6-1966

A COMPARISON OF THE ILLINOIS TEST OF PSYCHOLINGUISTIC ABILITIES AND THE PARSONS LANGUAGE SAMPLE IN ASSESSING THE LINGUISTIC FUNCTIONING OF INSTITUTIONALIZED MENTALLY RETARDED CHILDREN

Ralph Donald Lee Horner
Kansas State College of Pittsburg

A COMPARISON OF THE ILLINOIS TEST OF PSYCHOLINGUISTIC ABILITIES AND THE PARSONS LANGUAGE SAMPLE IN ASSESSING THE LINGUISTIC FUNCTIONING OF INSTITUTIONALIZED MENTALLY RETARDED CHILDREN

A Thesis Submitted to the Graduate Division in Partial Fulfillment of the Requirements for the Degree of Master of Science

By

Ralph Donald Lee Horner

KANSAS STATE COLLEGE OF PITTSBURG

Pittsburg, Kansas

June, 1966

ACKNOWLEDGMENTS

I wish to thank Dr. Henry Leland whose guidance made it possible to undertake the study of this topic; his staff who helped with the scheduling; Mr. Ray Foster and Dr. Kazou Nihiri who provided advice on the analysis of the data and secured the use of the UCLA Health Sciences Computer; Dr. James McCarthy of the University of Wisconsin, Dr. Gerry Siegle of the Universi of Minnesota, and Doctors Spradlin, Smith, and Leach of the Parsons State Hospital and Training Center who offered advice during the formative stages of the study; and my wife Patsy and daughter Heili whose understanding, encouragement, and willingness to accept long hours of neglect made possible the completion of this thesis.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTINUED)

ABSTRACT

The purpose of this research was to compare the Illinois Test of Psycholinguistic Abilities and the Parsons Language Sample in assessing the linguistic functioning of institutionalized mentally retarded children. A stratified sample of mentally retarded children was administered both the ITPA and PLS in counterbalanced order. The results were correlated through the Pearson product-moment correlation coefficient and the correlation matrix analyzed through a principle-components factor analysis. The ITPA, PLS comparison produced high correlation coefficients with the exception of the Intraverbal Gesture subtest of the PLS. The factor analysis of the correlation matrix revealed one or more subtests had high loadings on five different factors. Subtests of both the ITPA and PLS had high loadings on "Immediate Recall of Auditory Symbols," "General Linguistic Ability," and "Comprehension and Expression." The other two factors, "Imitation of a Motor Act," and "The Ability to Make a Gestural Response to a Verbal stimulus," each had one subtest of the PLS highly loaded on it. The results of the factor analysis suggested with mentally retarded children Vocal and Motor Encoding tests may be more a test of comprehension than encoding ability and that "Imitation of a Motor Act," is an aspect of linguistic functioning sampled by the PLS but not the ITPA.

LIST OF TABLES

CHAPTER I

INTRODUCTION

Statement of the problem. There are currently two tests, based on different learning models, that were designed for the purpose of diagnosing linguistic deficiencies in children. The first, which appeared in 1961, is the Illinois Test of Psycholinguistic Abilities (ITPA);[1] the second, the Parsons Language Sample (PLS),[2] appeared in 1963. The ITPA and the PLS were developed by their respective authors from different learning models. The former is based on a revision of Osgood's model of psycholinguistic abilities,[3] the latter upon a model derived from Skinner's system of verbal behavior.[4] The ITPA was designed "to meet the need for a comprehensive instrument for

---

[1] J. J. McCarthy and S. A. Kirk, Illinois Test of Psycholinguistic Abilities (Urbana: Institute for Research on Exceptional Children, 1961).

[2] J. E. Spradlin, "Assessment of Speech and Language of Retarded Children: The Parsons Language Sample," J. Speech Hear Disorders, Monogr. Suppl. 10, 1963a, pp. 8-97.

[3] C. E. Osgood, "A Behavioristic Analyses," Contemporary Approaches to Cognition (Cambridge: Harvard University Press, 1957a); C. E. Osgood, "Motivational Dynamics of Language Behavior," Nebraska Symposium on Motivation (Lincoln: Universit of Nebraska Press, 1957b); and C. E. Osgood, C. J. Suci, and P. H. Tannebaum, The Measurement of Meaning (Urbana: University of Illinois Press, 1957).

[4] B. F. Skinner, Science and Human Behavior (New York: McMillan Co., 1953); and B. F. Skinner, Verbal Behavior (New Yor Appleton-Century Crofts, Inc., 1957).

the assessment of language development in exceptional children"[5] such as the mentally retarded and cerebral palsied. The PLS was designed as "a systematic attempt to develop subtests and test items which sample language behavior according to the Skinnerian system,"[6] and was empirically evaluated through its administration to 275 ambulatory mentally retarded children. There has been, to this author's knowledge, no experimental comparison of the ITPA and the PLS reported in the literature.

Need for the study. When two tests are developed that purport to measure the linguistic functioning of different types of exceptional children, those in the field of mental retardation begin to wonder how one compares with the other in assessing the linguistic functioning of mentally retarded children. In making this comparison we may ask several questions. Which subtests of the two instruments, if any, sample the same type of linguistic behavior and which are sampling behavior unique to that particular test? What are the types of linguistic behavior sampled by both tests? Does one provide more or different information than the other? Is one easier to administer than the other? The absence of answers to such questions as these emphasizes the need for this research.

The tasks posed by some subtests of the two instruments appear similar. For example, the Motor Encoding subtests of

---

[5] J. J. McCarthy and S. A. Kirk, Illinois Test of Psycholinguistic Abilities--Examiner's Manual (Urbana: Institute for Research on Exceptional Children, 1961), p. vi.

[6] Spradlin, op. cit., p. 11.

the ITPA is very similar to the Intraverbal Gestural subtest of the PLS. The former is said to measure the ability to express one's ideas in gestures[7] and the latter to measure the ability to make a gestural response to a discriminative stimulus. The ability measured by showing an object to a subject and asking him to supply the appropriate motion for manipulating it may be essentially the same ability measured by asking the subject what you do with a particular object and expecting the appropriate gesture. If these subtests and others are measuring highly related abilities this fact might be exposed by a statistical factor analysis revealing common factors at work in both instruments. If there are not common factors present the highest factor loading for each subtest will be on a differen factor and the two instruments may be considered to be measuring different linguistic abilities. An attempt could then be made to ascertain what factors are being measured by the two tests through further research. This would also help to establish which subtest provides a better understanding of the extent of development or the nature of deficits in the linguistic functioni of mentally retarded children.

Purpose of the study. The purpose of this research is to establish what factors are common to both the Illinois Test of Psycholinguistic Abilities and the Parsons Language Sample in assessing the linguistic functioning of institutionalized mentally

---

[7]McCarthy and Kirk, op. cit., p. 6.

[8]Spradlin, op. cit., p. 13.

retarded children. The purpose of determining these factors is to present a comparison of two methods of assessing language development by establishing empirically whether the two tests are measuring the same or different aspects of linguistic ability.

Additional purposes of this research are: (1) to examine the philosophy or rationale behind each test; (2) to compare the construction of the two tests with the models upon which they are based; (3) to examine the reliability and validity data presented by the author of each test; and (4) to present a comparison of the various inter-relations of the two instrument showing which subtests are highly related and those which are not

Limitations of the study. The first limitation imposed upon this study is the use of subjects who have been diagnosed as mentally retarded and admitted for training and education at the Parsons (Kansas) State Hospital and Training Center. The reasons for this limitation are first in the purpose of the research, that of comparing two tests designed to detect areas of linguistic deficiency, and second the fact that mentally retarded children usually have some degree of linguistic deficiency.[9] This study is further limited to children who are capable of making a raw score of at least one on each instrument. There are two reasons for this; first, all else

---

[9]For a thorough discussion of linguistic deficiency in the institutionalized mentally retarded see J. E. Spradlin, "Language and Communication of Mentally Defectives," Handbook of Mental Deficiency, ed. Norman R. Ellis. (New York: McGraw-Hill, Inc., 1963b), pp. 532-534.

being equal they compose the group most likely to develop improved linguistic functioning if their deficiencies are diagnosed and remedied; and second, the correlations between the two tests would be adversely influenced by including children who could not respond successfully to at least some aspect of one or both of the tests. It is necessary to insure, in this manner, that each child responds to the demands of both instruments so a meaningful comparison can be made.

The study also has a limit of fifteen years as the maximum age for a subject with the lower limit set at ten years. This is to keep the sample within the same chronological age range as the original group upon which the standard score transformations of the PLS were computed. Mentally retarded subjects within the standardization range of the ITPA (normals two to nine years) would probably respond only limitedly to the demands of the test.

The ITPA has been provided with both language age norms and standard score norms. The standard score norms provide "a more versatile means of comparing S with his own standardization group than do the language age norms."[10] The language age norms allow the assessment of the language age of a subject who is above the chronological age of the standardization group yet within the language age range of the standardization group. The upper limit of language age for the ITPA ranges from eight

---

[10] McCarthy and Kirk, op. cit., p. 96.

years, eight months to nine years, six months.[11] Language age norms also allow a comparison of the ITPA with other measures whose results are expressed in terms of age.

THE ILLINOIS TEST OF PSYCHOLINGUISTIC ABILITIES

Rationale. The Illinois Test of Psycholinguistic Abilities is based upon Osgood's model of psycholinguistic abilities[12] which is an extension of Hull's learning theory.[13] Osgood's model encompasses two stages and three levels of organization that take place between the time a stimulus acts upon a human organism and a response is emitted. The first of the two stages is decoding which Osgood defines as "the total process whereby physical energies in the environment are interpreted by an organism."[14] The second stage is encoding defined as "the total process whereby intentions of an organism are expressed and hence turned again into environmental events."[15]

The three levels of organization are projection, integration, and representation. These occur in both the decoding and encoding stages of behavior.[16] The projection system is

---

[11] Idem., p. 109.

[12] Osgood, loc. cit.

[13] C. L. Hull, Principles of Behavior (New York: Appleton Century, 1943).

[14] C. E. Osgood, (1957a), op. cit., p. 76.

[15] Ibid.

[16] Idem., p. 77.

primarily concerned with the nervous system. The "isomorphism and inability to modify through experience characteristics of the nervous system"[17] are central to the projection system. The primary function of the projection system in Osgood's model is to "relate both receptor and muscle events to the brain via 'wire-in' neural mechanism."[18]

The second level of organization is the integrational level which "organizes and sequences both incoming and outgoing neural events."[19] There are several principles underlying the functioning of the integrational level of organization. One principle is concerned with the occurrence of sensory integration.[20] According to this principle when two stimulus events occur together with high frequency the occurrence of one will tend to evoke the other. An example of this is perceiving an entire object even though we actually perceive only parts of the object, the gaps being filled in by their having been paired numerous times in past perceptions. A second principle pertaining to the integrational level concerns motor integration.[21] According to this principle when response events occur together with high frequency the occurrence of one will tend to evoke the other. An example of this is engaging in a series of

---

[17] Idem., p. 78.

[18] Idem., p. 77.

[19] Ibid.

[20] Idem., p. 81.

[21] Ibid.

responses where part of the responses are evoked by others as in such nearly automatic behavior as brushing of the teeth.

Osgood states that integration in perceptual decoding is demonstrated by the fact that "subjects will come to reflect in their predictions about successive stimuli the sequential dependencies built into the series, even though they may be unaware that these dependencies exist."[22] An example of this is an individual seeking to purchase tobacco at a grocery market or drug store rather than a dry cleaning establishment because past experience leads him to predict it is more likely to be available there.

The integration in motor encoding takes place in three steps according to Osgood. The first is "a very slow and uncertain patterning or ordering of responses on the basis of exteroceptive controls, as in imitating the seen movements of another person."[23] After this has occurred a transfer is gradually possible to "proprioceptive controls (feedback), accompanied by considerably increased speed of execution," and finally "a transfer to central programming in the integrational motor system."[24] An example of this transfer occurs in "the formation of evocative relations among motor events," such as takes place when "opening the door" where the frequency of occurrence of the various complex acts required to accomplish

---

[22] Idem., p. 83.

[23] Idem., p. 84.

