The list of entries in this bibliography represents a selected sample, not an exhaustive one, of assessment instruments for prekindergarten and kindergarten children. Some standard instruments which are commercially available are included, but the majority are research instruments used successfully on an experimental basis by investigators. With few exceptions they are individually administered. The selection covers the major assessment areas in early childhood research, which include measurement techniques geared specifically to the urban child from a low socioeconomic background. Other pertinent considerations include test validity and reliability data, childlike appeal, ease of administration, and scoring. The entries are organized into specific test categories indicating the type of behavior measured, such as: cognitive status, cognitive abilities; memory, attention, and learning; perceptual skills (visual and auditory); reading readiness; techniques for assessing cognitive style characteristics; and personal social development. Each entry is labeled R (research instrument) or C (commercially available), and I (individual administration) or G (group administration). Further information is provided on publishing houses or individuals to contact, in the case of research instruments. (LH)
AN ANNOTATED
BIBLIOGRAPHY
OF MEASUREMENTS
FOR YOUNG CHILDREN

by Barbara Berger

October 1969

Center for Urban Education
AN ANNOTATED BIBLIOGRAPHY OF MEASUREMENTS FOR YOUNG CHILDREN

Barbara Berger

Center for Urban Education

October 1969
INTRODUCTION

The list of entries in this bibliography represents a selected sample of assessment instruments for prekindergarten and kindergarten children. (In some instances, measures are cited which have been designed to tap a wider age range, extending through Grade III.) Some standard instruments which are commercially available have been included, but the majority are research instruments, used successfully on an experimental basis by various investigators. With few exceptions, these measurements are individually administered, since a one to one testing situation tends to be optimal for young subjects.

It should be stated that this annotated bibliography is by no means exhaustive. It represents a selected sampling of available instruments covering the major assessment areas in early childhood research, which include measurement techniques geared specifically to the urban child from a low socioeconomic background. Other pertinent considerations were the following: test validity and reliability data; childlike appeal; ease of administration; and scoring.

The entries have been organized into specific test categories indicating the type of behavior measured; in some instances, instruments have been cross-referenced if appropriate to more than one test category. Age levels are specified for each entry. Each entry has also been labeled R (research instrument) or C (indication of commercial availability) and I (individual administration) or G (group administration); these designations precede all entries. Some of the research instruments also have reference numbers which refer to relevant publications. (See Appendix to this report for the bibliography of references.) In addition to this information, entries include the names of the publishing houses,
if commercially available -- or in the case of research instruments, the names of the individuals to contact for further information about the tests. The exceptions to these are the starred entries.

Starred entries represent instruments in process of development at Educational Testing Service, Princeton, New Jersey, in conjunction with a longitudinal assessment of disadvantaged children. These measurements now in preparation, include new instruments as well as modifications of existing measures for use with young children, currently in preparation. Inquiries about these listings should be addressed to: Dr. Virginia Shipman, Developmental Research Division, Princeton, New Jersey.
CHAPTER I
COGNITIVE STATUS MEASUREMENTS

The tests listed in this chapter fall into two categories: (1) intelligence tests and (2) general measures of cognitive status. The latter differ from conventional tests of intelligence in that performance shows a substantial correlation with age, in contrast to the IQ score. They also differ in terms of their greater sensitivity to experience and training. For this reason this group of measurements is appropriate for assessing the effects of intervention as well as academic readiness in early childhood.

A. IQ MEASURES

(C - I) Stanford-Binet (Form L-M) -- (2½ years and up)

Despite the fact that this is a highly verbal intelligence test, it is still considered to be the most appropriate instrument for assessing the intellectual functioning of socioeconomically deprived children. Since it appears less biased for this population than other well-standardized tests it is preferable to other IQ tests.

(Klaus R. and Gray, S., Beller, K.)

It has been reported to yield highly significant correlations


with other cognitive variables in evaluations of Headstart populations, and to be a good predictor of academic success in the kindergarten. Hess reported a correlation of .717 with reading readiness tests and an r of .702 with teachers' ratings. For these reasons as well as its high reliability and validity, it stands out as the most dependable IQ measure. A further consideration is the fact that the Stanford-Binet to date has been the most widely used intelligence test in research with younger children -- thus it has the advantage of permitting comparisons of current research findings with the results of previous investigations.


(4-1) Wechsler Preschool and Primary Scale of Intelligence -- (4 to 6½ years)

This test is essentially an adaptation of the Wechsler Intelligence Scale for Children, and similarly provides a verbal, performance, and full scale IQ score. It contains 5 verbal tests and an alternate, and 5 performance tests. Some of the items are new although most consist of items from the WISC, which have been adapted for younger children. Standardization includes six age levels, controls for sex, color, father's occupation, geographic region and urban vs. rural residence.

(Psychological Corporation, New York City)

Goodenough Draw a Person Test (DAP) -- (older preschool, K and primary)

This is a widely used test for evaluating intellectual functioning due to ease of scoring and administration, adaptability to group testing, and lack of reliance on verbal functioning. Notwithstanding these advantages, the DAP has certain drawbacks as an IQ test. It is not equivalent to the Binet; and by contrast with it, is not adequately sensitive to the effects of schooling at the pre-school level (Beller\(^4\)). Furthermore, with a pre-school population, it correlates significantly with chronologic age as an IQ test should not, and retest correlation is not as high as would be desirable (Hess\(^5\)). In view of these limitations, it is most useful as an approximation of intellectual functioning rather than a precise measure, when conditions impose restrictions on administration and time.

At the prekindergarten level, however, drawings of the human figure appear to provide a simple and effective screening device for the assessment of intellectual functioning, when the Goodenough Draw a Person Test is scored according to the Harris system. Harris\(^6\) reports moderate correlations between performance on the drawings and other intelligence measures for five year olds. He asserts that drawings so scored indicate the child's ability to form abstract concepts; his drawing score is more highly correlated with the handling of spatial and quantitative concepts than with perceptual speed or verbal abilities.

\(^4\)Beller. op. cit.

