages. Most test authors have tried to reduce the influence of motor behavior on test performance, but it is difficult to exclude entirely because of the close relation between motor and other types of behavior during the first year of life.

There is a similarity in the kinds of items included in infant intelligence scales because of the limited range of infant behavior and because test authors have all borrowed heavily from the normative items devised by Gesell. Most authors have avoided presenting any logical explanation of
their items. Gesell separated his tests into four general areas of behavior - motor, language, adaptive and personal-social - but he emphasized the impossibility of drawing hard and fast lines among these areas. A study by Bayley presented additional evidence that a response to a given test situation requires activities of more than one kind, making classification difficult. She found that sensory and motor items predominated during the first six months of life, while adaptive items gradually gained prominence after that period. This general classification is probably applicable to most of the infant intelligence scales.

The available infant scales have been found to have poor predictive value. In explanation, Bayley has suggested that there is no clear cut relation between intelligence in the infant and intelligence in the adult, even though the later behavior may depend in large part upon the earlier behavior. She further indicated that the most the existing tests can do is to measure development at successive age levels and offer an appraisal of development only for the level tested, rather than to measure a unit function of intelligence that extends from birth throughout life. This is a serious restriction since infant intelligence tests are used largely to predict intelligence at later ages.

Since 1904, when Binet published a few items suitable for evaluating infant behavior, there has been a large number of test items presented either as supplements at the lower end of tests for more advanced age levels or as separate infant scales. Kuhlmann extended the Binet scale from three
years down to three months in his revision and offered items for several levels during the first twelve months. The most widely known infant tests are those of Gesell, published originally in 1925 and in their latest form in 1947. The Gesell Developmental Schedules comprise a normative scale, rather than an intelligence test in the strict sense, but a developmental quotient evaluating the child's developmental level in the four different areas of growth can be derived.

Another normative scale was published in Vienna in 1928 by Hetzer and Wolf, and revised along the lines of the Binet Scale in 1930 by Buhler. Some of the Buhler items have been incorporated in more recent infant tests, but her scale has never been widely used in this country.

Other infant scales have included the Linfert-Hierholzer Scale (1928), the California First Year Mental Scale (1933), the Iowa Tests for Young Children (1936), The Cattell Infant Intelligence Scale (1940), and the Northwestern Infant Intelligence Scale (1943). The Linfert-Hierholzer, California and Iowa Scales are point scales, whereas the Cattell and Northwestern tests are age scales.

There are only five published studies reporting on the Cattell Scale. Fischer utilized Cattell test performances to demonstrate the retarding effects of institutional care on infants up to the six months level, describing a definite "hospitalism" syndrome which occurs. Her infants performed at a significantly lower level on most items than the infants in Cattell's standardization group. Two studies reported significantly higher
results of the Cattell Scale than the author reported, when working with more comparable infant populations. A study by Carter and Bowles presented results with two and three month old adoptive infants. Klatskin, working with infants at the twelve and thirteen month levels raised in accord with flexible methodology, found her infants to achieve at significantly higher levels of success on the majority of items when compared with Cattell's infants.

Two of the studies - those of Escalona and Gallagher - reported a significant relationship between the responsivity of the infant during testing and the predictive value of the examination. On the basis of her findings, Escalona urged a "Gestalt" view of testing, in which the infant's test behavior is considered in conjunction with his actual performance, for purposes of more effective prediction.

In a study similar to the present research, Nelson and Richards made a percentagewise comparison of two groups of infants - a group of infants tested at six months of age in the Fels longitudinal research and Gesell's normative six month group - and found that the items were passed by a significantly greater number of the Fels infants. Since several of the items were passed by from ninety to one hundred per cent of the Fels infants, the authors concluded that there was an overabundance of easy tests at that age level of the Gesell scales.
CHAPTER III

DESIGN OF THE RESEARCH

The Cattell Infant Intelligence Scale was published as an outgrowth of a longitudinal study of child health and development conducted at the School of Public Health of Harvard University. Constructed as an age scale and a downward extension of the Revised Stanford-Binet, Form L, the scale covers the age range of two to thirty months. Since Stanford-Binet items are interspersed with other items between the ages of twenty-two and thirty months, the author proposes that a continuous intelligence scale from early infancy to maturity has been attained. Five regular items and either one or two alternate items are presented for age levels one month apart during the first year, two months apart during the second year, and for the additional age levels of twenty-seven and thirty months.