[24] Ibid.

this event "provides a stability of customary action that frees it from constant voluntary supervision."[25] Another example is in the formation of "predictive motor relations" where the frequency of occurrence of the acts is lower but nevertheless somewhat predictive such as "unbuttoning one's shirt . . . predictive of peeling it off," and "lighting one's cigarette . . . predictive of blowing out the match. . ."[26]

As can be seen in the above discussion the integration level is primarily concerned with the more automatic or habitual types of behavior of the organism including the automatic aspects of perception, linking of responses, and predicting what will follow a particular stimulus event.

Osgood believes these integrative mechanisms are clearly present in language behavior especially in the grammar of language. The two quotations which follow are clearly illustrative of his position in regard to the automatic nature of language.

> When analyzed linguistically, the rules of grammar prove to be elaborate cases of redundancy or predictiveness. One such grammatical redundancy mechanism is congruence: in the present tense in English, the occurrence of a singular subject sets up a readiness for a verb ending in s (The boy runs but the boys run); a time marker sets up a readiness for the appropriate tense tag on the verb (Yesterday in the city I bought a hat); a dependent clause marker sets up a readiness for the major clause (When I come, open the door). In terms of our model, it is the frequency with which such grammatical

---

[25] Ibid.

[26] Ibid.

> redundancies have been heard and produced that sets up in the nervous system predictive integrations that match the structure of the language.[27]

> Being a relatively uninflected language, English depends heavily upon syntactical ordering mechanisms, another grammatical redundancy . . . if I say 'the happy, little _____,' all of you feel a strong tendency to fill in some noun. If I say 'the farmer killed the _____,' you have essentially two structural alternatives, a noun or a noun phrase (for example, duck or ugly duckling). If I say 'the old man eats _____,' the set of structured alternatives is larger, but still limited (a noun, dinner, meat; an adverb, swiftly, heartily; a prepositional phrase, with his hands, on the table, and so on). At each point in a language message, then, we have a hierarchy varying in its probabilistic character with the grammatical restrictions in the language as a whole. The closer the language user's nervous system can come to matching these restrictions with its own predictive integrations, the smoother become both decoding and encoding processed and the fewer decisions have to be handled by the semantic system.[28]

Osgood's theory so far has dealt with the relationship of stimuli to other stimuli and of responses to other responses but has not yet touched upon the relationship of the stimulus and the response. He deals with this in his discussion of the representational level of organization but states the relationship of stimulus to response discussed there takes place at each of the levels of organization and may be either learned or innate.[29] At the projection level are the innate unconditioned reflexes and the acquired conditioned reflexes. At the

---

[27] Idem., p. 87.

[28] Idem., p. 88f.

[29] Idem., p. 91.

integration level he feels the "complicated organization of instinctual sex behavior" may be innate as may be "the 'freezing' reaction of baby birds to certain complex retinal patterns."[30] Also at this level many "S-R relations originally organized on the 'voluntary' level will, if repeated sufficientl often, become autonomous integrations--most sensory-motor skills seem to suffer this fate, reading aloud and typing as well as tying one's shoes and brushing one's teeth."[31]

But Osgood feels that in the course of sensory events becoming associated with motor events a two-part process called a "two-stage mediation process" is necessary.[32] This concept is direct from Hull's learning theory and assumes "that in the course of associating external stimuli with overt behavior some representation of this overt behavior becomes anticipatory, producing self-stimulation that has a symbolic function."[33] Thus the representational level is concerned with the meaning or importance of linguistic symbols. Osgood makes this the basic form of learning in his model. Central to Osgood's concept is what he refers to as a "significate" that is a result of either a "wired-in" mechanism or previous learning. He defines a significate formally as "any stimulus that, in a given

---

[30] Ibid.

[31] Ibid.

[32] Ibid.

[33] Idem., p. 92.

situation, reliably elicits a predictable pattern of behavior."[34] But how does a nonsignificate stimulus become a significate stimulus? Osgood states it this way, "whenever a nonsignificate stimulus is associated with a significate, and this event is accompanied by a reinforcing state of affairs, the nonsignificate will acquire an increment of association with some fractional portion of the total behavior elicited by the significate."[35] Osgood calls this a "representational mediation process."[36] It is representational because it represents a part of the behavior originally produced only by the "significate" itself. As Osgood puts it, "this is why the (baby) bottle becomes a sign of milk-food object and not any of a thousand other things."[37] The mediation aspect arises because the behavior that is at least a partial significate produces self-stimulation that can become associated with still other responses or using Osgood's example, "sight of the (baby) bottle can thus mediately evoke 'yum-yum' noises and reaching out the arms."[38]

Osgood applies this model to the perceptual decoding process as follows:

> . . . take as illustration the object, BALL, stimulus characteristics of this object (its resilience, its shape, its weight, and so on

---

[34] Ibid.

[35] Idem., p. 92f.

[36] Idem., p. 93.

[37] Ibid.

[38] Ibid.

> . . . reliably produce certain total behavior rotary eye-movements, grasping, bouncing, squeezing, and even the pleasurable autonomic reactions associated with play-behavior) . . . according to the mediation hypothesis, the sight of the ball as a visual sensory integration, initially meaningless, will come to elicit some distinctive portion of the total behavior to the object as a representational mediation process. To the extent that this process occurs, the visual pattern becomes a perceptual sign signifying BALL object, e. g. a Unit in perceptual decoding. Long before the child begins to use language, most of the sensory signals from its familiar environment. . . have become perceptual signs of objects by virtue of association with representational portions of the same behavior the objects themselves produce.[39]

The application of the model to linguistic decoding is highly similar but varies slightly. Osgood states ". . . whereas perceptual signs bear a necessary physical relation to the objects they represent, linguistic signs bear an arbitrary relation."[40] He goes on to describe the linguistic decoding process as follows:

> Johnny is likely to hear the noise "ball," a linguistic sign, in frequent and close continuity with the visual sign of this object. . . . the linguistic sign must acquire, as its own mediation process, some part of the total behavior to the perceptual sign and/or object--presumably the mediation process already established in perceptual learning includes the most readily short-circuited components of the total behavior and hence should tend to be transferred to the linguistic sign. Thus, a socially arbitrary noise becomes associated with a representational process and acquires meaning, e. g., a unit in linguistic decoding."[41]

---

[39] Ibid.

[40] Idem., p. 94.

[41] Idem., p. 95.

To complete the presentation of the representational level we need to apply it to the other side of Osgood's two stage process, encoding. He proposes several steps in presenti a theoretical interpretation of what takes place at the representational level of organization in the encoding process. The first step occurs in the "babbling" stage where the child learns to repeat his own vocalization. This sets the stage for the second step which is the imitation of others. This second step, according to Osgood, "involves nothing more than primary generalization--the tendency to repeat a heard sound spreads from self-produced cues to other-produced cues--and upon hearing mother say 'ball' the child says 'bah,' his neares skill unit."[42] Now if we assume the stimulus object BALL has some significance for the child (through seeing it, playing with it or contact with it in some manner), hearing the label "Ball" in connection with the object's presence should produce: "(1) a single-stage association between the sight of the object and imitative labeling, and (2) a two-stage, mediated associati between sight of the object and imitative labeling . . ." Osgood considers the single stage association of little value, "sheer labeling that requires the physical presence of the object."[43] The two-stage association is quite important and to Osgood represents the formation of a unit of linguistic encoding. The association of a representational process "frees

---

[42] Ibid.

[43] Idem., p. 96.

the child's language from the immediate here-and-now--any antecedent condition--desire for the object when it is missing, for example--which elicits the critical representational process is now capable of mediating the correct, socially communicative vocalization."[44] This to Osgood "is the essence of abstraction in the use of language."[45]

Construction. Using Osgood's model of communication processe as a basis, McCarthy developed a communication model composed of three levels; Representational, Integrative and Projection, three types of input; auditory, visual, and tactile, and two types of output; vocal, and motor, combined into six channels of communication; auditory-vocal, auditory-motor, visual-vocal, visual-motor, tactual-vocal, and tactual-motor, and three processes; decoding, encoding, and association.[46]

If all these different channels, levels, and processes were combined into subtests the test would be excessively large and unmanageable. McCarthy and Kirk therefore decided to use only two of the channels of communication (auditory-vocal and visual-motor) because, as they put it, "these seemed most intimately related to the acquisition and use of language in children from two and a half to nine years of age," and to eliminate the projection level since it "deals primarily with

---

[44] Ibid.

[45] Ibid.

[46] J. J. McCarthy, "A Test for the Identification of Defects in Language Usage Among Young Cerebral Palsied Children," Cer. Palsy Rev. 21, 1960, pp. 3-5.

innate physiological processes" and "cannot be altered through learning."[47]

The original plan called for six tests at the representational level (Auditory Decoding, Visual Decoding, Auditory-Vocal Association, Visual-Motor Association, Vocal Encoding and Motor Encoding) and instead of developing decoding, association, and encoding tests for the integrative level, tests were to be designed to cover each part of the whole level. The integrative level was split because of a refinement in the theory which separated "evocative" learning (where one stimulus regularly evokes another stimulus such as takes place when reciting the alphabet) and "predictive" learning where the probability is less than 1.00 that one stimulus will evoke another stimulus (such as the word "boy" following the word "the"). The construction of the test, using "whole level" tests at each of the two levels of the integrative level rather than developing separate tests of decoding, association and encoding at these levels is a slight departure from the theory at the integrational level. Therefore Kirk and McCarthy renamed the "integrative" level the "automatic-sequential" level after the types of tests used to measure this level. The original plans for the "whole level" tests of this section included: Auditory-Vocal Automatic, Visual-Motor Automatic, Auditory-Vocal

---

[47] J. J. McCarthy and S. A. Kirk, The Construction, Standardi tion and Statistical Characteristics of the Illinois Test of Psycholinguistic Abilities (Madison: Photo Press, Inc., 1963) p. 1.

Sequential, and Visual-Motor Sequential. These whole level tests were considered instead of separate decoding, association, and encoding tests because of problems encountered when attempti to "distinguish the processes psychometrically at this level" which "would have required extensive experimentation" and would have been "prohibitive in terms of time and cost."[48] This placed the number of potential tests at ten; but, a suitable visual-motor automatic test could not be successfully developed so the number of tests was reduced to the present nine.

The tests at the representational level are classified as Decoding tests (understanding the meaning of symbols), Associati tests (relating symbols on the basis of their meaning), and Encoding tests (expression of ideas in symbols). The two Decoding tests are Auditory Decoding and Visual Decoding. Auditory Decoding as defined by McCarthy and Kirk is "the ability to understand the spoken word."[49] This is measured by having the subject "decode" sentences such as "Do balls bounce?" and "Do you rain?" If the subject does not respond or responds with an incorrect verbal or gestural answer it is assumed his failure is the result of an inability to decode.

Visual Decoding is "the ability to comprehend pictures and written words."[50] Written words could not be used for the level of child for whom this test was being designed so pictures

---

[48] Idem., p. 2.

[49] Idem., p. 7.

[50] Ibid.

are the sole stimuli. The test requires that a subject look at a picture of a single object or scene and then pick a perceptually similar object or scene from the four pictured on the next page. For example, the subject is shown a hand saw and then is required to "decode" a coping saw from the four stimulus pictures on the second page. Since responding requires only pointing, encoding ability minimally influences the response. Inability to select the correct comparison picture is assumed to be due to defective decoding ability.

The two Association tests are Auditory-Vocal Association and Visual-Motor Association. Auditory-Vocal Association as defined by McCarthy and Kirk is "the ability to relate spoken words in a meaningful way."[51] This is measured by simple analogies such as "Soup is hot, ice cream is _____." The test was constructed in such a way that children should be familiar with each word in the statement (thereby reducing the possibility of failure due to decoding deficits) and should also have the word for the correct response in their vocabulary (thereby reducing the possibility of failure due to encoding deficits). Failure on this test is assumed to result from defective association ability. The second Association test, Visual-Motor Association, is defined as "the ability to relate visual symbols in a meaningful way."[52] The test requires the subject to establish a meaningful relation between a particular stimulus object and one of four others pictured on the page

---

[51] Idem., p. 8.

[52] Idem., p. 9.

(except at the lower end of the scale where real objects are used). The test requires relating the stimuli on a "transitional basis" such as sock and shoe, or on a "substitutional basis" such as boy and girl being people.[53] The influence of decoding ability is minimized by keeping the pictures as simple and familiar as possible. Encoding ability is minimized by requiring only that the subject point to the correct object or picture. Lack of success is attributed to faulty association ability.