\(^5\)Hess. op. cit.

(C-I) **Columbia Mental Maturity Scale (revised edition) -- (3 to 12 years)**

This is an untimed, non-verbal pictorial classification test, involving perceptive discrimination relating to color, shape, size, use, number, kind, missing parts and symbolic material. Although this test was originally designed to measure the intelligence of handicapped children and provides mental age quotients, it does have limitations as an IQ test. It is probably most useful as a measure of concept formation in young children, and in this connection makes a good addition to a test battery, particularly when non-verbal instruments are preferable. It has the advantage of being enjoyed by young subjects and being quick and easy to administer.

Testing time: 15 to 20 minutes.

(Harcourt, Brace and World Publishing Company, New York City)

(C-I) **Peabody Picture Vocabulary Test -- (2½ years and up)**

This is a widely used non-verbal intelligence measure, which has the advantage of being time saving, simple, and quickly and easily administered. It consists of a series of 15 test plates, each containing four pictures. E presents each set of pictures, reads the accompanying stimulus word, and asks S to respond by indicating the picture which best illustrates it. The fact that it makes no verbal demands on the child makes it particularly appropriate for testing children with speech impairment.

Testing time: 15 minutes.

(American Guidance Service Inc., Minneapolis, Minnesota; 1959)
Merrill-Palmer Scale of Mental Tests -- (2 to 6 years)

This test is structured in terms of six month periods. It consists largely of performance items, in contrast to the verbal abilities indexed by the Stanford-Binet. This feature is an advantage insofar as the Merrill-Palmer measures skills which the Stanford-Binet does not. Moreover, young children enjoy this test. Another consideration is the fact that it has been widely used in research studies, and is reported to show a substantial correlation with the Stanford Binet (r. in the 60's) for two, three and four year olds.

However, the Merrill Palmer does have certain limitations: a large proportion of timed items, method of presenting scores, and standardization on small samples at each age level.


Arthur Point Scale of Performance Tests, Revised Form II -- (5 to 15 years)

This non-language measure of intellectual ability is especially appropriate for deaf children, youngsters with severe reading disability, cases of delayed or defective speech, and the non-English speaking child. However, it is not normed for use with children below five years of age. Standardization data is based on an age range of 5 to 15 years.

(Psychological Corporation, New York City)
B GENERAL MEASURES

(C-I) Gesell Developmental Schedules - (Pre-School and K)

This instrument consists of clinical, qualitative measures of motor development, adaptive behavior, language development and personal social behavior. It affords a total developmental quotient as well as age norms for various parts of this test.

(Psychological Corporation, New York City).

(C-I) The Cooperative Pre-School Inventory (Caldwell) -- (3½ through K).

This instrument was designed for use by classroom teachers to assess the effects of Head Start training. It is particularly sensitive to experience, and can be used to measure changes associated with educational intervention. Substantial correlations with Mental Age on the Stanford Binet have been reported. Achievement is assessed with respect to knowledge of the child's personal world; ability to follow directions; knowledge of time sequences and locational associations; the characteristics of social roles and basic numerical concepts.

In addition to the commercial edition of this test, the research department of Educational Testing Service has developed a modified, shortened version. The latter includes items measuring General Knowledge; Listening (word meaning and comprehension); Writing (form copying).

Testing time: 15 to 20 minutes.

(Commercial test available from Educational Testing Service, Cooperative Test Division, Princeton, New Jersey).
(R-G) **STAR or Screening Test of Academic Readiness (Ahr)** — (Older preschool and beginning kindergarten)

Rf. 1

The STAR was designed as a preschool screening test, for the purpose of identifying children with intellectual deficits or learning disabilities. It has been used both with upper middle class and Head Start samples. This instrument includes eight subtests: Picture Vocabulary; Letters (recognition and printing); Picture Description; Copying Geometric Designs; Relationships (size, space, direction, and position of mutilated pictures); Numbers. The STAR has the advantage of systematically tapping a variety of numerical, receptive verbal, visual-motor, and relational abilities. Correlations with individual and group intelligence and achievement tests are reported to compare very favorably with the predictive coefficients of group intelligence tests at the kindergarten level. Thus it appears to be a promising instrument for estimating general level of intellectual functioning as well as specific abilities, and may prove to be more useful in this connection than established intelligence measures. This test also has the advantage of administration by the classroom teacher on a large group basis.

Testing time: Approximately one hour.

(For information, contact A. E. Ahr. See psychological directory listing.)
This test was developed as part of the Pittsburgh public schools program; it was designed as a diagnostic instrument to facilitate planning of an appropriate cognitive curriculum for each child. Items were developed on the basis of an analysis of first grade materials to determine those skills that were necessary prerequisites to the skills taught in first grade. Initial piloting and subsequent revision was based on the administration of this test to sizeable samples of four and five year olds. Factor analysis reveals ten factors to account for most of the inter-item correlations: Vocabulary, Color Naming, Classification, Functional Relationships, Visual Matching, Auditory Matching, Picture Arrangement, Symbol Series, Counting, and Verbal Concepts.

Correlation studies of performance on this test with achievement on the Metropolitan Readiness Test for five year olds show high correlation ($r = .80$), indicating this instrument to be useful in predicting first grade readiness and achievement, and to constitute a sound basis for curriculum development in early childhood.

(For information, write directly to the co-authors, Esther Kresch, Pittsburgh Public Schools or Bert Green, Jr., Carnegie Mellon University, Pittsburgh.)
(R-I)* The TAMA General Knowledge (3½ through Grade III)

The TAMA is another cognitive status measure which differs from the preceding tests in scope. It measures general knowledge only.

This test covers a range of areas such as: social and physical environment, health and safety, practical arts, consumer behavior, sports and games, art, literature, TV and comics, etc. Procedure consists of an orally presented stimulus and picture responses -- three choices. No verbal response is required.