In standardizing the test, 1346 examinations on 274 children were used. The tests were administered at the ages of three, six, nine, twelve, eighteen, twenty-four, thirty and thirty-six months (Stanford-Binet). It was not possible to test all of the children at those ages but they averaged five examinations each. Percent passing was the only method of item analysis used by Cattell in placing her tests on the scale. For the age levels between the standardization ages - two, four, five, seven, eight, ten, and eleven months during the first year - items were placed by estimation, based
on the percent passing at the immediately preceding and following standardized age levels.

Items were adapted largely from Gesell and Buhler. A lesser number were taken from other sources. Items were eliminated from the scale if they failed to show sufficient increase in the percentage of passes from one age group to another, or if they increased irregularly in the number of passes from age to age, showed plateaus or failed to approach closely the one hundred percent mark at any age. Additional reasons for eliminating items were the following: (1) Items which were difficult to administer or score, or which required an undue amount of subjective judgment on the part of the examiner; (2) items which did not hold the attention of the child; (3) items requiring cumbersome apparatus; (4) items which were thought to be unduly influenced by home training; (5) items planned to test control of the large muscle groups; (6) items which appeared to test abilities similar to those covered by other items at the same age level, and (7) items at age levels for which a sufficient number of more or equally satisfactory items were available.

On the basis of Stanford-Binet results on thirty-five children whose records were complete, Cattell rearranged her items to bring the median IQ for each age level as close as possible to the median IQ of 106 obtained on the Stanford-Binet at thirty-six months. She found that at no age did the median IQ differ by more than two points from the Stanford-Binet median. The mean Stanford-Binet IQ at thirty-six months for these children was 105.
Cattell found her scale to be of doubtful validity before twelve months, but of increasing validity thereafter. She measured validity in terms of the scale's ability to predict later Stanford-Binet scores. For the age levels of six, nine and twelve months and the Revised Stanford-Binet at thirty-six months, the correlations were .10, .34 and .18, respectively. The median IQ changes were found to be greater before than after twelve months. The correlations between the age levels of twelve, eighteen, twenty-four and thirty months were much higher, .56, .67, .71 and .83, respectively.

The corrected odd-even reliability coefficients found by Cattell were as follows: .56 at three months, .88 at six months, .86 at nine months, .89 at twelve months, .90 at eighteen months, .85 at twenty-four months, .71 at thirty months and .87 with the Stanford-Binet at thirty-six months.

According to Cattell, the fairly rigorous requirements for enrollment in her study group probably resulted in a standardization sample somewhat above the general population composition, a conclusion which she considered to be partially substantiated by the mean IQ of 105 obtained at thirty-six months. In general, Cattell characterized her group as being of the lower "middle classes." Enrollment requirements included good physical health and normal delivery, a background of primarily North European stock, more or less permanent employment of the father, and willingness of the mother to cooperate with the study group over a period of years. A few of the parents were professional people, but the majority were employed in such positions as policemen, clerks, storekeepers, and the like.
The administration of the Cattell Scale is similar to that of the Stanford-Binet, with the exception that serial testing is permitted in order to secure the infant's best efforts and attention. Several items can be scored frequently on the basis of observation of one activity, such as the degree of fine motor coordination displayed by the child in securing a small sugar pellet. The testing manual includes a complete description and an accompanying photograph for every item, lessening the possibility of inadequate administration and scoring. Record forms for the complete scale are available.

Scoring is the same as for the Stanford-Binet. A basal age is established and to this month level are added additional credits for all succeeding successes in computing the mental age. The test is continued until two levels have been completely failed. Since there are five items placed one month apart during the first year, each item receives one fifth or .2 of a month credit. Thus, an infant who achieves a basal age at the six month level and has three additional successes beyond that level has a mental age of 6.6 months. Similarly, the chronological age is estimated in terms of tenths of months, every three days comprising an additional one tenth of a month. The IQ is computed in the same manner as for the Stanford-Binet.