The two Encoding tests are Vocal Encoding and Motor Encoding. McCarthy and Kirk define Vocal Encoding as "the ability to express ideas in spoken words."[54] This test requires the presentation of a simple object to the child and requesting that he tell you all about it. The objects: ball, chalk, block, etc., were selected in the hope that most children would readily recognize them and any failure to tell about them in a number of different ways would be due to poor encoding ability. Motor Encoding, the second of the Encoding tests, is defined as "the ability to express one's ideas in meaningful gestures."[55] This test requires showing a picture of an object (except at the lower end of the test where actual objects are shown) and asking the subject to "Show me what you should do with this." The subject is expected to supply an appropriate gesture indicating intactness of motor encoding ability. The possibility

---

[53] Ibid.

[54] Ibid.

[55] Idem., p. 10.

of failure due to lack of familiarity with the object (poor decoding ability) was again minimized by employing as simple objects as possible.

The remaining three tests measure psycholinguistic abilities at the Automatic-Sequential level of organization. They are classified as either automatic (where the language has been used so frequently its structure becomes largely free of conscious effort) or sequential which is defined by McCarthy and Kirk as "the ability to correctly reproduce a sequency of symbols" and further states, "it is largely dependent upon visual and/or auditory memory."[56] The first of the tests at the automatic-sequential level deals with automatic linguistic ability and is termed the Auditory-Vocal-Automatic test. It measures the ability "to predict future linguistic events from past experience."[57] This ability is assessed by having the subject complete such sentences as "Here is a man. Here are two _____." and "Father is hanging a picture. Now the picture has been _____." accompanied by the appropriate pictures. The pictures are supplied for support to younger subjects and are not intended to supply information. To be truly Auditory-Vocal the visual component should not be present and research will be necessary to determine the significance of this difference. As mentioned above attempts to develop a Visual-Motor-Automatic test met with failure.

---

[56] Idem., p. 12.

[57] Idem., p. 11.

The last two tests of the ITPA are Sequencing tests. The two are Auditory-Vocal Sequencing and Visual-Motor Sequencing. The Auditory-Vocal Sequencing test is defined by McCarthy and Kirk as "the ability to correctly repeat a sequence of symbols." It requires the subject to listen to and repeat a series of digits and is basically "a test of immediate auditory recall."[58] The Visual-Motor Sequencing test measures "the ability to reproduce a sequence of visual stimuli from memory."[59] This ability is measured in the following manner: (1) the examiner arranges a series of small square chips (with pictures or geometric forms) in a certain order, (2) the subject is allowed to view these for five seconds, (3) they are then mixed up, and (4) the subject is required to reproduce the order from memory. This procedure is repeated for various arrangements starting with a few of the pictured chips and advancing to orders using several geometric forms.

Standardization. The present experimental edition of the ITPA was standardized on 700 children between the ages of two years, six months and nine years, zero months.[60] School age subjects were randomly selected from a list of all school children in the Decatur, Illinois, school system. The sample of preschool children were largely the siblings of the school age subjects. The city of Decatur was chosen because its

---

[58] Idem., p. 12.

[59] Ibid.

[60] Idem., p. 14.

distribution of social classes approximated that of the state of Illinois. The IQ range of the standardization group was 80 to 120 as measured by the 1937 revision of the Stanford-Binet. Children with sensory or physical defects were excluded as were those of the Negro race. Fourteen age groups were used ranging from two years, six months to nine years, zero months. The same number of males and females were included in the study. The proportion of various socioeconomic classes in the sample approximated that of the population from which the sample was drawn. Subjects from homes where a language other than English was commonly used were excluded. Testing was divided into two sessions with a day interval to avoid fatigue. The first day the Stanford-Binet was administered with the ITPA being administered at the second session. The examiners were mainly pre-doctoral students in psychology and special education with courses in individual mental testing and with additional testing experience beyond this. Considerable effort was expended to insure uniformity of administration and close supervision provided a basis by which problems of administration could be dealt with immediately.

Reliability. Reliability of the ITPA was assessed by determining internal consistency and test-retest stability. The authors found the "overall internal consistency estimates are comparable with other instruments of this type."[61] Test-retest data was gathered on sixty-nine children in the six

---

[61] Idem., p. 28.

years, zero months and six years, six months age groups with a test-retest interval of three months or more. The stability coefficients ranged from a low of .18 for Visual-Motor Sequencing to .86 for Auditory Vocal Sequencing. The ITPA total score showed a stability coefficient of .70. A full age range estimate formula was used to transfer this into an estimated stability coefficient of .97 for the entire age range. Overall split-half reliability coefficients were quite high ranging from .90 (Motor Encoding) to .99 (ITPA total). In comparing the various reliability data the authors conclude that the selection of the six year, six months age level was a poor choice for a test-retest stability measure probably due to the fact it is a period of rapid change with entrance into school at this time contributing to further change.

Validity. The total raw score of the ITPA was compared with mental age and social class in an attempt to establish a form of concurrent validity. These variables have been found to be related to linguistic ability. The data suggest that "ITPA scores depend substantially, but not entirely upon mental age," and that "the influence of social class is less marked" but added "the social class index was relatively crude."[62] Additional validity data has appeared in subsequent publications.[63]

---

[62] Idem., p. 36.

[63] J. J. McCarthy and J. L. Olson, Validity Studies on the Illinois Test of Psycholinguistic Abilities (Madison: Photo Press, Inc., 1964); and J. J. McCarthy, "Notes on the Validity of the ITPA," Mental Retardation, 1965, Vol. 3, No. 2, pp. 25-26.

Concurrent as well as predictive validity data was obtained by giving the ITPA and eight criterion measures to a group of eighty-six children specially selected to compare with the overall characteristics of the standardization subjects. The eight criterion measures were given a second time after a three month period to provide a means of assessing predictive validity. The eight criterion tests were the "reading and spelling section of the Stanford Achievement Battery, the reading section of the Durrell-Sullivan Capacity Test, the Raven's Progressive Matrices, the Goodenough Draw-a-Man test, the Peabody Picture Vocabulary Test, and the mean length of response and sentence complexity scores derived from a sample of the subject's speech."[64] The results suggest the whole test as well as the Visual Decoding, Visual-Motor Association and Auditory-Vocal Sequential have an "excellent" degree of concurrent and predictive validity. Somewhat below "excellent" in concurrent and predictive validity were Auditory Decoding, Auditory-Vocal Association and Visual-Motor Sequential. Vocal Encoding demonstrated "questionable" concurrent and predictive validity with Motor Encoding and Auditory Vocal Automatic showing "doubtful" concurrent and predictive validity.[65] Content validity was demonstrated by showing statistically that the subtests are "internally consistent and fairly heterogeneous with respect to one another" and that through the use of the standard error

---

[64] J. J. McCarthy, (1965), op. cit., p. 25.

[65] Ibid.

ranges in the interpretation of scores one can further insure a "statistical content universality."[66] Further evidence of content validity is the statistical revelation that all but the Automatic-Sequential level tests are "basically 'single-ability' in character," which is the way the tests were designed.[67] The authors point out that several abilities pertinent to language are not assessed by the ITPA because of inability to design adequate scales and for economy in the length of the test which imposes limits on the validity of the instrument's content.

Construct validity was assessed by determining what factors have an influence on the test and which do not. Social class, sex, mental age, birth order and number of siblings have been demonstrated in the past to have an effect on language ability. The effects of mental age and social class have been discussed above but it should be added that McCarthy and Kirk found the mental age of those in the higher social classes to be above that of those of the same age in the lower class thus confounding the determination of the exact influence of social class on ITPA test scores.[68] Additional findings are: (1) the ITPA is not influenced by the sex of the subject, (2) other things equal first born and single children in a family tend to score higher

---

[66] Ibid.

[67] Ibid.

[68] McCarthy and Kirk, (1963), op. cit., p. 57.

than other children, and (3) again with other things equal the fewer the children in a family the greater the tendency for these children's scores each to be higher than would be those in a more populous family.[69]

A fourth type of validity study concerns diagnostic validity.[70] This was assessed by comparing ITPA scores with teachers rankings of linguistic ability and by having "experts" on the ITPA sort ITPA profiles into various disabilities which were then compared to the clinical diagnoses. Teachers rankings failed to correlate significantly with ITPA performance. A closer look at the ITPA test data showed the range of talent was severely restricted in most instances. The greater the range of talent a particular teacher had to rate the closer their ratings came to ITPA scores. The study will have to be repeated with a wider range of linguistic talent to be rated and a greater number of subjects in each sample. The "experts" were significantly beyond chance in their ability to classify subjects according to type of linguistic deficiency. This indicates the potential value of the test as a diagnostic tool for those who are skilled in its use.

The authors conclude that "generally, the data suggest the concurrent, construct, and predictive validities to be adequate,

---

[69] McCarthy and Olson, op. cit., p. 38.

[70] Idem., p. 52.

followed by the content and diagnostic."[71] They point out a few cautions concerning the use of the test: (1) Motor Encoding, Vocal Encoding and Auditory-Vocal Automatic may be measuring something other than that defined by the Examiner's Manual, (2) of the three Vocal Encoding appears the most valid, and (3) "In the diagnosis of children with linguistic defects, particularly dyslexia, it is recommended that auxiliary tests accompany the use of the ITPA."[72] Aside from the cautions just mentioned the ITPA appears to be a valid measure of several aspects of linguistic functioning and will likely see increasing importance and use as a means of measuring linguistic functioning.

THE PARSONS LANGUAGE SAMPLE

Rationale. The Parsons Language Sample is based upon Skinner's system of Verbal Behavior.[73] This system places the emphasis upon observable causes for behavior which can be described objectively. The system includes no unobservable hypothetical constructs or other theoretical propositions but relies entirely upon simplicity and objectivity in explaining behavior. Skinner emphasizes the relation of environmental conditions to verbal behavior rather than the meaning of the verbal behavior stating ". . . meaning is not a property of

---

[71] Idem., p. 66.

[72] Idem., p. 67.

[73] Skinner, (1957), loc. cit.

behavior as such but of the conditions under which behavior occurs,"[74] thus the behavior can be studied apart from a constr such as meaning. Skinner includes both vocal and non-vocal classes of responses stating that by "defining verbal behavior as behavior reinforced through the mediation of other persons we do not, and cannot, specify any one form, mode, or medium."[75] Thus gestures, sign language, writing as well as speaking are all verbal behavior. The Parsons Language Sample samples two classes of verbal behavior, vocal and non-vocal.

Skinner's system also includes various conditions under which vocal and nonvocal verbal behavior may occur. The first is mand behavior defined formally as "a verbal operant in which the response is reinforced by a characteristic consequence and is therefore under the functional control of relevant conditio of deprivation or aversive stimulation."[76] Words describing behavior that include mand as part of the word are examples of this type of verbal behavior, (e. g., demand, command, countermand) as is the related behavior of requesting and asking. The mand often specifies the action or the reinforcem to be provided by the listener or perhaps both. A person may mand that another stop, look, or listen thus specifying the action of the individual or he may mand both the action and the reinforcement such as pass the bread or give me a drink.

---

[74] Idem., p. 13.

[75] Idem., p. 14.

[76] Idem., p. 35f.

A second condition is Verbal Behavior under the Control of Verbal Stimuli[77] which means that some aspects of verbal behavior come about as the result of the verbal behavior of other people. One example of this is the echoic response. An echoic response is repeating a response made by another person. This type of verbal behavior is especially prevalent during the early years of an individual's life. When a child is learning his first words an adult will attempt to get the child to repeat as many words as he can. The ability to do so reinforces the echoic response and helps to establish words in the child's repertoire.

Adults reinforce echoic behavior when they repeat directions or some other response produced by another person since it often facilitates their ability to carry them out, etc. Another form of echoic behavior is a person echoing himself. A person may repeat verbally something he has said in order to strengthen the self stimulation he receives and reinforce his own verbal behavior. Such behavior is often used when attempting to memorize.