*(This entry and all subsequent starred entries are in process of development in the research division of Educational Testing Service, Princeton, New Jersey. See introduction to this report for explanation.)
CHAPTER II
MEASUREMENTS OF COGNITIVE ABILITIES

(Concept Formation; Memory, Attention and Learning; Language)

These instruments are designed to assess specific cognitive functions. The majority are uni-dimensional measurements; however, the concluding portion of this chapter includes two multi-dimensional batteries indexing a range of cognitive as well as some perceptual functions.

A. CONCEPT FORMATION

1. Understanding of Basic Concepts (size, color, shape, quantity, space, etc.)

(R-I) Institute for Developmental Studies Early Childhood Inventories --
(preschool and K)

Rf. 12

These inventories, especially designed for and standardized on low income Negro urban children, consist of a series of measurements assessing primarily receptive understanding of a range of concepts.

Concept categories included are the following:

- Body parts
- Colors, shapes
- Classroom objects vocabulary
- Relational concepts (pre-mathematics and pre-science)
- Quantity (set and number)
- General (same-different)
- Linguistic (alphabet letters, comparatives, superlatives, prepositions)

(Experimental versions of these inventories may be obtained from Dr. Jack Victor, Institute for Developmental Studies, New York University, Washington Square Branch, New York City).
Knowledge of Relational Concepts subtest from the Experimental Early Childhood Battery\(^1\) -- (4 years through K).

This test utilizing play objects as stimulus materials assesses understanding of various relational concepts: size, length, directionality, quantity, weight, on two levels: (1) recognition of a positive instance of the concept (2) ability to produce/demonstrate the concept without benefit of examiner of stimulus cues.

Concept Familiarity Index from the Institute for Child and Experimental Education Preschool Battery\(^2\)--(2 years to 3 years, 8 mos.).

This measure parallels the preceding, but is designed for younger children. It taps understanding of a range of basic concepts including color, form, tactile, quantity, length and positional spatial concepts. Performance is assessed both in terms of recognition of an instance of the concept and ability to demonstrate the concept.

Berger's Concept Attainment Test -- (prekindergarten and K).

This test measures understanding of body concepts, directional-spatial concepts, positional concepts, serial concepts (first, last, end, beginning), speed concepts, and quantity concepts. The child's grasp of concepts is assessed solely in terms of ability to produce the concept. Part I (body concept) is presented in a Simon Says game format. Part II (directional-spatial concepts) uses the child's body as a reference point. The succeeding parts of this test employ play objects as stimulus materials. Preliminary piloting of this instrument

---

\(^1\)See description of multi-dimensional batteries in the final portion of this chapter.

\(^2\)See description of multi-dimensional batteries in the final portion of this chapter.
was conducted on day care samples of three and four year olds.

Testing time: 15 to 20 minutes.

(A copy may be examined in the library, Center for Urban Education, New York City.)

2. Piagetian Concepts

(C-1) Concept Assessment Kit: Conservation -- (4-12 thru K)

This instrument comprises three scales. Forms A and B are parallel forms which assess sophistication of conservation concepts with respect to: Two-Dimensional Space; Number; Substance; Continuous Quantity; Weight; and Discontinuous Quantity. Form C affords an indication of the child's ability to generalize conservation concepts, and includes two tasks: Area and Length. This measure is useful for longitudinal and learning studies because there are two parallel forms of the scales, which provide pre and post measurements. One of the scales also affords an indication of generalization of conservation.

This instrument has been cross-validated with middle class children, shows high levels of internal consistency, and homogeneity, and significant correlations with school grades.

Testing Time: 25 minutes for Form A or B and C.

(Educational and Industrial Testing Service, San Diego, California)

(R-I)* ETS Spatial Egocentrism -- (4-1/2 years through K)

Ref. 9, 23

This ten minute test measures capacity for taking a non-egocentric approach in solving problems. It requires S to identify objects from different points of view, using concrete stimulus objects and pictorial items. An object (toy animal or building) is placed in the middle of
the table. Another object (doll or dog) is moved to different positions around the table and S must identify (verbally and pointing to a picture) what part of the central object is seen.

(R-I)° Conception of Natural Events -- (4½ years through K).

This procedure assesses egocentrism in a different context. It consists of a modification of the Piaget technique, in which S is asked a series of general and specific questions about natural events. The selection of tasks is based on the work of Laurendeau and Pinard. ³

3. Classification and Reasoning

(R-I) Institute for Developmental Studies Concept Sorting Test -- (older preschool and K)⁴

Ref. 2

The purpose of this task is to measure the child's ability to categorize stimuli on some logical basis. The child is presented with sixteen cards in random order and instructed to sort them into piles. The cards depict vehicles, buildings, men at work, and animals. The child is then asked to explain his groupings. Both sorting and verbalization scores are obtained.

(R-I) Institute for Developmental Studies Adaptation of the Osler Concept Formation Technique -- (older preschool and K)

This test was designed to compare classification on the basis of


⁴These techniques are described in the 1965 Annual Report, of the Institute for Developmental Studies, New York University.
qualities in the stimulus field ranging from perceptual identity to communality. The child must match one of three objects depicted on a comparison card to a set of three objects on a stimulus card. For each of eight stimulus cards the child must make four matchings, two of which are possible on a perceptual basis, and two of which require identification of properties not perceivable in the field. Four levels of classification are tapped: perceptual identity -- the objects are identical; perceptual similarity -- the objects are similar, for example, two different kinds of dogs; class specificity -- two objects are related to the same class, for example, dog and cat; and class generality -- classification on this basis requires knowledge of broader class concepts, for example, boy and dog go together because they're both living things. The child makes his choices by pointing but is then asked why the objects go together. From his responses, we obtain both choice and verbalization scores.

(R-I) Grouping and Sorting Measure from the Institute for Child Development and Experimental Education Preschool Battery\(^5\) --

\(^5\) For description of this battery, see final portion of this chapter.

19
(R-I) Matrix Test (Bank Street College, Research Division) -- (4 years through Grade III).