Plan of Present Research

The data for the present study were taken from the Cattell test records of infants examined in the Guidance Department of the Chicago Catholic Charities for the purpose of determining their suitability for adoption. The
infant testing program has been a part of the agency adoption procedure since 1948 and the records of several hundred administrations of the Cattell Scale were available. However, in line with Cattell's procedure in standardizing her scale, only those records of infants who had been tested within one week of their month birthday - for the purposes of this study, six and seven months - were selected. Additional criteria used in selecting records were as follows: (1) Reasonable indication, based on a consideration of test behavior and the opinion of the examiner included in the report accompanying each test record, that the responsiveness of the infant permitted complete and, in so far as could be ascertained, valid testing; (2) placement in an adoptive home during the first month of life, and (3) full term gestation. In all, one hundred and fifty-eight test records of six-month-old infants and eighty test records of seven-month-old infants were adjudged to be suitable for analysis.

As in Cattell's group, several factors were operative which prevented an entirely random sample but the present study group was probably typical of the infants for whom the scale is largely being used. The policy of the agency precludes early placement of infants for whom adoption is contraindicated by reason of birth injury, serious physical disorders or background incidence of mental illness. The infants were, with very few exceptions, born out of wedlock in one of the agency's maternity hospitals. The girls who request agency care for illegitimate pregnancy have usually been found to be of average to low average intellectual endowment and socio-economic status. The adoptive homes selected by the agency, on the other
hand, can be described as "middle class" in character. It is probable that a minimum of average mental ability is necessary to fulfill the agency requirements for employment and living standards. A genuine desire for a child can be assumed for the adoptive families selected after careful investigation. Most acceptable applicants must wait for two or three years before a child is placed with them. For these reasons, the environment of a typical adoptive home, in terms of the physical advantages and emotion satisfactions the infant receives, can probably be considered as more than ordinarily motivating.

All of the tests had been administered by the same qualified examiner over a period of two and one half years. The infants were brought to the clinic by one or both of the adoptive parents, and the physical environment in which the examinations were conducted was the same for all of the babies. In accord with the test directions, all of the six and seven month items were administered while the infant was in a sitting position, usually on the lap of the adoptive mother. The testing equipment was that specified in the testing manual. For the six and seven month levels, this includes four red wooden one inch cubes, a large aluminum cup, a large mirror, a metal door key, a red sugar pellet, an embroidery ring, and a sheet of onion skin paper.

The items which were examined for the purpose of comparative analysis were those placed at the six and seven month levels of the Cattell Scale. Cattell presented five regular items for each of these levels, and one alternate item for six months and two alternate items for seven months. Since the
alternate items were not administered routinely with the adoptive infants, they were not included in the analysis. The exception to this was alternate item at the six month level, which was administered as a substitute for the fifth regular six month item - persistent reaching - and which therefore was analyzed in place of the regular item. The test items for these two age levels are as follows:

**Six Months**
1. Cube, attains
2. Cup, lifts
3. Mirror, approaches and manipulates
4. Reaching, unilateral
Alt. a. Cube, approaches second

**Seven Months**
1. Pellet, attempts
2. Mirror, pats and smiles
3. Ring, inspects
4. Cube, takes two
5. Paper, exploits

In accord with the prescribed scoring procedure, the infant's performances on these items were scored plus or minus at the time of the administration of the examination, on the basis of the successful or unsuccessful nature of his responses. This permitted a comparison of the successes on each item of the total group, in terms of percentages, with those reported by the author for the standardization group.

The age level of six months was selected as the chief focus of study for several reasons. Infants between the ages of six and seven months are considered to be fairly stable, as far as the quality of their test performance is concerned. They are typically very much interested in their
surroundings and their attention to the test objects is probably more quickly and readily elicited than at any other age during the infancy period. These factors increase the possibility of obtaining valid test scores. Further, their attention span for individual objects is sufficiently long to permit adequate observation by the examiner. Cattell found that six months was the single age level below twelve months to have the highest correlation with later Stanford-Binet scores. As an additional practical reason, since the majority of infants referred to the Guidance Department for testing are about six months of age, there was a greater number of records for this group available and it was felt that findings for this age group would be of some value when using the Cattell Scale in the future.

The test records of eighty seven-month-old infants were included for study to permit an inter-group comparison of performance on the six and seven month items, as a further means of evaluating the suitability of placement of these items. Since an incidental analysis of sex differences in the performances of the six month infants was also made, the total six month group was separated into two groups of eighty-two boys and seventy-six girls. Thus, a total of four groups of adoptive infants was studied.

As a preliminary analysis, the test performances of the four groups of adoptive infants were analysed in the following ways. The mean scores on the total scale and the standard deviations were obtained for each group. The mean IQ's for the six and seven month adoptive infants were compared for significant differences, as were the mean scores for the six-month-old boys
and the six-month-old girls. The formula for determining the standard error of the difference between two uncorrelated means was used. The standard test was applied to determine the significance of the differences.