Another example of verbal behavior under the control of verbal stimuli is textual behavior which is vocal behavior under the control of a text in some form. Response to written text, pictured text, and symbols are all examples of textual behavior. Skinner defines it as "a vocal response . . . under

---

[77] Idem., p. 52.

the control of a nonauditory verbal stimulus."[78] Reinforcement of textual behavior occurs in an educational setting when children are praised for responding with the correct vocal response upon viewing printed or written words (reading) and upon viewing a printed or drawn picture (visual recognition). Textual behavior is reinforced in adult life when a person's reading or visual recognition leads to some favorable circumstance. Reinforcement may occur so frequently, that at the extreme people may attend nearly every picture they see and read everything from books to match covers. As in echoic behavior where a person responds to his own verbal behavior a person can also respond to a text he has produced himself. An example would be making a note to "remind ourselves to do something or help ourselves to say something . . ." Skinner also states "the relatively permanent nature of a text, as compared with the echoic stimulus makes self-textual behavior ordinarily more important than self-echoic."[79]

Nonvocal verbal behavior is also a part of Skinner's system and the first example of this in his system is transcription. This behavior, when it consists of a writing response to a written stimulus, is similar to echoic behavior except the sense modality is visual rather than auditory. Another type of transcription behavior is a writing response to a vocal stimulus

---

[78] Idem., p. 66.

[79] Idem., p. 69.

such as occurs when taking dictation. Both types of transcription behavior are greatly reinforced in everyday life.

In both the echoic and transcription types of verbal behavior, the correspondence between the stimulus and the response is identical. This is not the case in many instances of verbal behavior. For example, to the stimulus "What is two plus two?" the correct response is four, to the stimulus "What is the capitol of Kansas?" the correct response is Topeka. Skinner calls this type of behavior intraverbal. Intraverbal behavior plays a large part in such things as general conversation, and learning how to count and compute. Intraverbal behavior is often dependent upon more than just the first preceding stimulus. Examples of this are seen when we try to recall from memory part of a poem or musical composition and must make a running start before it can be recalled. Intraverbal behavior is also a prominent part of word association and translation of one language to another.

A third type of verbal behavior under the control of verbal stimuli is comprehension. Spradlin states, "a person comprehends what is being said to him when the person makes a differential response which bears some relation to a specific stimulus situation."[80] A person demonstrates comprehension when he responds to a command such as "Come here," or "Close the door." A person may comprehend without demonstrating it

---

[80]Spradlin, (1963a) op. cit., p. 11.

immediately, such a response would occur to "Go to the store tomorrow." The comprehension may be auditory, visual, or tactile depending upon the type of stimulation to be comprehended.

Another important part of Skinner's system is the verbal behavior termed "tact." It is defined as "a verbal operant in which a response of given form is evoked (or at least strengthened) by a particular object or event or property of an object or event."[81] An example is a child knowing the name of some object because of the many reinforcements that have occurred in the past when he gave the name and was praised or acknowledged in some manner.

Construction. The Parsons Language Sample in its original form had seven subtests containing a total of 123 test items.[82] These subtests are Tact, Echoic and Intraverbal on the vocal section of the test and Echoic Gesture, Comprehension and Intraverbal Gesture on the non-vocal section of the test. The seventh subtest, Mand, sampled both vocal and nonvocal responses.

The first subtest of the vocal section is the Tact subtest which consists of twenty-eight items the subject is asked to identify. There are seven real objects ranging from a ball to a C-clamp, seven miniature objects ranging from a duck to a bottle brush, seven colored pictures ranging from a motherly type woman to an anchor, and seven noncolored pictures ranging from a fatherly type man to a propeller. The examiner presents

---

[81] Skinner, (1957), op. cit., p. 81f.

[82] Spradlin, (1963a) loc. cit.

each item and says "What is it?" or "What do you call it?" A ceiling is five consecutive errors. The raw score is obtained by summing the correct answers. The Tact subtest comes directly from Skinner's system and samples the child's discriminative naming of objects and pictures. There may also be nonvocal tacts but the PLS does not sample this type of response.

The second subtest of the vocal section is the Echoic subtest. This subtest requires repeating an auditory stimulus presented by the examiner. Section A requires repeating ten words or sentences ranging from "ball" to "In the summer time the little children like to eat black walnut ice cream." Section B requires repeating one or a series of numbers ranging from "2" to "1-5-8-9-3-7." The word or sentences and number or numbers are said once by the examiner and must be repeated exactly as presented. A ceiling consists of three consecutive errors for Section A and three consecutive errors for Section B. The Echoic subtest is also direct from Skinner's system and samples a child's ability to repeat words and numbers and establish them in his repertoire through the reinforcement he receives from this repetition.

The third subtest of the vocal section is the Intraverbal subtest. This subtest requires responding to a discriminative stimulus by answering questions, supplying a last word, making simple analogies, and determining similarities. The items range from the simple question "What do we do when we are hungry?" to the more difficult similarity item "In what way is an egg and a seed alike?" The Intraverbal subtest is also

directly related to Skinner's system and samples a child's ability to respond to a stimulus that provides cues to the response rather than being identical to the response.

The non-vocal section of the PLS samples only one major classification of Skinner's system as compared to the two sampled by the vocal section. The three subtests of the non-vocal section: Echoic Gesture, Comprehension and Intraverbal Gesture, would all be classified as verbal behavior under the control of verbal stimuli (again remember that verbal behavior does not have to be vocal). The mand classification was represented by a subtest that sampled either vocal or nonvocal behavior but was abandoned after the empirical evaluation proved it highly unreliable.

The fourth subtest of the PLS is the Echoic Gesture subtest of the nonvocal section. This subtest requires imitation of the examiner's motor response. Where the Echoic subtest of the vocal section requires repeating an auditory stimulus presented by the examiner, the Echoic Gesture subtest of the non-vocal section requires imitating a visual stimulus presented by the examiner. This subtest also consists of two sections. Section A requires a subject to imitate eight motor acts ranging from pointing to a light to tapping a block with a finger. Section B requires the imitation of five patterns of tapping blocks in a left-right sequence. This subtest samples a child's ability to mimic the actions of others and establish them in his repertoire through the reinforcement he receives from this mimicry.

The fifth subtest of the PLS is the Comprehension subtest of the non-vocal section. This subtest requires following directions that are given vocally, by gestures, or both. The items range from calling the child's name to see if the child looks toward the examiner to "Look at the light, open the door, and put the cup to the right of the purse." This subtest samples the child's ability to make a differential response to verbal cues which bear a relation to the stimulus situation.

The sixth subtest of the PLS is the Intraverbal Gesture subtest of the non-vocal section. This subtest requires making a gestural response to an auditory discriminative stimulus. The items range from "Where is the light?" to "What do you do with a handkerchief?" This subtest samples the extent a child makes use of geatural behavior when responding to a stimulus that could elicit either a vocal or gestural response. Spradlin states that "Initially the aim of the intraverbal gesture was to measure both vocal and gestural behavior. However, during the initial testing one examiner focused on vocal responses and failed to record gestural responses. At this point examiner's were instructed to score only the gestural responses."[83]

Several types of verbal behavior included in Skinner's system are not included in the Parsons Language Sample. Many of these, such as self-echoic, textual, self-textual and transcription either do not readily lend themselves to inclusion in a subtest or are more complicated verbal behavior than is often present in the mentally retarded.

---

[83] Idem., p. 14.

Standardization. The Parsons Language Sample was evaluated empirically by several methods. The first was the development of standard score transformations from the results of administration to 275 mentally retarded children.[84] They were all ambulatory and ranged in age from seven years, eleven months to fifteen years, eight months. They were all patients at the Parsons (Kansas) State Hospital and Training Center. The four examiners, two junior college students and two housewives without previous psychometric training, were thoroughly trained in the proper techniques of administration and recording.

Reliability. Examiner equivalence was tested by comparing the results of a sample of each examiner's administrations. The size of the samples ranged from fifty-eight for examiners two and four, to seventy-six for examiner one and eighty-three for examiner three.[85] The Kolmogorov-Smirnov two sample tests revealed that only the vocal and nonvocal parts of the Mand subtests had significant differences between examiners.[86] The Harley Test revealed variance between examiners was within chance expectations with the exception of the Mand subtest.[87] A rank-order correlation between examiners in terms of the percentages of subjects responding appropriately to the different items in each subtest revealed very high correlations with the

---

[84] Idem., p. 15.

[85] Idem., p. 16.

[86] Idem., p. 17.

[87] Idem., p. 17f.

exception of the Intraverbal Gesture subtest (the Mand subtest was not included in this analysis due to the small number of items). Spradlin states that "evidently they (the examiners) are eliciting gestures differently on the various items or are scoring them differently when they occur."[88]

Another method used to empirically evaluate the PLS was the calculation of split-half and test-retest reliability coefficients.[89] The odd-even reliability coefficients were quite high for the subtests of the vocal section (.95 to .98) with the exception of the vocal Mand items (.40). The coefficients were not as high for the subtests of the nonvocal section (.91 to .93) but were still above the traditional .90 with the exception of the nonvocal Mand items (.29). The low split-half reliabilities and examiner equivalence on the vocal and nonvocal parts of the Mand subtest was sufficient justification for the author of the PLS to eliminate the Mand subtest from the rest of the empirical evaluation.

The stability of the PLS over time was measured by the test-retest method. One of the original examiners randomly selected twenty subjects from among the ones the other three examiners had tested and twenty she had tested to arrive at a sample of forty from the original 275 subjects. The time between the test and retest ranged from two to five months. Eleven of the forty had to be eliminated as untestable. Only

---

[88] Idem., p. 18.

[89] Idem., p. 19.

those with scores above zero on each subtest at one or both administrations were retained in the sample. The retest scores were higher than those in the first testing but were only significantly different for the echoic subtest. This was true when the examiner was the same one who administered both tests and for the Echoic Gesture score and total score when the test-retest examiner differed. It is possible that language ability significantly improved over the two to five month period in the areas measured by the PLS that were significantly higher. It is also possible that the Echoic, Echoic Gesture and total score are somewhat unreliable. A third possibility is that memory of the nature of the task facilitated performance in the retest situation. As Spradlin states "additional data are needed before further interpretations can be made."[90]

In a similar test of reliability thirty-two children from the original sample were retested after twenty-nine weeks. The vocal section showed a significant increase at the .05 level but the test-retest reliability coefficient was a respectable .93.[91] A test-retest situation was possible on twenty-seven patients ages six through fifteen who were first tested two to sixty-two days after arrival at the hospital. The second test was administered by the same examiner approximately fifteen months later. The test-retest correlations were low, ranging from a high of .74 to a low of .15. Spradlin

---

[90] Idem., p. 21.

[91] Idem., p. 22.

states, "In summary, six subtests exhibit substantial score stability both in terms of within-administration stability (correlation of half scores) and in terms of subsequent administration by either identical or different examiners. As was found for the examiner equivalence data, the two mand subtests appeared distinctly less stable than did the remaining six subtests. The lower retest correlations involving newly admitted patients retested more than a year later suggest more hesitant predictive statements regarding such subjects."[92]

The above findings served as the basis for eliminating the unreliable mand subtest from the PLS. Two investigators in the Department of Speech Pathology of the Parsons hospital provided additional modifications based upon their experience with the test and suggestions from others experienced in the use of the test.[93] The modifications can be summarized as follows: (1) omitting the stimulus "What is it?" or "What do you call it?" in the Tact subtest if the child identifies the items spontaneously, (2) crediting an echoic response preceded by "say" as when the examiner says "say ball" and the child responds "say ball," (3) crediting an echoic gesture response that is a

---

[92] Idem., p. 23.

[93] Sandra K. Edson and J. C. Rolland, "Rationale for Modification of Instructions on PLS and Standardization, Administration and Scoring of the PLS," (Parsons Demonstration Project Report No. 14A and 14B, September, 1964a, Parsons State Hospital, Parsons, Kansas) (Mimeographed); and S. K. Edson and J. C. Rolland, "Revision of the Parsons Language Sample, September 1964," (Parsons Demonstration Project Report No. 15, September, 1964b, Parsons State Hospital, Parsons, Kansas) (Mimeographed).

mirror image of the response made by the examiner rather than require an identical response on Section A but requiring an identical response on the block tapping of Section B, (4) allowing a second repetition of the comprehension items if the child fails to respond the first time, and (5) eliminating the possibility of an incorrect intraverbal gestural response being scored positively by specifying that not only must the child use gestures but must use correct gestures.

The original test plus these modifications and elimination of the Mand subtest composes the current revised Parsons Language Sample.