Ref. 39

This test was designed to assess classification, sorting, and related cognitive skills associated with inferential reasoning. It is based upon a format used by Inhelder and Piaget to study classification behavior in young children, and consists mainly of newly constructed items combined with a few devised by Inhelder and Piaget. This test also resembles Raven's Progressive Matrices, but its format and content are more suited for young children -- it includes representational as well as abstract items, requires a less abstract attitude, presents items individually on separate cards rather than in a booklet.

The test is made up of 44 items falling into the following categories: Perceptual Matching; Class Membership; One-Way Classification; Two-Way Classification. Social class differences in performance are readily apparent on this test, with middle class children consistently demonstrating superior achievement as compared with socioeconomically deprived youngsters. Four year olds can understand these tasks and experience a moderate degree of success on the Matrix test. However, at this age level, experimentation has resulted in the use of a simplified version, which excludes the Two-Way Classification.

(For information, contact Dr. Herbert Zimiles, Research Division, Bank Street College, New York City)
(R-I) **ETS Logical Reasoning Tests** -- (Kindergarten, Primary)

This pencil and paper test, designed for older children, taps concept utilization in terms of the ability to group and classify objects on the basis of common properties. It consists of 20 picture items, where S is required to mark the picture he thinks does or does not belong with a particular group of pictured objects.

Testing time: 15 to 20 minutes.

(For information, contact Dr. Mitzi Tanaka, Developmental Research Division, Educational Testing Service, Princeton, New Jersey.)

(C-1) **Columbia Mental Maturity Scale** -- (4 to 12 years)

Although this is primarily an intelligence test, it is also very useful as a non-verbal measure of concept formation and has, therefore, been included in this test category. (See Chapter I, section A for description.)
B. MEMORY, ATTENTION AND LEARNING

1. Memory and Attention

(R-I) Stanford Memory Test -- (3½ through Grade III)

Ref. 14

This procedure is designed to measure both short term and long term memory (non-verbal). It consists of a series of line drawings of children, each presented twice. On the first presentation E names the picture; on the second S must produce the name. Time interval between presentation and response varies, yielding a short-term retention curve. Approximately 1/2 hour after the end of this series, S is shown an array of all the pictures and asked to find the matching picture -- this part of the test yielding a long term retention function.

Testing time: 10 minutes.

Additional measures of memory include the following subtests:

The Immediate and Delayed Memory subtests from the Experimental Early Childhood Battery 6-- (older preschool and K).
Parallel subtests from the Institute for Child Development and Experimental Education Preschool Battery 6-- (2 years through 3 years, 8 months).

2. Learning Ability

(R-I) Discrimination Learning Subtest from the Institute for Child Development and Experimental Education Preschool Battery (3 years to 3 years, 8 months).

6 See concluding portion of this chapter for a description of these batteries.
(R-I) Parallel subtest and the Single and Double Alternation Problem-solving Test from the Experimental Early Childhood Battery — (Clark preschool and K).

Both of these measurements make considerable demands on the child with respect to learning proficiency. Each assesses capacity for concept acquisition, thereby tapping basic learning ability.

(C-I) Animal House Subtest from the Wechsler Preschool and Primary Scale of Intelligence — (4 years through K).

Although this is basically a measure of sign-symbol association ability, it is also revealing of the child's learning efficiency, since an optimal score depends on learning ability. Performance on this test is moreover diagnostic of weakness with respect to: memory, attention, ability to follow directions, dexterity.

C. LANGUAGE

(R-I) * ETS Story Sequence Task, Part II — (3½ through K)

Rf. 27

This procedure measures verbal recall of story material. It provides data on structured speech patterns as well as comprehension and interpretation. Similar cartoon style picture cards are used here. E tells a story and instructs S to listen carefully because he will have to repeat the same story. As the story is being told, E places the appropriate cards in front of S. Following the oral presentation, S is asked to retell the story and his production is taped.

Testing time: 5 to 10 minutes.
(R-I) *ETS Communication Skills Test -- (kindergarten, first grade)

Rf. 27

This test measures recall, comprehension and interpretation skills. It consists of 10 items, each with 3 pictorial choices. E reads a short story, then retells a short section at a time and asks two questions per section, which S responds to by pointing to the appropriate picture. Although developed as a classroom test for first grade children, and standardized on this age group, it is also suitable for assessing story listening comprehension skills of kindergarten children.

(C-I) Illinois Test of Psycholinguistic Ability -- (pre-school, K, and primary)

Rf. 5, 28.

This instrument consists of nine subtests, each designed to measure some function contributing to language development. Critics of this test have specified the need for defining the meaning of the sub-scale scores and the use of cross-validation techniques. One of the sub-tests (the auditory-vocal automatic) has been shown to be highly sensitive to intervention and is particularly suitable for assessing the effects of training with respect to expressive language — application of knowledge of word and sentence properties.

(University of Illinois Press, Urbana, Illinois)

In addition to these structured measurements of language behavior, techniques have been developed to elicit and study the spontaneous speech of young children. The next two procedures, developed by the Institute for Developmental Studies, New York City, afford representative samples of expressive language, which provide important data for
the study of individual children with reading and language difficulties, as well as comparisons of children at different developmental levels, socioeconomic group comparisons, and studies of effects of differing intervention programs.

(R-I) Institute for Developmental Studies Telephone Interview -- (pre-school, K and primary)
Rf. 2
This technique consists of a standardized interview conducted during free play period using two telephones, with E asking six general questions and S responding in a play telephone conversation. Responses are taped and analyzed in terms of three variables: (1) degree to which child uses more than routine, unelaborated, factual responses (2) his ability to make himself understood (3) his use of complete sentences or phrases as against single-word and non-functional, fragmentary answers.

(R-I) Institute for Developmental Studies -- Adaptation of the Clown Technique (K. Salzinger) -- (preschool, K, and primary)
Rf. 2
This procedure employs an operant conditioning technique. Younger children are introduced to a papier-mache Clown, whose nose lights up when the child talks to him. (The child's speech is recorded through a microphone in the clown's neckpiece.) With older children, a rocket device has been used (a toy rocket mounted on plywood climbs closer to the moon as the child speaks). A five minute speech sample is taken for each subject and various linguistic analyses carried out, relating to the range of oral vocabulary, mean sentence length, and content (range of topics talked about).