As the main part of the study, the percentages of success incurred by the six-month-old adoptive infants on the six and seven month items were calculated and compared with those reported by Cattell for the six-month-old infants in her standardization group. The formula for determining the standard error of the difference between uncorrelated percentages was used. The standard test of significance was applied to determine the significance of the differences between percentages.

For the related investigation of suitability of the six and seven month item placement, the performances of the six and seven month adoptive infants on these items were analysed, using the same statistical methods.

Finally, for the purpose of determining sex differences in performances on individual items, the performances of the six-month-old adoptive boys and the six-month-old adoptive girls were analysed, using the same statistical methods.
CHAPTER IV

ANALYSIS OF RESULTS

The mean scores and related statistics for the four groups of adoptive infants are presented in Table I.

TABLE I

MEANS, STANDARD DEVIATIONS AND RELIABILITY FOR THE FOUR GROUPS OF ADOPTIVE INFANTS

<table>
<thead>
<tr>
<th>Group</th>
<th>Number</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Six months</td>
<td>158</td>
<td>112.89</td>
<td>9.84</td>
<td>.782</td>
</tr>
<tr>
<td>Boys</td>
<td>82</td>
<td>112.64</td>
<td>10.30</td>
<td>1.13</td>
</tr>
<tr>
<td>Girls</td>
<td>76</td>
<td>113.15</td>
<td>8.10</td>
<td>.929</td>
</tr>
<tr>
<td>Seven months</td>
<td>80</td>
<td>107.09</td>
<td>7.80</td>
<td>.872</td>
</tr>
</tbody>
</table>

The mean score for the total group of six-month-old adoptive infants was 112.89. Cattell did not report mean scores or standard deviations for her standardization age groups. Therefore, only the mean Stanford-Binet IQ of 105 which she found for her group of thirty-five children tested at thirty-six months is available for reference. Since Cattell reported her median IQ's, the median IQ for the six-month infants was computed. It was
found to be 113.49, as compared with Cattell's median of 108 (Cattell N was 103).

The mean IQ for the six-month-old adoptive boys was found to be 112.64, as compared with the mean of 113.15 for the six-month-old adoptive girls. Since the difference was not statistically significant, the sex difference was not influential for this group of infants on their total performance on the scale.

The mean IQ for the group of seven-month-old adoptive infants was 107.09. When compared with the mean IQ of 112.89 for the six-month-old adoptive infants, the t value of 4.95 was found to be significant at the .001 level of confidence. The difference between the two means represents a decided drop in general level of performance, especially when the brief time interval of one month is considered. It appears to be a function of the scale itself, since it is unlikely that the two infant populations differed significantly. For definitive conclusions, this trend should be critically examined through further study. One possible explanation is that the items at the seven month level are more correctly placed than those at six months, bringing the total scores downward. However, the findings on the seven month items indicate that this was not the case with the present group of infants.

Another explanation is related to Bayley's findings that test items tapping an adaptive type of behavior gain prominence after the six month level. In other words, the six-month-old infants would have the advantage
of being the most advanced age level performing on items largely sensori-motor in character, whereas the seven-month-old infant represents the lowest age level performing within the range of the more truly adaptive items. It is only possible to conjecture with the present limited data which does not include the full range of performance of the two groups.

Table II lists the percent passing of the adoptive and Cattell six-month-old infants on the six and seven month items.
### TABLE II