Validity. Attempts to validate the Parsons Language Sample have centered around comparisons with non-language test behavior and with the ranking of language behavior by psychiatric aides. A type of validation study concerned with the relationship of the PLS and Intelligence was completed by Horowitz. She felt that if the two correlate highly, they may sample the same type of behavior and if this is true the necessity of having the PLS would be open to serious question since "there is more known about the standard I. Q. score."[94] At the time of this study 157 children had received either the Wechsler-Bellvue, Wechsler Adult Intelligence Scale or Wechsler Intelligence Scale for Children and the Parsons Language Sample. There is no mention of who administered the tests, nor the conditions

---

[94]Frances D. Horowitz, "The Relationship Between the Parsons Language Sample Scores and Wechsler Intelligence Quotients of Mentally Retarded Subjects" J. Genet. Psychol. (In Press), p. 2.

under which they were administered, nor the time interval between the administrations of the PLS and the intelligence test. Those who had received both tests were divided into two groups. Those between eight and thirteen (called the younger group) and those between fourteen and sixteen (called the older group). A random sample of forty subjects was drawn from each of these groups. Pearson correlation coefficients were computed between (1) Wechsler composite scores (W-B, WISC and WAIS) and the means of the Parsons Language Sample's six subtests, (2) Wechsler Verbal Scales and the Vocal section of the Parsons Language Sample, and (3) Wechsler Performance scales and the Nonvocal section of the Parsons Language Sample. The correlations of the younger group for the above were (1) .48, (2) .70, and (3) -.18. The older group had correlations of (1) .61, (2) .83, and (3) -.03. For younger and older subjects combined the correlations for the above were (1) .55, (2) .75, and (3) -.10. Thus significant correlations exist between the Wechsler total score and Parsons Language Sample Vocal Section. The correlations between the Wechsler Performance Scale and the Parsons Language Sample Nonvocal section were nonsignificant. This led the author to conclude that "language as a concept independent of intelligence is best defined by non-vocal communicatory behaviors."[95] Even though several correlations were significant the author states "their absolute value is such that it cannot be said that the tests are entirely overlapping."[96]

---

[95] Idem., p. 4.

[96] Idem., p. 3.

She felt that for the population under study, the Parsons Language Sample and the Wechsler Scales "are, in part, independent measures"[97] and that the nonsignificant correlation between the Nonvocal section of the Parsons Language Sample and the other measures indicates it "might prove useful as a measure of language . . . relatively independent of the measure of intelligence."[98]

A second validation study compared the Parsons Language Sample with aide rankings of a child's language behavior in his day to day cottage life.[99] Comparisons were made between a child's PLS vocal score and his rank in speech communication and between his PLS nonvocal score and rank in nonvocal communication. One hundred ten children from five cottages were compared on PLS vocal score and aide ranks. Five different correlations (one for each cottage) were computed. They ranged from .33 to .86 with a median correlation of .64. These correlations are of sufficient strength to show the vocal section of the PLS can with some validity be predictive of a mentally retarded child's vocal communication in a cottage setting. One hundred eleven children were compared on PLS nonvocal score and aide ranks. The five correlations on this measure ranged from .18 to .80 with a median correlation of .40. This wide a range indicates either the PLS nonvocal scale does not

---

[97] Idem., p. 4.

[98] Idem., p. 5.

[99] Spradlin, (1963a) op. cit., p. 22.

systematically predict nonvocal communication or the aide's of various cottages varied in their criteria for assigning ranks. The correlations between the PLS vocal scores and nonvocal communication were similar to the correlations between nonvocal scores and nonvocal communication. They ranged from .13 to .52 with a median correlation of .23 thus showing the vocal scale of the PLS is nearly as good a predictor of nonvocal behavior as is the nonvocal scale.

CHAPTER II

REVIEW OF RELATED RESEARCH

Previous Measures of Language Ability. Meacham was one of the first of present day investigators to systematically explore linguistic functioning.[1] As part of his doctoral dissertation he developed a scale to measure several aspects of speech in mentally defective children. The aspects of speech measured by this scale are articulation, auditory discrimination, auditory digit memory span, and oral language development. The articulation test consists of ten monosyllables each with a consonant-vowel-consonant combination. These were recorded and replayed through earphones to the child who is asked to repeat what he hears. The auditory discrimination test in its final form consists of eighty-six items each requiring a word-picture matching response by the subject. Each item consists of three pictures mounted on a five by eight card and three recorded words presented through the loud speaker of the recorder. One word identifies a picture on the card and the other two sound like the names of the other two pictures. A typical item includes pictures of feet, the sun, and money accompanied by the auditory stimuli of seat, fun, and money: the subject being instructed to point to the picture named in the series of words. The auditory digit memory span test consists of sixteen recorded

---

[1] M. J. Meacham, "The Development and Application of Procedures for Measuring Speech Improvement in Mentally Defective Children," Amer. J. Ment. Defic., 1955, 60, pp. 301-306.

sequences of digits (four sequences of two digits, four of three, four of four, and four of five) presented to the subject with a one second interval between each digit and a ten second interval between sequences. The one syllable digits were select randomly with the exceptions that a number could appear only once in each sequence and no consecutive numbers could appear together. The test of oral language development used the ten pictures of the Children's Apperception Test with questions designed to encourage verbal response to the cards. The examine was allowed to encourage oral expression through the use of such questions as : "(a) What do you see on this picture? (b) Will you tell me a story about this picture? (c) What's happening? (d) What else? (e) Can you tell me more about that?"[2] In administering the tests to a sample of mentally defective children, test-retest and split-half reliability coefficients showed the tests of articulation, auditory discrimination, and auditory memory span were found to be reliable measures of these aspects of the speech of mentally defective children. The tests of auditory discrimination and auditory memory span were found to be highly related and most likely do not measure independent areas. Eight weeks of speech therapy brought significant improvement in articulation, auditory discrimination, auditory memory span, and the average sentence length section of the oral language development scale. There was no significant improvemen in control subjects with the exception of auditory memory span

---

[2] Idem., p. 303.

suggesting some form of learning or adaptation occuring in this ability. The author also points out that in the initial scores (prior to speech therapy) of the experimental and control groups there was "noticeable difference."[3] He accounts for this by stating the experimental group was purposely composed of children who needed speech therapy and that control subjects were not selected on the basis of their speech needs. This appears to be an obvious confounder within the study in that variation between groups cannot be empirically related to a specific variable. Absence of overall improvement in the control subjects may have been the result of their being systematically different rather than their lack of speech therapy. This also casts doubt on the use of Meacham's scale without further researc It has yet to be shown that improvement of score is different when one of two like groups receives speech therapy and the other does not. An increase in score may be a function of time or familiarity with the test itself as shown in the improvement of auditory memory span in subjects who received no speech therapy.

In 1958 Dunn published his well-known Peabody Picture Vocabulary Test (PPVT).[4] The test consists of a series of 150 plates each with four pictures. A stimulus word is introduced at varying levels of difficulty depending on the chronological

---

[3] Idem., p. 305.

[4] M. Dunn, Expanded Manual-Peabody Picture Vocabulary Test (Minneapolis: American Guidance Services, Inc., 1965).

or mental age of the child. The task consists of listening to a stimulus word presented by the examiner and in some way indicating which picture of the four corresponds to that particular word. Thus the PPVT is generally considered an auditory word comprehension test. Dunn states the PPVT "is designed to provide as estimate of a subject's verbal intelligence through measuring his hearing vocabulary."[5] The test is administered only over the critical range for a particular subject. This is accomplished by establishing a basal of eight consecutive correct answers and not presenting any additional plates to the subject after a ceiling of six incorrectly identified plates in any series of eight.

The test itself was constructed by first selecting a total of 3,885 words that could be illustrated from the 1953 edition of Webster's New Collegiate Dictionary. These words were illustrated and administered to 360 subjects ages two through eighteen to establish the level where the stimulus word was matched with the correct picture by forty to sixty per cent of a particular age group. Following this a series of 200 plates was developed, 8 for each of 23 age levels from $2\frac{1}{2}$ years to 18 years plus 16 additional plates for below the $2\frac{1}{2}$ year level to enable the use of the test with either very young or mentally retarded children. The plates were then administered in three forms, which totaled 600 stimulus words, to another 750 subjects in counterbalanced order. The data derived from

---

[5] Idem., p. 25.

this testing was used to arrive at the final 150 plates. The best 300 words were divided to provide two forms of the test.

The Peabody Picture Vocabulary Test was standardized on 4,012 individuals distributed over nineteen age levels starting at 2½ years and progressing upward at 6 months intervals to age 5 then year intervals to 18 years of age. The number of individuals in each group ranged from 92 in the 2½ and 3 year age groups to a high of 384 in the 11 year group. The average for all groups was 211 and there were 227 in the median group. The data obtained from the testing of these individuals was used to establish age norms derived by the Mean Test Age method. Intelligence quotients were derived by setting an IQ of 100 equal to the mean raw score for each age level then, using a standard deviation of 15 points, converting the raw score distribution to a standard score scale. This scale has seen widespread use but is limited in the scope of information it provides.

Dorothea McCarthy in 1959 presented a language evaluation procedure which she used as part of a general psychological evaluation battery.[6] The total battery consists of thirty-one subtests grouped under six headings: Gross Motor, Manipulative, Apperceptive, Numerical, Language and Conceptual. The Language part contains five measures, three based on tests and two on ratings. The tests are Pointing to Pictured Objects, Picture

---

[6]Dorothea McCarthy, "A Preliminary Report on the Verbal Items of a New Psychological Appraisal Test with Institutionalized Mentally Retarded Children," (paper read at the Council for Except. Child., Northeast Assn., Atlantic City, New Jersey, April, 1959.

Naming, and Vocabulary. Pointing to Pictured Objects requires, as one would expect, pointing to the picture named by the examiner. Picture Naming requires the name of the pictured object be given by the subject. Vocabulary requires a verbal definition to thirty different items with the score dependent upon the number defined and the quality of the definition. The ratings are made on a five-point scale for Oral Language Usage and Intelligibility of Speech during the testing session.

The measures were administered to a group of ninety-six mentally retarded children at Letchworth Village in the state of New York by two experienced clinicians. Children were selected for the sample from the age groups of six, eight, ten, twelve and fourteen years. Intelligence quotients taken from the most recent testing ranged from fifteen to seventy-two.

Complete testing was possible on eighty-four of the cases. The Language section was found to correlate .82 with the Conceptual Section, .63 with Numerical, .59 with Manipulative, and .42 with the Gross Motor section. Metropolitan Achievement Test scores on twenty-three of the children had a mean grade level of 2.1 and correlated .55 with the overall scores of the Language section.

McCarthy presented the Means of each of the five measures for each of three groups (Nontrainable, Trainable and Educable) at each of five age levels (6, 8, 10, 12, and 14 years). Any comparison would necessarily involve a knowledge of the composition of the three groups. The "Educable" group consisted of forty-six children who through tests, observations, and tryouts

were able to profit from attendance at the institutional school. Nineteen children were being tried in "sense training" classes and these children were considered the "Trainable" group. The thirty-one considered "Not-trainable" were those who merely sat each day because they were unable to participate in an educational setting or work placement. Use of McCarthy's scale would involve administering the five measures and comparing the scores derived in this manner with the Means presented in her research. This is at best quite limited but improvements could be easily made through additional research.

In 1960 Lassers and Low made one of the first attempts to measure speech and language within a structured social setting.[7] They viewed language as essentially a social process and developed the San Francisco Inventory of Communicative Effectiveness to rate and measure various attributes of language occurring in a standardized interview situation.

The San Francisco Inventory of Communicative Effectiveness (SFICE) is composed of four sections. The first section is called the "Greetings" section and "was designed to initiate a communicative relationship between child and examiner through a common social greeting and to test the child's ability to make the appropriate social responses."[8] The examiner provides

---

[7]Lassers and Low, "A Study of the Relative Effectiveness of Different Approaches of Speech Therapy for Mentally Retarded Children," (Report to the Office of Education on Contract No. 6904, 1960).