* * *
Finally, there are two multi-dimensional batteries which complete the list of cognitive ability measurements. The following two research instruments also include some visual perceptual measures:

(R-I) An Experimental Early Childhood Battery -- (older preschool and K)\(^7\)

This battery was initially piloted with four- and five-year-old children in day nurseries and subsequently used to assess training effects of differing preschool programs with inner city children. It consists of seven subtests, each defining a task sequence representing a progression of increasing complexity:

1. **Perceptual discrimination**: This measure taps chiefly visual analytic and visual-motor integration skills. Specific tasks here include:
   - (1) Seguin Formboard
   - (2) Mannikin (from the Merrill Palmer Scale of Mental Tests)
   - (3) Bear Puzzle (original)
   - (4) Block Designs (from the Wechsler Preschool and Primary Scale of Intelligence).

2. **Delayed Memory**: This measure indexes retentive capacity with intervening time delays:
   - (1) Two minute delay
   - (2) Four minute delay
   - (3) Eight minute delay

3. **Immediate memory**: This memory test assesses immediate recognition, recall and memory for patterns. It includes measures of:
   - (1) Object recognition
   - (2) Object recall
   - (3) Memory for bead designs.

\(^7\)This battery was developed by a child assessment team at the Center for Urban Education, and used in a longitudinal comparative investigation of Montessori and conventional prekindergarten training practices, conducted by Barbara Berger. (See Appendix for reference.)
4. **Discrimination Learning:** This subtest measures learning ability with respect to the acquisition of simple and more complex discriminatory responses. It also provides a measure of learning flexibility in the inclusion of an extra-dimensional shift problem, where solution demands the ability to shift perceptual set. It consists of the following three tasks:

   (1) **Simple form discrimination** — Test materials are cards picturing a circle, square, and triangle. The correct stimulus is always the square.

   (2) **Varied form discrimination** — Test materials are the same, but with these shapes in three colors (red, yellow and blue). Correct choice is always the triangle.

   (3) **Extra-dimensional shift** — Test materials are same as in preceding task, but beginning in the reverse order. The correct choice is always yellow, regardless of the shape.

5. **Single and Double Alternation Problem-solving:** This subtest also measures learning ability, but with respect to the acquisition of complex positional concepts involving single and double alternation patterns. The model for this measure is the Jenkins-Pascale Test of Concept Formation, which demonstrates discriminatory power up and down the phylogenetic hierarchy. This particular subtest is regarded as the best single indicator of learning efficiency. It consists of two tasks:

   (1) **Single alternation.**

   (2) **Double alternation.**

6. **General Information and Comprehension:** The items here are based on comparable sections of the WISC. (General Information section also assesses child's knowledge of parts of his body.)
7. **Knowledge of Relational Concepts:** This subtest assesses the child's grasp of basic concepts (size, length, directionality, quantity, weight) on two levels: (1) recognition of a positive instance of the concept; (2) ability to produce/demonstrate the concept without benefit of examiner or stimulus cues.

Subtest correlations with Stanford-Binet scores for day care samples is moderate, ranging from an r of .37 to .67, with the exception of the two measures tapping learning ability (Discrimination Learning and the Single and Double Alternation Problem-solving). In fact, these subtests show a near zero correlation with the Stanford-Binet, indicating that they are tapping a different dimension of ability than is measured by conventional intelligence tests.

It is recommended that this battery be administered in three testing sessions, reserving the Single and Double Alternation Problem-solving for the last session, and completing the rest of the subtests in the first two sessions.

Detailed description of this battery is available from the Library, Center for Urban Education, New York City.

(R-I) Institute for Child Development and Experimental Education Preschool Battery -- (2 years through 3 years, 8 months).

This instrument is appropriate for younger children, and includes some additional measures. At present, the lower and upper age limits of the standardization data (based on a lower and middle class Negro sample) goes down as far as two years, but currently extends upward only to the 3 year, 8 month level. Additional normative data, however, is currently
being collected on older preschool populations, both Negro and Puerto Rican children. Assessment areas are the following:

1. **Language**
   - Labeling (simple object labeling task);
   - Vocabulary (a multiple choice vocabulary test, employing the pictures from the Peabody Picture Vocabulary Test).

2. **Memory**
   - Immediate and delayed recall (a delayed reaction test tapping immediate memory and recall following intervening time delays);
   - Memory for sequence (a sequence task testing ability to imitate a sequence of behaviors performed by the examiner).

3. **Persistence** (a technique for assessing persistence at a boring task).

4. **Perceptual Discrimination**
   - A series of tasks which measure ability to make simple perceptual discriminations among similar objects and carry out behavior appropriate to these discriminations: geometric form board completion; puzzle solution; block design reproductions;
   - Additional tasks at the upper age level which test form discrimination with cutouts and puzzles. A test of figure-ground differentiation.

5. **Motor development** (a battery of various tasks assessing fine motor coordination).

6. **Concept Assessment**
   - Techniques for assessing child's knowledge of basic concepts which employ concrete manipulable materials. Two procedures are used. For some items, the child is merely required to identify instances of the concept when presented with positive and negative examples. Other items require the child to demonstrate the concept with the test materials.
7. **Classification**

Grouping task (tests the ability to select objects along a given dimension);

Sorting task (tests the ability to identify relevant dimensions by indicating which one of a series of objects differs from the others).

8. **Discrimination Learning**

The technique in this case is a test of simple discrimination learning, presenting a series of progressively more difficult discrimination problems, and above the three year level, an extradimensional shift task.

(For information, contact Dr. Francis Palmer, New York State University at Stonybrook.)
CHAPTER III

MEASUREMENTS OF PERCEPTUAL SKILLS

(Visual Perception; Auditory Perception.)