**COMPARISON OF PERCENT OF ADOPTIVE AND CATTELL SIX-MONTH-OLD INFANTS PASSING SIX AND SEVEN MONTH CATTELL ITEMS**

<table>
<thead>
<tr>
<th>Test items</th>
<th>Percentages</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Adoptive</td>
<td>Cattell</td>
<td>t</td>
<td>P</td>
</tr>
<tr>
<td></td>
<td>(N = 158)</td>
<td>(AvN=190)*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Six month tests:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Cube, attains</td>
<td>100</td>
<td>79</td>
<td>6.77</td>
<td>.001</td>
</tr>
<tr>
<td>2. Cup, lifts</td>
<td>99</td>
<td>79</td>
<td>6.29</td>
<td>.001</td>
</tr>
<tr>
<td>3. Mirror, approaches</td>
<td>85</td>
<td>80</td>
<td>1.19</td>
<td>-</td>
</tr>
<tr>
<td>4. Reaching, unilateral</td>
<td>82</td>
<td>63</td>
<td>4.00</td>
<td>.001</td>
</tr>
<tr>
<td>Alt. a. Cube, approaches second</td>
<td>94</td>
<td>77</td>
<td>3.11</td>
<td>.01</td>
</tr>
<tr>
<td>Seven month tests:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Pellet, attempts</td>
<td>80</td>
<td>58</td>
<td>4.27</td>
<td>.001</td>
</tr>
<tr>
<td>2. Mirror, pats and smiles</td>
<td>58</td>
<td>50</td>
<td>1.43</td>
<td>-</td>
</tr>
<tr>
<td>3. Ring, inspects</td>
<td>58</td>
<td>49</td>
<td>1.34</td>
<td>-</td>
</tr>
<tr>
<td>4. Cube, takes two</td>
<td>85</td>
<td>45</td>
<td>5.97</td>
<td>.001</td>
</tr>
<tr>
<td>5. Paper, exploits</td>
<td>80</td>
<td>40</td>
<td>5.80</td>
<td>.001</td>
</tr>
</tbody>
</table>

*Represents an average of the numbers of infants tested with these items.

It is evident from Table II that the adoptive infants scored significantly above Cattell's infants on the majority of the six and seven month items. Six of the total of ten items were easier for the adoptive infants to
a very significant degree. On only three of the items - the tests involving the two levels of response to the mirror image and inspection of a ring - did the two groups of infants achieve more comparable levels of performance. On each of these items the adoptive infants achieved a higher percentage of success but the differences were not statistically significant. Therefore, it can be safely concluded that the items of these two age levels were in general too easy for the group of adoptive infants.

Further, since three of the six month items were passed by ninety to one hundred percent of the group and since the level of performance by the six-month-old infants on three of the five seven month items was on a par with that usually accepted as appropriate for the at-age group (above seventy percent), the present findings would lend some validity to the conclusion that these items have been incorrectly placed on the scale.

The findings presented in Table III which lists the percents passing of the six and seven-month-old adoptive infants on the six and seven month items lends further weight to this conclusion. Since the seven-month-old infants scored significantly above the six-month-old infants on all but three items (items 1, 2, and Alt. a, on the six month level, on which the six-month-old infants also scored in the ninety to one hundred percent range), the present data supports Cattell's conclusion that these items meet the criterion of increase in ability to pass with age. However, the high percentage of passes by the seven-month-old infants on the seven month items (percents passing on items 1, 4, and 5 were in the ninety to one hundred percent range)
indicates that this age level of the scale was too easy for the seven-month-old adoptive infants.

TABLE III

COMPARISON OF PERCENT OF SIX AND SEVEN-MONTH-OLD ADOPTIVE INFANTS PASSING SIX AND SEVEN MONTH CATTELL ITEMS

<table>
<thead>
<tr>
<th>Test items</th>
<th>Percentages</th>
<th></th>
<th></th>
<th>t</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Six Month</td>
<td>Seven Month</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Infants</td>
<td>Infants</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(N - 158)</td>
<td>(N - 80)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Six month tests:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Cube,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>attains</td>
<td>100</td>
<td>100</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>2. Cup,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lifts</td>
<td>99</td>
<td>99</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>3. Mirror,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>approaches</td>
<td>85</td>
<td>95</td>
<td>2.69</td>
<td>.01</td>
<td></td>
</tr>
<tr>
<td>4. Reaching,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>unilateral</td>
<td>82</td>
<td>91</td>
<td>2.12</td>
<td>.05</td>
<td></td>
</tr>
<tr>
<td>Alt. a. Cube,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>approaches second</td>
<td>94</td>
<td>96</td>
<td>1.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seven month tests:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Pallet,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>attempts</td>
<td>80</td>
<td>95</td>
<td>3.78</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>2. Mirror,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pats and smiles</td>
<td>58</td>
<td>76</td>
<td>2.82</td>
<td>.01</td>
<td></td>
</tr>
<tr>
<td>3. Ring,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>inspects</td>
<td>58</td>
<td>80</td>
<td>3.72</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>4. Cube,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>takes two</td>
<td>85</td>
<td>95</td>
<td>2.69</td>
<td>.01</td>
<td></td>
</tr>
<tr>
<td>5. Paper,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>exploits</td>
<td>80</td>
<td>90</td>
<td>2.18</td>
<td>.05</td>
<td></td>
</tr>
</tbody>
</table>
There are two possibly influential factors which cannot be excluded with certainty from the present study and which preclude a definite conclusion that the items at these two age levels, although too easy for the two adoptive age groups, have been placed too high on the scale. These are (1) real differences between the adoptive and the Cattell infants in their ability to perform on the items, and (2) differences in the standards used to evaluate success on the individual items.