[8]Idem., p. 112.

such standard stimuli as "Hello," "How are you today?" "My name is _____," and if no response is elicited by this last stimulus he asks "What is your name?" Section two is referred to as the "House and Family" section. It "was designed to evoke conversational speech about a familiar topic."[9] The stimuli provided by the examiner for this section range from "Where do you live?" to presenting a picture of a house and family and saying, "Look, (child's name). Here's a house. Here is a living room where the children watch television. Here is the kitchen where mother cooks. Father is cutting the grass. Mother is holding the baby. And the children are playing ball. Tell me more about your house and what you do there."[10] Section three is termed "Situational Pictures." The two pictures, one the scene of a birthday party and the other a classroom "present a simple situation familiar to the child and often evoke verbal responses which may not otherwise be easily or spontaneously elicited."[11] The examiner informs the child "I'm going to show you a picture. Tell me a story about it." The results of the first nine responses to this section are the data analyzed by the verbal output scale.[12] The fourth section uses a story titled "Surprise." It is a simple story designed "to motivate

---

[9] Ibid.

[10] Idem., p. 113.

[11] Ibid.

[12] Idem., p. 136.

interest and elicit verbal response."[13] The examiner reads the story and displays a series of pictures pertaining to the story. After the entire story has been read the examiner shows the child each picture and asks a series of questions about the part of the story portrayed in the picture. All responses to the test are tape-recorded and later analyzed by two scales; the Communicative Attributes Rating Scale and the Verbal Output Scale.

The Communicative Attributes Rating Scale rates the quality of communicative relationships, the quality of communicative behavior, and overall communicative effectiveness.[14] The quality of communicative relationships is determined by placing a subjects test responses somewhere on an eight-point scale with such ratings as "almost none" to "all" on Immediacy (quickness to respond to stimuli), Frequency (how often he responds to stimuli), Enthusiasm (judged by voice quality as well as the above measures), Involvement (degree subject participates in the communicative relationships), and Ease in Social Relationships (judged by the quality, pitch, rate and loudness of the subjects voice).

The quality of communicative behavior is determined by rating Understandability (degree the subject's word can be recognized or understood), Articulation (the extent speech sounds are produced correctly), Average length of response

---

[13] Idem., p. 114.

[14] Idem., p. 130.

(ranging from "almost no response to "very long responses"), Appropriateness (degree the responses are relevent to the test stimuli), and Complexity of Ideation (judged by the complexity of the content of the subject's responses). Overall Communicative Effectiveness is determined by judging the overall impression of a subject's communicative adequacy.

The attributes listed above were determined by the investigators "to be significant components of communicative effectiveness after an extended period of observation and analysis of communicative behavior of mentally retarded children. The attributes were either not identified or were insufficiently defined in the literature or available tests."[15] Three judges rating the language of ninety mentally retarded children through the use of this scale were in high agreement with "median inter-correlations ranging from .61 to .89 with the median of this range at .75"[16] The ratings on Overall Communicative Effectiveness were in such harmony that the median correlation was .89.[17]

As mentioned above the Verbal Output Scale analyzes the first nine responses to section three of the SFICE (situational pictures). This section was selected for analysis "because of their (the pictures) demonstrated motivating power and because

---

[15] Idem., p. 4.

[16] Ibid.

[17] Ibid.

they occurred approximately at a mid-point in the test at which time the majority of subjects communicated relatively freely and comfortably."[18] These responses are analyzed in terms of sentence complexity, grammatical complexity, and length of response. Sentence complexity is scored by a numerical classification system ranging from (0) for not responding or responding unintelligibility to (4) for producing a complex sentence.[19] The grammatical complexity of each sentence is scored by identifying eleven different grammatical elements ranging from pronouns, adjectives, and adverbs to direct objects, compound verbs, and coordinative interjections; allowing one point for each appearance in the sentence.[20] The length of response is scored by simply counting the number of words used by the subject in each phrase of sentence response.

As can be seen the test provides a quite comprehensive look at the language ability of a mentally retarded child. Some disadvantages apparent in the test are the necessity of using a tape-recorder, the possibility of bias in the ratings, and the exclusion of measures of nonvocal language behavior.

Sievers pioneered in the development of a test of linguistics by creating the Differential Language Facilities Test (DLFT).[21]

---

[18] Idem., p. 136.

[19] Idem., p. 133.

[20] Ibid.

[21] Dorothy J. Sievers, "Development and Standardization of a Test of Psycholinguistic Growth in Preschool Children," Selected Studies on the Illinois Test of Psycholinguistic Abilities. (Madison: Photo Press, Inc., 1963).

She cites as the purpose of her research that of: (1) devising "a series of tests which would measure the various aspects of language in preschool children according to a theory of communication as postulated by Osgood, and (2) obtain normative data on the scale with a sample of average children between the ages of two and six."[22]

The first step in the development of this test was the compilation of a list of fifty-seven words that could be represented by both pictures and objects. These words were taken from the Thorndike-Lorge List, the picture vocabularies of the Stanford Binet (forms L and M), the Van Alstyne Picture Vocabulary Test, the Smith list, and the Minnesota Preschool Scale. A total of 198 children two to six years of age were asked to identify the pictured version of the word. The twenty words that showed the least variability over age and no statistically significance difference between sexes were chosen for the test. These words were incorporated into the first subtest, Labeling of Objects and Pictures.

The second subtest of the scale, Object Association, requires an association between such things as a doll dress and a coat hanger when the latter is placed with several other objects; e. g., block, box, and cup. The third subtest of the scale is named Word Association. This subtest requires the presentation of each of the twenty words mentioned above in the form of the question "What other things does _____ make you

---

[22] Idem., p. 1.

think of?" Use definitions such as "build" for the word "block" receive one point, noun associations such as "concrete" for the word "block" receive two points. The fourth subtest consists of a series of Mutilated Pictures. Part of each picture is blanked out and the child receives one point for each picture he is able to identify. The fifth subtest, consisting of four dual line pencil mazes, is called Visual Form Tracing. Each maze has a value of five points if completed without error. One point is subtracted for each error up to five. Subtest number six is called Gesture Sequence Matching. The examiner gestures through various combinations of hand and knee claps and the child receives one point for each correctly copied sequence. Speech Sound Mimicry is the seventh subtest of the series. This requires that the subject listen to nonsense words and repeat them correctly. The words range from "feem" to "skallest"[23] with the subject receiving one point for each correct repetition. In subtest eight, Nonsense Grammatical Mimicry, the child is asked to repeat sentences containing nonsense words. He receives one point for each correctly repeated word in the sentence. The ninth subtest is named Gestural Conversation. In this subtest the examiner says, "See if you can guess what I want you to do," and attempts to elicit a gestural response from the child by actions ranging from the offering of a block to a child to writing on a piece of paper and handing it and an eraser to the child. Points are based on

---

[23] Idem., p. 4.

the number of gestures required to correctly complete the actions. The tenth subtest is a Picture Series Description one. This test consists of six cards each showing a sequential event. The first card, as an example, has three pictures showing the sequence of a woman pouring a glass of milk for a boy. The child is shown each card and asked to tell what has happened. A score of two points is given if the description of the action is complete and one point is given for a partial response. Subtest number eleven is referred to as Vocal "Cloze." This consists of sentences with certain words missing; for example, "Mother _____ a dress." The child is asked to supply such a word as "has," "bought," or "tore," with scoring based on grammatical form and appropriateness of meaning for the sentence.

The entire test was administered to "normal appearing" children two years, zero months to five years, eleven months who were members of sixteen different nursery schools and kindergartens in Illinois.[24] Each child was administered the short form of the Stanford-Binet, Form L, and only those with intelligence quotients between 89 and 123 (one standard deviation each way from the mean) were included in the sample. The children were also taken from several occupational levels due to language development "favoring the upper occupational levels."[25] The final sample contained 228 children.

---

[24] Idem., p. 5.

[25] Idem., p. 6.

There were no consistent sex differences in performance on the test so the data for males and females was combined. Age norms were based on the median total scores of each half-year group. The medians were plotted, smoothed and converted into a table showing the age equivalent of various ranges of language scores; for example, a score of 132 yields an age equivalent of four years, zero months.[26] Sievers felt it was too early in the development of the test to derive separate norms for the subtests.[27]

A study by James McCarthy explored the language behavior of cerebral palsied children using the forerunner of the ITPA the Differential Language Facilities Test mentioned above.[28] The Gestural Conversation subtest was not included due to its inadequate discrimination at older age levels. A total of eighty-four subjects were secured from every available source and included both public and private institutions, speech clinics, public nursery, public school, private medical practice, and free medical clinic patients. The following criteria were adopted for subjects: (1) mental age between 2-0 and 6-4, (2) intelligence quotient of 70 or over, (3) chronological age between 2-0 and 8-7 (automatically set by 1 and 2 above), (4) no serious

---

[26] Idem., p. 11.

[27] J. J. McCarthy, "Qualitative and Quantitative Differences in the Language Abilities of Young Cerebral Palsied Children," Selected Studies on the Illinois Test of Psycholinguistic Abilities (Madison: Photo Press, Inc., 1963).

[28] Sievers, loc. cit.

sensory or motor defects and (5) medical diagnosis of spastic or athetoid cerebral palsy. The subjects were tested by five trained examiners with all final scoring completed by McCarthy. As many children as possible were seen prior to testing to aid rapport. Testing sites varied but were chosen on the basis of lack of distractions. The results led to the following conclusions: (1) normal children are superior to cerebral palsied children (both spastic and athetoid) of the same mental age, in overall language ability; (2) this linguistic superiority being mainly do to superior expressive and receptive language, inner language did not significantly differ from the normals; (3) with mental and chronological age statistically equated, spastic children proved superior to athetoid children in overall language ability; and (4) this superiority was primarily in the area of expressive language with little difference existing between the two in inner and receptive language.

Research using the Illinois Test of Psycholinguistic Abilities. The ITPA although relatively new has stimulated a great deal of research. The studies presented in this section are felt to be representative of the many ways the ITPA has been studied and been used in studies.

Olson's study compared the linguistic behavior of receptive aphasic, expressive aphasic, and deaf children on the ITPA.[29]

---

[29] J. L. Olson, "A Comparison of Receptive Aphasic, Expressive Aphasic, and Deaf Children on the Illinois Test of Psycholinguistic Abilities," Selected Studies on the Illinois Test of Psycholinguistic Abilities (Madison: Photo Press Inc., 1963).

The author felt differing patterns of linguistic functioning, if present, would aid in making differential diagnoses in future cases. The procedures for determining if such differences exist were: (1) comparing the responses of the three types of subjects on the nine subtests of the ITPA, (2) comparing the linguistic profiles of the three types of subjects ITPA scores with the "theoretically expected profiles predicted from the known language characteristics of the children,"[30] and (3) comparing the ITPA linguistic diagnoses with the clinical diagnoses. The children in the sample were residents of two institutions. The deaf and receptive aphasic were from the Central Institute for the Deaf in St. Louis, Missouri; the expressive aphasic children from the Institute of Logopedics in Wichita, Kansas. Subjects were chosen according to the following criteria: "(1) CA between 5-0 and 9-6, (2) IQ as close to the normal range (80-120) as possible, (3) a competent clinical diagnosis of deafness, receptive aphasia or expressive aphasia, and (4) no compound sensory defect."[31] The final sample contained twenty-seven receptive aphasic children, twenty-five deaf children, and fourteen expressive aphasic children. The statistical techniques used in the analysis of the data were analysis of covariance (to control differences in CA and MA) and binomial expansion (to compare theoretical predictions to actual test scores).

---

[30] Idem., p. 49.

[31] Idem., p. 51.

The results of the study can be summarized as follows: (1) the ITPA does adequately discriminate among receptive aphasic, expressive aphasic, and deaf children: (2) receptive aphasics differed significantly from deaf subjects on four of the nine subtests; the deaf being significantly superior on the Visual Motor Association, Auditory Decoding, Vocal Encoding, and Auditory Vocal Automatic subtests; (3) expressive aphasic children were significantly superior to receptive aphasic children on the Auditory Vocal Association, Auditory Decoding, Auditory-Vocal Automatic, and the Auditory-Vocal Sequential tests with receptive aphasic children achieving significant superiority on the Motor Encoding test, (4) the linguistic profiles of the ITPA scores compared with the theoretically predicted scores far beyond chance for the receptive aphasic and deaf children but were "at or near"[32] chance for the expressive aphasic group, and (5) the ITPA diagnosis and the clinical diagnosis agreed in the case of receptive aphasic and deaf children but not in the case of expressive aphasic children.