A. VISUAL PERCEPTION

(R-I) * ETS Adaptation of the Johns Hopkins Perceptual Test -- (3½ years through Grade III).

Ref. 30

This is a test of form recognition where S is presented with a pictorial form stimulus and asked to identify the matching stimulus figure. Performance for lower socioeconomic class shows .657 correlation with the Columbia Mental Maturity Scale and positive but lower correlations with the Peabody Picture Vocabulary ($r=.449$) and the Draw A Person Test ($r=.356$).

Testing time: 10 to 15 minutes.

(C-I) Seguin Formboard from the Merrill Palmer Scale of Mental Tests -- (3 years through K)

This test indexes both form discrimination and eye-motor coordination. The formboard and blocks are placed in front of the child, and he is asked to "see how quickly he can put the blocks in place."

Although age norms are available for timed performance, this test can also be used as an untimed, power test, giving 1 point credit for all pieces correctly placed. The next two measurements tap integrative capacities on the perceptual level, and are effective in identifying brain damaged children.
(R-I)* ETS Adaptation of Analysis of Visually Perceived Forms (Birch Rf. 7,8 and Lefford) -- (4½ through Grade III).

This test of form analysis demands recognition of segments of abstract figures. S is presented with model card showing a whole geometric form. Test card with a segment of the form is then presented (placed above the model), and S is asked to find the identical lines on the model card and trace this segment with his finger.

Testing time: 10 minutes.

(R-I)* ETS Adaptation of Synthesis of Visually Perceived Forms (Birch Rf. 7,8 and Lefford) -- (K through Grade III).

This test consists of a series of problems presented on 6' x 8" cards. On the left is the test figure; on the right, there are four sets of lines varying in length and spatial orientation. S is asked to pick out the bundle of lines which could be used to make the test figure.

(C-I) Frostig Developmental Test of Visual Perception -- (Prekindergarten through Grade III).

This test includes five separate sections covering: Eye-motor coordination; figure-ground differentiation; shape constancy; position in space; spatial relationships. Perceptual age equivalents are provided for each part. With younger prekindergarten children individual administration is advisable; with older children the test may be given to small groups.

Testing time: approximately 30 minutes.

(Consulting Psychologists' Press, Palo Alto, California).
B. AUDITORY PERCEPTION

(R-I) **Children's Auditory Discrimination Inventory** (Stern) -- (3½ through 4½)

This is an appropriate measurement for the younger prekindergarten child. The procedure is to show two pictures which are orally labeled by examiner. S is then presented with one of the oral labels and asked to point to the pictorial match.

Testing time: 15 minutes.

(For information, contact Dr. Carolyn Stern, Dept. of Education, University of Los Angeles, Los Angeles, California).

(R-I) **ETS Adaptation of the Wepman Auditory Discrimination Test** -- (Kindergarten through Grade III).

This test may be used for kindergarten children. The procedure is for the examiner to read pairs of words to S. S must indicate whether each pair is the same or different.

**********

Tests which have been found to be diagnostically useful in identifying youngsters with visual or auditory deficits are the following measurements:

(R-I) **The Rosner–Richman Perceptual Survey** -- (Kindergarten

This test can be administered by the classroom teacher or paraprofessional in about 15 minutes. It has been successful in identifying children with perceptual-motor dysfunctions.

(For information, contact Jerome Rosner or Vivian Richman, Pittsburgh Public Schools).
(C-I) The Auditory Vocal Sequencing Test and the Visual Motor Sequencing Test from the Illinois Test of Psycholinguistic Abilities -- (Prekindergarten through Grade I).

These two subtests measuring auditory and visual memory are effective in screening young children with specific perceptual deficiencies, who subsequently have learning difficulties in school.
CHAPTER IV
READING READINESS

(R-I) Feldmann and Mahler Reading Prognosis Test (1966 revised edition) --
(last half of kindergarten, and beginning first grade).

This has proven to be an effective diagnostic test in predicting reading achievement in first grade. It has the advantage of measuring specific beginning reading skills and providing separate scores in contrast to most global measurements available; and parts of it may be used to supplement group readiness tests, when more precise diagnostic information is desired. Furthermore, the administration time is fairly short. Six subtests are included, covering three major areas: language skills (meaning vocabulary and story-telling); perceptual discrimination (auditory and visual); basic reading skills (sight vocabulary, knowledge of alphabet letters).

Testing time: 25 minutes.

(For information, contact Dr. Shirley Feldmann, New York City College, or Dr. Jack Victor, Institute of Developmental Studies, New York University, New York City).

(R-G) A Group Reading Readiness Test -- (latter half of kindergarten).

This group pencil and paper test consists of parts of various standardized reading measurements. It includes nine subtests, which primarily assess visual and auditory discrimination skills, as well as story comprehension (marking picture responses to oral questions). Standardization data on Negro and Puerto Rican kindergarten children from low-income areas are available. Where limitations of testing time and staff
do not permit a complete individual diagnostic, this group instrument
can be supplemented by additional subtests (as for example, sections of
the Feldmann-Mahler Reading Prognosis test), to obtain a more precise
picture of the child's strengths and weaknesses.

Testing time: Approximately 50 minutes. Advisable to administer in
three sessions.

(Enquiries may be directed to Dr. Miriam Goldberg, Teachers College,
Columbia University, New York City).

(C-G) Metropolitan Reading Readiness Tests — (Kindergarten and beginning
first grade).

This test is one of the most widely used readiness measures, the
choice of many school administrators as well as researchers. It has the
advantage of coverage in the numbers area as well as pre-reading skills.

It includes separate measurements of:

1. Recognition of word meaning, listening.
2. Recall, listening; comprehension, listening; interpretation, listening.
3. Item discrimination, word discrimination.
5. Number recognition, number writing; counting; number concepts;
   operation concepts; function and relation (comparison, simple ratio);
   measurement.

Time: 35 to 60 minutes. Recommended procedure is administration to small
groups.