The sample of adoptive infants included in the present study group may have been superior in ability to the infants in Cattell's standardization group, who were in turn considered to be above the population as a whole. This is impossible to determine exactly, but the differences in median scores obtained for the two groups suggests that this may have been the case.

It is unlikely that there were significant differences in the degree of motivation afforded to the adoptive and the Cattell infants. The previously cited study by Fischer pointed out the importance of environmental stimulation as a determinant of test performance, and it is true that the adoptive infants as a group were in all probability well motivated by their physical and emotional environments. However, there is no valid reason for concluding that the cooperative mothers with whom Cattell was dealing in her longitudinal study were any less motivating in their handling of their children than the adoptive mothers of the infants in the present study group.

A more likely factor of influence is that of individual differences - in the direction of greater leniency - in evaluating success on the test items.
In Klatskin's study, two examiners evaluated the same items but no significant differences were found in their results. However, in the studies conducted by Nelson and Richards on the Gesell items, and by Carter and Bowles on the Cattell Scale, both groups of authors indicated differences in scoring standards as possible factors in their higher results. Cattell indicated that the most frequent cause of error she found in working with the scale was the tendency to give the infant the benefit of the doubt, and to credit for chance rather than for purposeful reactions. It is very difficult to be entirely accurate in scoring infant test responses. For example, item 1 on the six month level demands a voluntary rather than a reflex grasp of the one inch cube for success, but frequently even careful observation leaves some room for doubt as to the quality of the infant's response. Similar complications occur in the scoring of many of the other test items.

Therefore, the factors of real differences in ability between the infant groups and individual differences in scoring may have been influential in the present higher results. However, as a general conclusion the findings of the present study add to the growing body of evidence that the scale is in need of further standardization.

It is interesting to note the possibility of some support in the present data for one of the conclusions of Fischer's study. She concluded on the basis of her results that the grasping activities of the six-month-old

---

1 In a letter to the writer.
Infants are especially affected by a lack of stimulation. The six and seven month items which specifically involve grasping behavior are items 1, 2, 4 and alt. a. on the six month level and items 2 and 4 on the seven month level. Fischer found that only fifteen and thirty-one percent of her infants were successful on the six month items 1 and 2, respectively, and none of her infants solved the other grasping items. If the reverse should be true—that is, if a high degree of motivation would elicit strong grasping behavior, the well motivated adoptive infants could be expected to perform well on items demanding grasping activity. An inspection of Table II reveals that on the items in question the adoptive infants did attain a high level of success.

A comparison of the performances of the six-month-old adoptive boys and girls on the six and seven month items revealed that they were alike in their performances on all of the items with the exception of item 3 on the seven month level—involving prolonged inspection of an embroidery ring. The difference in performance on this item was significant at the .05 level. No explanation can be offered for this difference, except that in a comparison of numerous differences performance on one item can conceivably differ significantly through chance alone. Since the boys and girls were found to perform alike on the total scale and on nine of the ten individual items, it can be concluded that sex was unrelated to test performance for the present group of six-month-old adoptive infants.

3 See Table IV in appendix.
CHAPTER V

SUMMARY AND CONCLUSIONS

Because of similarity in construction to the Revised Stanford-Binet, Form L, and convenience in administration and scoring, the Cattell Infant Intelligence Scale, published in 1940, is now in common use in the clinic. It is mainly used as an aid in evaluating the suitability of young infants for adoption. Very little research on the scale beyond the original work of the author has been reported in the literature. However, there is some evidence from two studies of infant groups that the scale is in need of re-standardization. In both studies, by Klatskin,\(^1\) and by Carter and Bowles,\(^2\) high mean scores on the total scale were found, and Klatskin further reported significantly higher results on individual items when her infants were compared with the infants in Cattell's standardization group.