The results for (2) above were explained as follows: (a) the deaf were superior on Visual-Motor Association probably because the subtest is designed to assess central language and receptive aphasics usually suffer central nervous system damage rather than a peripheral sense injury which would hamper performance on such a test, (b) Auditory Decoding may have been higher for the deaf not because of better auditory receptive language

---

[32] Idem., p. 58.

ability but because "the deaf can usually learn to lipread while the receptive aphasic usually cannot . . ."[33] which would enhance performance on this test, (c) superiority of the deaf on the Vocal Encoding test, where the child tells all he can about a common object, was felt to reflect that while vocal language ability is severely handicapped by both conditions deaf children are less severely handicapped than receptive aphasic children, and (d) the same reasoning was applied to the Auditory-Vocal Automatic test which "measures a child's functional knowledge of the grammar or structure of the language,"[34] the deaf children being less handicapped linguistically than the receptive aphasic children.

The results for (3) above were explained as follows: (a) the four subtests on which the expressive aphasic children were significantly superior "utilize the auditory mode of input"[35] which would favor expressive aphasic over receptive aphasic children; and (b) the Motor Encoding subtest, on which the receptive aphasic children were significantly superior, requires communication through gestures which is often characteristic of receptive aphasic children.

The interpretation offered for (4) above is as follows: "The deaf and receptive aphasic groups present relatively stable linguistic handicaps with the major deficiency in the

---

[33] Idem., p. 55.

[34] Ibid.

[35] Idem., p. 57.

auditory decoding channel, thus their psycholinguistic profiles are predictable far above the chance level. On the other hand the expressive aphasic group was not at all easily predictable. It is the writer's opinion that this unpredictability is due to the fact that the ITPA provides a psycholinguistic diagnosis more specific than the rather global clinical diagnostic connotations of the term 'expressive aphasic.'"[36]

The discrepancy between ITPA diagnoses and clinical diagnoses for the expressive aphasic children (5 above) were explained by the possibility of expressive aphasia as a diagnosis being so general that expressive aphasic children as a homogeneous group is not an actuality. The author cites: (a) the dissimilar profiles of the ITPA scores, and (b) evidence in the cases of expressive aphasia in the sample of major language disability other than expressive aphasia, to support his argument of heterogeniety of the group.

The areas explored by Bateman's study of the partially sighted were: the level of their reading ability, the kinds of errors made in reading, the relation of visual defect to reading ability, the level of their psycholinguistic ability, the relation of visual defect to psycholinguistic ability, and the relation of reading to psycholinguistic functioning.[37]

---

[36] Idem., p. 58.

[37] Barbara D. Bateman, "Reading and Psycholinguistic Process of Partially Seeing Children," Selected Studies on the Illinois Test of Psycholinguistic Abilities (Madison: Photo Press Inc., 1963).

A partially sighted child was defined for the purpose of the study as "a child who was currently enrolled in a special class or resource room program for partially sighted children which is recognized by the Illinois State Department of Public Instruction."[38] Reading ability was measured by Monroe's Diagnostic Reading Examination with psycholinguistic ability being measured by the Illinois Test of Psycholinguistic Abilities. Subjects were 131 children selected from twenty special class and resource room programs for partially sighted children in the state of Illinois. The following concerns the subjects of the study. The ratio of boys to girls was 3:2. Binet Intelligence scores were normally distributed. Mild visual deficiency accounted for 38% of the defects, 42% had moderate visual deficiencies, and 20% were legally blind. The visual handicaps of females were more severe than those of males. Eye conditions included myopia (29%), popularly called nearsightedness; cataracts (15%), a disease of the eye in which the crystalline lens becomes opaque; retrolental fibroplasia (13%), a defect of the eyes due to excessive oxygen administered at birth; esotropia (9%), a condition in which one eye fixes the object and the other deviates inward; nystagmus (8%), a quick, jerky movement of the eyes, followed by a slower return; albinism (5%), congenital absence of pigment from the iris; and miscellaneous (13%). The severely visually handicapped scored higher in intelligence than the mildly handicapped. Children with errors of refraction

---

[38] Idem., p. 71.

(failure to focus the optical image normally upon the retina) tended to be less severely handicapped, older, and to have lower IQ's than the children with nonrefractive eye conditions.

The results of the study can be summarized as follws: (1) reading achievement of the partially sighted children does not differ significantly from that of normally sighted children; (2) silent reading comprehension of the partially sighted is higher than oral reading scores; (3) discrepancy between actual grade placement and reading grade scores of the partially sighted was less than one-half month; (4) slowness in reading on the part of the partially sighted was not a major problem; (5) results of the reading errors appeared questionable but slight evidence was present that partially sighted have more trouble with vowel discrimination and reversals than other errors; (6) there were no significant differences in the reading achievement of the mild, moderate, and severe visual defect groups; (7) the mild defect group read less well in relation to grade placement and made more errors than the moderate and severe groups; (8) those with errors of refraction read less well than those with non-refractive eye defects; (9) there was no significant differences found in the error profiles of children with errors of refraction and those without; (10) partially sighted children did not significantly differ from normals on the auditory-vocal channel subtests (Auditory Decoding, Auditory-Vocal Association, Vocal Encoding, Auditory-Vocal Automatic, and Auditory-Vocal Sequential) but were significantly lower on the visual-motor channel subtests

(Visual Decoding, Visual-Motor Association, and Visual-Motor Sequential) as would be expected; (11) psycholinguistic profiles of the high IQ (114 or above) and low IQ (86 or below) paralleled each other with the exception of Auditory-Vocal Association which was dis-proportionately low for the low IQ group; (12) psycholinguistic profiles of mild, moderate, and severe visual defectives paralleled each other with slight deviations on Motor Encoding and Auditory-Vocal Sequencing; (13) the mild, moderate and severe visual defectives were all low on the Motor Encoding subtest with the mild defect group scoring higher than the other groups; (14) only the cataract and retrolental fibroplasia groups, who had the more severe visual handicaps, manifest a clear visual channel problem on the ITPA, (15) the degree of defect, rather than type, had the greatest influence on ITPA performance; and (16) reading achievement showed a significant positive correlation with the Auditory-Vocal Sequential subtest (.44), Visual-Motor Sequential (.35), and Auditory-Vocal Automatic (.43) subtests. These are all measures of the Automatic-Sequential level which like reading involve use of symbol sequences. The author states she feels the "ITPA as a whole appears to be an excellent diagnostic aid to be utilized in determining level and mode of visual functioning in partially seeing children."[39]

---

[39] Idem., p. 83.

The purpose of a study by Kass was to "discover some psychological correlates of reading disability (dyslexia) . . ."[40] This reading disability was defined as "retardation in reading skills which occurs after adequate instruction and which is not due to mental retardation or to sensory defects . . ."[41] The author used the ITPA and some additional tests to assess several facets of psycholinguistic abilities. The additional tests were felt to be necessary to build a more complete clinical model for studying reading than could be accomplished with the ITPA alone. The tests she added were: (1) Visual Automatic (where the child views an unfinished picture and is asked to guess what the picture will be when completed), (2) Sound Blending (where the child listens to sounds and attempts to fuse them into words), (3) Mazes (where the child must predict which route leads to successful completion), (4) Memory-for-Designs (where the child is asked to view a design and reproduce it from memory), and (5) Perceptual Speed (where the subject is required to make rapid comparisons of figures visually).

A sample of twenty-one children was composed who met the following criteria: "(1) Chronological age between 7-0 and 9-11, (2) Normal intelligence (determined by Stanford-Binet,

---

[40] Corrine E. Kass, "Some Psychological Correlates of Severe Reading Disability (Dyslexia)," Selected Studies on the Illinois Test of Psycholinguistic Abilities (Madison: Photo Press Inc., 1963), p. 88.

[41] Ibid.

Form L-M), (3) In the second, third, or fourth year in the primary grades (including repeaters and not counting kindergarten), (4) Retarded in reading on a battery of diagnostic reading tests . . .; one-half year retarded if in second year in school; one and one-half, if in third year; and two and one-half, if in fourth year, and (5) no known auditory or visual defects."[42] The twenty-one were administered the ITPA and the five additional tests and the results were statistically analyzed via the z-test to determine if the sample significantly differed from the theoretical average.

The results of the study can be summarized as follows: (1) Children with reading disability were above the theoretical average for "normal" children in Visual Decoding (ability to understand what is seen); (2) Children with reading disability were similar to the theoretical average for "normal" children in Auditory Decoding (ability to understand what is heard), Visual-Motor Association (ability to draw relationships from what is seen), Vocal Encoding (ability to express verbally), Motor Encoding (the ability to express ideas gesturally), and Auditory-Vocal Sequential (ability to reproduce a series of symbols presented auditorily); (3) Children with reading disability "appeared to have marginal deficit"[43] in Auditory-Vocal Automatic (ability to use the structure of language automatically), and Visual Automatic (ability to predict a whole

---

[42] Idem., p. 91.

[43] Idem., p. 93.

from a part); and (4) the tests which children with reading disability exhibited the greatest deficit were Auditory-Vocal Association (ability to draw relationships from what is heard), Visual-Motor Sequential (ability to reproduce a series of symbols presented visually), Sound Blending (ability to blend parts into a whole), Mazes (ability to manually execute a visual prediction), Memory-for-Designs (ability to manually represent a visual image from memory), and Perceptual Speed (ability to visually compare detailed figures rapidly).

These results indicate the following: (1) children of normal intelligence with reading difficulties may compensate for their difficulty in handling symbols in reading by gathering information from pictures, and (2) children with reading disability tended to have more deficiencies at the integrational level than at the representational level of psycholinguistic functioning.[44]

The purpose of an article by Bateman was to point out some of the ways the ITPA can be put to use in providing information concerning treatment or remediation that will maximize the functioning of a mentally retarded person.[45] Intelligence tests are for the most part too global to meet this need; yet they have been used to "categorize" children who perhaps differ

---

[44] Ibid.

[45] Barbara D. Bateman, "The Role of the ITPA in Differential Diagnosis and Program Planning for Mentally Retarded," Amer. J. Orthopsych., 1965, 35, pp. 465-472.

markedly in the more molecular patterns and ranges of abilities and disabilities.

The ITPA, providing a profile of nine specific psycholinguistic areas, point out the areas that are relatively intact and need only be improved or that can be used to help develop other areas shown by the test to be deficient. Another distinct advantage of the ITPA is that it attempts to provide a measure of nine areas where "only one ability is required"[46] whereas on an item such as a Binet verbal absurdities, "it is difficult to know if he failed to understand the auditory stimulus, failed to detect the absurdity in the statement or was unable to adequately express his understanding."[47] She points out the test "is designed to help fill the gap between diagnosis of language disorder and the development of remedial techniques for those areas of disability."[48] Four areas of deficiency have been commonly revealed by the ITPA. These are: (1) Auditory Vocal channel disability, (2) Visual-motor channel disability, (3) Perceptual level, and (4) Encoding process. Disability in the Auditory Vocal channel manifests itself in low scores on the Auditory Decoding Test (ability to understand what is heard), Auditory Vocal Automatic (ability to use the structure of language automatically), and Auditory Vocal Sequential (ability to reproduce a series of symbols presented auditorily). According

---

[46] Idem., p. 466.

[47] Ibid.

[48] Idem., p. 467.

to the author "remediation usually must begin with auditory decoding, or sometimes with auditory discrimination."[49] Those who show difficulty on the Visual-motor channels are frequently those with "'minimal brain injury,' 'perceptual disorganization' or the 'Strauss syndrome.'"[50] They are most likely to show defects on Visual Decoding and Motor Encoding.[51] Deficiency in the perceptual level (ITPA Automatic-Sequential) would appear in the Auditory-Vocal Automatic, Auditory-Vocal Sequential and Visual Motor Sequential. Bateman states "mentally retarded children most often show lower scores on the memory and closure tests at this level . . ."[52] A fourth area is in the Encoding process which would suggest deficiency in the ability to express verbally (Vocal Encoding) and the ability to express ideas gesturally (Motor Encoding). The author states "children, often labeled expressive aphasics, whose profiles are characterized by much higher performances in decoding (comprehension) than in encoding (expression)"[53] display this deficiency.

There is need for much additional research to develop further remedial techniques but as the author states, "the research described here is hopefully one step in this direction."[54]

---

[49] Idem., p. 466.