(Harcourt-Brace and World Publishing Company, New York.)
CHAPTER V

TECHNIQUES FOR ASSESSING COGNITIVE STYLE CHARACTERISTICS

The instruments in this category provide information about process or stylistic aspects of cognitive functioning.

A. FIELD INDEPENDENCE-ANALYTIC FUNCTIONS

(R-I) Embedded Figures Test from the Cincinnati Autonomy Test

Battery -- (Banta)\(^1\) (3 and 4 year olds).

This test measures preference for a field independent, analytic style vs. a field dependent, global perceptual style, and is suitable for the young pre-kindergarten child. It consists of a series of problem solving tasks; where S is presented with a pictorial stimulus figure and must then locate it in an embedded pictorial contest (The stimuli include both social and non-social stimulus material and geometric figures). In addition to the Banta items, this test includes some of the easier figures from the Karp Embedded Figures Test for older children.\(^2\)

Testing time: 10 to 15 minutes.

(Now available from S. A. Karp, Sinai Hospital of Baltimore, Baltimore, Maryland.)

---

\(^1\)See concluding portion of this Chapter, following Section E, for a description of the Banta battery.

(R-I) The Coates Embedded Figures Test -- (3½ through K).

This procedure consists of a similar technique using different stimulus material. Substantial correlations (.70 to .86) are reported with Witkin's adult embedded figures test, as well as satisfactory reliability at the prekindergarten level (Karp and Konstadt, 1963). Testing time: 10 to 15 minutes

(For information contact Dr. Susan Coates, Downstate Medical Center, State University of New York).

(C-I) Picture Completion, Block Designs, and Object Assembly from the Wechsler Pre-School and Primary Scale of Intelligence -- (older pre-school and K)

This cluster of subtests can also be used to assess analytic functioning, since they show a high loading of the analytic factor assessed on the embedded figures tests.

(R-I) Siegel Conceptual Style Sorting Task -- (3½ through K)

This test measures cognitive style; analytic-descriptive, relational-contextual, and categorical-inferential orientations. It has been used with low income preschool children and proves to be a discriminating test for 3-year-olds. The stimulus materials consist of familiar objects which can be related in various ways. On each trial, a different

---


5Karp and Konstadt.
object is selected by E, and S is asked to select the things that go with it. S is then questioned about the reasons for his grouping, with this aspect of the test tapping verbal ability.

Testing time: 25-30 minutes.

(For information, contact Dr. Irving Siegel, Merrill Palmer School, Detroit, Michigan).

B. REFLECTIVITY -- IMPULSIVITY

(R-I) The Kagan Picture Matching Test -- (kindergarten and primary grades) Ref. 16, 17, 18, 19, 20.

Test materials consist of a series of finely detailed drawings and a standard (which may be either pictorial or a haptic wooden cut-out stimulus figure). The task demands that S select the one drawing from a series which is identical with the standard; scoring takes into consideration both response latency and errors. Only the sample items and easier pictures are appropriate for kindergarten children.

Research evidence indicates that for older children in the early elementary grades, shorter latencies are usually associated with higher scores indicating a pattern of impulsivity -- and latency and errors show stable individual differences over periods as long as a year. For this age group, performance on this test has also predicted number of errors in reading (Kagan, 1965) and errors in tests of inductive reasoning (Kagan, Pearson and Welch, 1966) as well as categorizing style (Kagan, Moss and Siegel, 1963). However, further research will be necessary to determine the degree of comparability and variation in response styles of kindergarten children.

This procedure is also based on the Kagan technique but uses different stimulus material. The child must select the one drawing that corresponds to the standard (in this case, a wooden cut-out figure).

Following presentation of sample items; six meaningful and six geometric forms are presented. Measures obtained are response latency to choice of drawing, number of errors, and time spent in exploring the cut-out form before indicating readiness to identify it.

Instructions are identical to the Kagan procedure except that the child is given only one response choice on each figure and is not told when his choice is wrong.

Testing time: 10-15 minutes.

(For information, contact Dr. William Ward, Educational Testing Service, Princeton, New Jersey).

(R-I) Reflectivity Test from the Cincinnati Autonomy Test Battery — (3 and 4 year olds).

This is a different adaptation of the Kagan technique, specifically designed for the younger preschool child. Scoring here, however, is based on correct responses only, omitting latency data (the author of this test reports that experience with three-and-four-year-olds does not show a higher probability of error with higher latencies as is indicated for older children).

Testing time: 10-15 minutes.

6 See concluding section of this chapter for a description of this battery.
(R-I) * Lewis Matching Familiar Figures Test -- (3½ through K).  
Rf. 25

This modification of the Kagan procedure utilizes four pictures and a single standard, requiring S to identify the one comparison figure identical to the standard. The content includes both meaningful and geometric stimuli. Pilot work indicates this test to be appropriate for low income children, three and four years old; shows an adequate range of mean response time and error scores, and the expected negative relationship between time and error scores.

C. MOTOR IMPULSE CONTROL

These procedures tap a different aspect of impulsivity than the measures of reflectivity-impulsivity. Actually, performance data shows low correlations between these two types of measurements.

(R-I) Motor Inhibition Test (Macoby, Dowley, Hagen, and Degerman) -- (3 and 4 year olds)  
Rf. 26

This test requires the child to perform three simple motor acts as slowly as he can (drawing a line between two points, winding a car up to the rear of a toy truck, walking a distance of six feet). S practices each act and then performs it as slowly as he can. The most widely used technique has been the draw-a-line slowly task, reported by various investigators as the best single index of the behavior being measured.

D. CREATIVITY

The following techniques consist of modifications of the Wallace and Kogan procedures specifically for younger children:

(R-I) **Test of Alternate Uses** (Ward) — (1st through K).

Rf. 35

This procedure assesses divergent thinking ability. The child is presented with different objects (newspaper, table knife, cup and coat hanger) and asked in each instance to tell all the things he can do with the object, or make with it, or ways he can play with it.

(R-I) **Naming Category Instances Test** (Ward) — (1st through K).