The purpose of this paper was to extend the type of comparative item analysis done by Klatskin and study the level of performance of a group of six-month-old infants on the total scale and on the individual items placed at the six and seven month levels of the scale. The performances of a group of seven-month-old infants were also studied as an additional investiga-

---


tion of the suitability of placement of these items. A secondary purpose of the study was to explore further the findings from previous studies that sex was unrelated to infant test performance.

Accordingly, by employing appropriate statistical procedures, the performances on the total scale and on the individual items of the two age levels by a group of six-month-old adoptive infants were compared with the levels of performance reported by Cattell for her group of infants. Similar comparisons were made between the performances of the six and seven-month-old adoptive infants, and between the performances of six-month-old adoptive boys and girls. On the basis of these comparisons, the following conclusions and results were obtained:

1) Significantly higher results on the majority of the individual six and seven month items were attained by the six-month-old adoptive infants, as compared with Cattell's infants. An inspection of the median scores for the two groups of infants indicated that the adoptive infants also performed at a higher level on the total scale than Cattell's infants. These findings add to the growing body of evidence that the scale is in need of more extensive standardization.

2) Since three of the five six month items were passed by ninety to one hundred percent of the six-month-old adoptive infants, this can be taken as evidence that the at-age level was too easy for the six-month-old adoptive infants. Further, since three of the five seven month items were passed by ninety to one hundred percent of the seven-month-old adoptive
infants, this can also be taken as evidence that the at-age level was too easy for the seven-month-old adoptive infants.

3) The extent of influence of two factors on the above findings could not be demonstrated exactly and therefore could not be excluded from consideration. These are (1) real differences between the Cattell and adoptive infants in ability and degree of motivation, and (2) differences in scoring standards. It was concluded that the latter factor in particular could have contributed to the higher results obtained in the present research.

4) The significantly higher percentages of success attained by the seven-month-old adoptive infants on seven of the ten items at the two age levels, when compared with the percentages attained by the six-month-old infants on these items, supports Cattell's inclusion of these items in the scale on the basis of increase in the ability to pass with age.

5) Sex was unrelated to performance on the total scale and to performance on individual items.
## APPENDIX I

### COMPARISON OF PERCENT OF SIX-MONTH-OLD ADOPTIVE BOYS AND GIRLS PASSING SIX AND SEVEN MONTH CATTELL ITEMS

<table>
<thead>
<tr>
<th>Test Items</th>
<th>Percentages</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Boys (N = 82)</td>
<td>Girls (N = 76)</td>
<td>t</td>
<td>p</td>
</tr>
<tr>
<td>Six month tests:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Cube, attains</td>
<td>100</td>
<td>100</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2. Cup, lifts</td>
<td>99</td>
<td>99</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3. Mirror, approaches</td>
<td>88</td>
<td>82</td>
<td>1.07</td>
<td>-</td>
</tr>
<tr>
<td>4. Reaching, unilateral</td>
<td>83</td>
<td>83</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Alt. a., Cube, approaches</td>
<td>93</td>
<td>92</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>second</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seven month tests:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Pellet, attempts</td>
<td>79</td>
<td>80</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2. Mirror, pats and smiles</td>
<td>57</td>
<td>59</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3. Ring, inspects</td>
<td>51</td>
<td>67</td>
<td>2.10</td>
<td>.05</td>
</tr>
<tr>
<td>4. Cube, takes two</td>
<td>87</td>
<td>85</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5. Paper, exploits</td>
<td>74</td>
<td>86</td>
<td>1.14</td>
<td>-</td>
</tr>
</tbody>
</table>
BIBLIOGRAPHY

BOOKS

Bayley, Nancy, The California First Year Mental Scale, Berkeley, 1933.
Cattell, Psych, The Measurement of Intelligence of Infants and Young Children, New York, 1940.
Fillmore, Eva, Iowa Tests for Young Children, Iowa City, 1936.
Gesell, Arnold, The Mental Growth of the Pre-School Child, New York, 1925.
Kuhlmann, Frederick, A Handbook of Mental Tests, Baltimore, 1922.
ARTICLES


The thesis submitted by Patricia Bledsoe has been read and approved by three members of the Department of Psychology.

The final copies have been examined by the director of the thesis and the signature which appears below verifies the fact that any necessary changes have been incorporated, and that the thesis is now given final approval with reference to content, form, and mechanical accuracy.

The thesis is therefore accepted in partial fulfillment of the requirements for the degree of Master of Arts.

October 1, 1954  
Date

Edmund P. Marx  
Signature of Adviser