Rf. 36

The child must name as many objects as he can falling into common categories -- things that are round, things that have wheels, and things that are red.

Scoring procedures on each of these measurements include three types of scores:

1. Ideational fluency score -- total number of appropriate responses given.
2. A uniqueness score -- total number of responses which no other child gave.
3. Task involvement score -- total time spent with each task.

Ward reports high intercorrelations of scores for the Uses and Instances tests with kindergarten children, and a significant correlation of involvement scores with total response scores in each case — whereby results on these two measurements indicate a unitary creativity dimension at this age level.

Testing time: test of alternate uses (10 minutes); naming category instances (5 minutes).

(For information, contact Dr. William Ward, Educational Testing Device, Princeton, New Jersey).
A Non-Verbal Uses Test -- (3 1/2 through K).

This creativity procedure is more appropriate for children manifesting language difficulties, as well as younger children, who lack verbal facility. It utilizes a play situation where S is presented with different toys (one at a time) and told to play with each toy for a few minutes. The assessment here entails recording each instance of a different category of play behavior and time spent in each variety of play, as well as any spontaneous verbalization.

E. ACHIEVEMENT MOTIVATION

Gumpcookies -- (4 1/2 through K).

This instrument measures achievement motivation, and taps feelings of competency, work vs. non-work orientation, and positive attitudes toward school. The procedure presents in story format 100 items that focus on the behavior of imaginary figures called Gumpcookies. Each item includes dichotomous behavior response patterns, where the options indicate the strength of motivation to achieve. E reads the story and points to each Gumpcookie as it is described and S then identifies his own Gumpcookie (the one that behaves exactly like he does). Items are scored 1 or 0, where 1 indicates a response consistent with achievement motivation. Preliminary testing data indicates this to be a very promising instrument for four and five year olds. Testing time: 25-30 minutes.

* * * *
In addition to these measurements, indexing specific aspects of cognitive organization, the multi-dimensional pre-school battery (described next), assesses a range of cognitive characteristics, including a number of the variables measured by the preceding tests:

(R-I) The Cincinnati Autonomy Test Battery (Banta) -- (3 and 4 year olds).

This battery was specifically developed to measure various stylistic aspects of cognitive functioning -- each emphasizing a separate aspect of self-regulating behavior having relevance for autonomous problem-solving strategies. Normative data is presently being collected in preparation for standardization of this instrument. In its present form, this battery includes a series of problem-solving measurements as well as rating scales, covering the following assessment areas:

1. Curiosity and Exploratory Behavior -- Task Initiation and Curiosity Box.
2. Innovative Behavior -- The Dog and Bone Test.
6. Incidental vs. Intentional Learning -- Find the Color Green Test;

Examiner rating scales include:
Task Competence Rating Scale;
Social Competence Rating Scale; and
Kindergarten Prognosis.

Testing time: approximately 1 hour.

(Inquiries may be directed to Dr. Thomas Banta, Department of Psychology, University of Cincinnati, Cincinnati, Ohio).
CHAPTER VI
PERSONAL SOCIAL DEVELOPMENT

(R-I) Headstart Inventory of Factors Affecting Test Performance
(Zigler and Butterfield) -- (Prekindergarten and K).

This rating scale was adapted from the face sheet of the Stanford-Binet, but has the advantage of specifying the nature as well as the degree of response, which are confounded in the Binet procedure. Variables measured include the following: degree of adverse effects and factors affecting performance; response time, attention, persistence; feelings of competency and ease with examiner; dependency, activity level; enthusiasm; verbalness and spontaneous verbalization; interests; tolerance for frustration, impulsivity.

(R-I) Behavior Inventory (Hess and colleagues) -- (preschool). Ref. 13

This instrument was designed to measure certain behavioral and emotional tendencies, and was used in a Headstart evaluation project. Teachers rate children on 7 point scales. Factor analysis of the preliminary version resulted in a revised and abbreviated scale of 20 items clustering around the following factors: aggression; verbal-social participation; timidity; independence; and achievement motivation. The inventory is easy to administer and analysis is simplified by the provision of summary scores in each of these behavior areas. Despite the fact that this measurement is only minimally related to task oriented behavior, it affords a more than moderate indication of cognitive performance in assessing behaviors which are integrally related to academic success. It has proved to be a reliable measurement which also relates to other behavior instruments and cognitive measures.

(For information, contact Dr. Robert Hess, Stanford Research Institute, Menlo Park, California).
Brown Self-Concept Referents Test -- (3½ through Grade III).

This instrument measures self-concept, including perception of others (teachers, peers), and perception of self. For younger prekindergarten children, only the first part, self-perception, is recommended. But for older children, the test may be subsequently readministered, changing the referents to assess how the child thinks he is seen by significant others (teacher, peers, and/or mother).

The technique is to present the child with a full size colored photograph of himself, and then ask him to report his perceptions on the basis of a bipolar list of 14 adjectival items presented orally in an either-or format. This procedure is one of the few measures which index the child's evaluation of self as subject and object, and has reliability and validity with four year old children. It also has the virtues of being easy to administer, making only simple verbal demands and minimal demands for sustained attention, and is appealing to young children.

(Copies of this test may be obtained from Dr. Jack Victor, Institute of Developmental Studies, New York University, New York City).

Open Field Test -- (3½ years through K).

This technique consists of observations of play behavior in a standardized but relatively free play situation, where the examiner does not participate, but merely records behavior on a variety of dimensions (sex role appropriateness, social responses, complexity, approach to tasks, demonstration of manipulative and analytic skills, etc.)
(R-I)* Ability to Delay Gratification Test (Mischel Technique) -- (3½ through Grade 3).

S is shown two rewards, and is told he can have the smaller one now or the larger one at a later specified period, and asked to make a choice.

Testing time: 2 to 5 minutes.
REFERENCES


23. Laubengayer, Nancy. *The ability of the preschool child to take the point of view of another person and the effects of training on his spatial egocentrism*. Paper presented at the meeting of the Midwest Association for the Education of Young Children, Toronto, 1965.