[50] Ibid.

[51] Idem., p. 467.

[52] Idem., p. 468.

[53] Ibid.

[54] Idem., p. 465.

Another study by Bateman with Wetherell as co-author surveys the psycholinguistic functioning of mentally retarded children.[55] Some of the deficits of the mentally retarded were mentioned briefly in the article just reviewed.[56] In this article the authors attempt to answer the question of what are the psycholinguistic aspects of mental retardation by posing four questions: "(1) Do these children differ psycholinguistically from average children of comparable mental and chronological age? (2) Are there differences between the psycholinguistic profiles of children of various IQ levels and diagnostic categories? (3) Are there differences between the psycholinguistic performances of urban and rural low IQ children? and (4) What psycholinguistic patterns can be seen in the ITPA profile of a 'typical' retarded child?"[57]

The data indicate the following in relation to (1) above: (a) all the mean language ages of the subtests for the retarded children were below their mean mental age except the Visual motor association subtest, (b) greater strength of performance was present at the representational than at the automatic-sequential level. The results of (b) above suggest: (1) retarded children "have experienced a greater educational emphasis on meaningful than on automatic language usage, (2) have greater

---

[55]Barbara D. Bateman and Janis Wetherell, "Psycholinguistic Aspects of Mental Retardation," Mental Retardation, 1965, Vol. 3, No. 2, pp. 8-13.

[56]Bateman, loc. cit.

[57]Bateman and Wetherell, op. cit., p. 8f.

'actual' retardation or developmental lag in the habitual or automatic manipulation of linguistic symbols, or (3) that the test items at the automatic-sequential level are less motivating to retarded children than to the standardization population."[58]

The following was presented as the answers to "Are there differences between the psycholinguistic profiles of children of various IQ levels and diagnostic categories?" ITPA profiles of children of different IQ levels and diagnostic categories suggest: (a) no difference in expressive ability (vocal and motor encoding) of those with IQ's above 90 and those with IQ's of 80 to 90, both groups being high in vocal encoding and low in motor encoding; (b) Visual-motor sequencing did not significantly differ across the high (90 and above), middle (80-89), and low (58-79) IQ groups; (c) those with IQ 58-79 had a flatter ITPA profile suggesting "more generalized mental and/or psycholinguistic retardation . . .;"[59] (d) children in the high IQ group were quite deficient in motor encoding and visual motor sequencing; (e) children in middle IQ group presented a flat profile with the exception of high vocal encoding and quite low visual motor sequencing; (f) children in low IQ group were low in motor encoding and the entire automatic-sequential level; and (g) Mongoloids were similar to an equated group of Non-Mongoloids except the Mongoloid group was superior in motor encoding and inferior in auditory vocal automatic, both groups had greater

---

[58] Idem., p. 9.

[59] Idem., p. 10.

deficit in the automatic-sequential level than in the representational level.

There were no significant differences between low IQ rural and urban children in psycholinguistic ability after both had received a year and one half of schooling nor were there significant differences for the same group after three years of schooling (about half in regular class and the other half in special classes). The psycholinguistic pattern seen in the ITPA profile of a "typical" retarded child consists of (a) automatic-sequential level deficiency, (b) some strength in the visual channel over the auditory channel at the representational level. This knowledge points out the need for "repetition, overlearning and 'mechanical' drill"[60] in language training as well as other learning situations. The authors also feel "the ITPA has . . . contributed substantially to our knowledge of the language process of retarded children and may point the way for the development of remedial language training."[61]

A study by Olson, Hahn and Herman attempted to derive some guidelines for consideration in planning curriculum for mentally retarded children based on psycholinguistic strengths and deficits.[62] A program was developed by the authors of this publication in which the goals were focused upon the language

---

[60] Idem., p. 12.

[61] Idem., p. 13.

[62] J. L. Olson, R. Hahn, and Anita L. Hermann, "Psycholinguistic Curriculum," Mental Retardation, 1965, Vol. 3, No. 2, pp. 14-19.

development of mildly retarded children. The seven children selected for the program were educable mentally retarded under seven years of age who through intelligence tests or clinical interview gave evidence of capacity to benefit from an individualized program. Each child received the Stanford-Binet, the Peabody Picture Vocabulary Test, and the Illinois Test of Psycholinguistic Abilities. A staff conference was held to plan an individualized program for each child. The program was eight weeks long and the children attended class two and one-half hours a day, five days a week. The program involved the entire group, during certain periods, where the nine psycholinguistic areas of the ITPA were emphasized. This was accompanied by individual tutoring in specific areas of deficit. The type of instruction provided followed that outlined by Wiseman.[63] The authors present several case histories and explain where remediation was attempted and their results. No statistical analysis was attempted due to the small number of subjects in the sample. There was a slight increase in intelligence test scores on the Stanford-Binet and Peabody Picture Vocabulary Test but as the authors point out this could be due to either random errors of measurement or improvement as a result of the program.

This increase did cause the authors to raise the following questions: "Does a curriculum emphasizing an individualized approach to remediating language deficits increase measured

---

[63] D. E. Wiseman, "Program Planning for Retarded Children with Psycholinguistic Disabilities," (Unpublished Doctoral dissertation, The University of Illinois, Urbana, 1964).

intelligence? Would significant IQ changes have occurred if the program had been continued beyond the two-month period? and Does the PPVT measure more adequately than the Stanford-Binet the abilities stimulated by this type of program?"[64]

The pre- and post-test ITPA's were the same for three of the children with the other four demonstrating some gain in language age. The mean gain on representational level subtests amounted to ten and one-half months which created the following questions in the authors' minds: "Are some children with psycholinguistic disabilities better able to profit from an individualized remedial language program? If so, how can we identify these children? Does the greater mean gain at the representational level (19.5 months) indicate that the abilities assessed by the automatic-sequential level (three months' gain) are more resistant to educational remediation? Does the differential growth rate indicate that the abilities assessed at the automatic-sequential level are more dependent upon innate characteristics than are the abilities assessed at the representational level? Do organically impaired children respond as readily as culturally deprived children to an individualized remedial language program?"[65]

Wiseman has taken the initial step in bridging the gap between the diagnosis of linguistic deficiencies and their

---

[64] Olson, Hahn, and Herman, op. cit., p. 18f.

[65] Idem., p. 19.

remediation.[66] He proposes a systematic method of remediating areas of language disability as diagnosed by the ITPA. The proposals cover five areas: understanding or decoding, association of ideas, expressing ideas or encoding, automatic or closure processes, and memory.

Auditory understanding or decoding ability according to Wiseman can be improved through such techniques as asking a child questions that are answered by yes, no or maybe; giving him verbal directions to follow; and having him repeat or answer questions about a story. Visual understanding or decoding ability can be facilitated by having the child identify objects in pictures; identify colors, forms, numbers, letters; and explain the action of various pictures.

Improvement in auditory association can be brought about by having the child classify things ("Name all the birds you can think of"); categorize ("plow, pitchfork, tractor belong to?") establish similarities ("In what way are a horse and a cow alike?"); and associate cause and effect ("What would happen if . . ?). Visual Association can be improved by having the child classify pictures; establish similarities between pictures ("car and truck"); identify the odd picture in a series; and arrange pictures to tell a story.[67]

---

[66] D. E. Wiseman, "A Classroom Procedure for Identifying and Remediating Language Problems," Mental Retardation, 1965, Vol. 3, No. 2, pp. 20-24.

[67] Idem., p. 22.

Vocal encoding ability (expression of ideas through speech) can be helped by having the child describe an object and make up a story concerning it; verbally instruct another child or group about something; and by solving some verbal problems ("If you couldn't find a toy in your toy box, how would you go about finding it?"). Ways of enhancing the expression of ideas through actions (motor encoding) include impersonating the movements of an animal and having the child imitate this; having the child draw pictures; and playing charades.

The auditory part of the automatic or closure process can be improved through separating the letters or syllables of a word and presenting this with several pictures and asking the child to identify the correct picture (auditory sound blending); through saying "This is a very nice _____." and having the child complete the sentence by identifying a particular object; and teaching the child to give a specific response to various words (such as the response "tree" to the word "elm"). Visual closure improvement is brought about by having the child identify shadows of objects; identifying forms hidden in a background; and identifying incomplete forms.

Auditory memory development is helped by having the child repeat sentences; retell stories in his own words; repeat unrelated words; and repeat letters, numbers, etc., in sequence. Visual memory development is helped by showing a child a picture that contains a story and having him tell the story upon removal of the card; look at a page of pictures and recall them from memory; having the child close his eyes after looking at an

arrangement of objects, rearrange them and have him put them back in their original arrangement.

Wiseman feels "the correction of defective areas should promote a more orderly development of the language process, hopefully leading to greater mental efficiency."[68]

A recent publication by Bateman reviews fourteen additional studies as well as four of the ones presented in this review.[69] Of these fourteen, three are statistical studies that employed a factor analysis of the ITPA.

The first study reviewed, an unpublished one by Semmel and Mueller of George Peabody College, explored "the factor validity and structure of the ITPA as applied to mentally retarded children."[70] The centroid factor analysis performed on the results of one hundred eighteen administrations of the ITPA to mentally retarded children revealed nine specific factors, one for each of the abilities defined by the test. They also found each subtest correlated positively with total language age (with the correlations ranging from .52 to .81) and with mental age (with the correlations ranging from .33 to .62).

---

[68] Idem., p. 22.

[69] Barbara D. Bateman, The Illinois Test of Psycholinguistic Abilities in Current Research (Urbana: Institute for Research on Exceptional Children, 1964).

[70] M. I. Semmel and M. W. Mueller, "A Factor Analysis of the Illinois Test of Psycholinguistic Abilities with Mentally Retarded Children," Cited by Bateman in The Illinois Test of Psycholinguistic Abilities in Current Research (Urbana: Institute for Research on Exceptional Children, 1964), p. 1.

A second study reviewed by Bateman, this one by Center, sought to determine the number of factors represented by a battery of tests which included the ITPA, an "Achievement Battery made up of selected subtests," and the "Partial PMA Battery."[71] These tests were administered to forty-eight children who were "a culturally homogeneous third grade group of 23 boys and 25 girls between the ages of 8 and 9 years in a Georgia county school system."[72] The results for the ITPA can be summarized as follows: (1) Vocal and Motor Encoding loaded on an Extraversion factor; (2) Visual Decoding was highly loaded on a Visual, Non-Verbal Perception factor and moderately loaded on a Reasoning factor; (3) Auditory Decoding had a high loading on a Verbal or Auditory Comprehension Factor; (4) Visual Motor Association and Visual Decoding loaded significantly on a Visual, Non-Verbal Perception Factor; (5) Auditory-Vocal Sequencing, Auditory-Vocal Automatic, and Auditory-Vocal Association loaded significantly on an Inner Language Factor; (6) Auditory Decoding loaded moderately on the Perception factor but was highly loaded on Verbal or Auditory Comprehension; and (7) Visual Motor Sequencing loaded very high on a Memory factor.

The third study reviewed, that by Loeffler, used the ITPA as part of a battery of 32 tests given to one hundred mentally

---

[71] W. R. Center, "A Factor Analysis of Three Language and Communication Batteries," Cited by Bateman in The Illinois Test of Psycholinguistic Abilities in Current Research (Urbana: Institute for Research on Exceptional Children, 1964), p. 2.

[72] Ibid.

retarded children in an attempt to isolate a possible seven primary abilities.[73] The seven hypothesized abilities were: "Figural Identification, Verbal Comprehension, Fluency, Figural Reasoning, Semantic Reasoning, Immediate Memory for Symbolic Units, and Immediate Memory for Figural Units."[74] The tests loaded significantly on six of the seven primary abilities. The tests that were predicted to load highest on Semantic Reasoning were highest on Verbal Comprehension.

Research using the Parsons Language Sample. The Parsons Language Sample is not as readily available as the ITPA and therefore has not been used as extensively in research as the ITPA. The studies have been mainly unpublished since the original monograph on the PLS appeared in 1963. Its main use in research has been to classify the children in the sample of these research studies as "high" or "low" in verbal ability.

One such study by Siegel used the PLS in this manner.[75] Subjects who scored within the upper 25% on the PLS were classified as high in verbal ability while those in the lower 52% were considered low in verbal ability. Siegel's study tested the hypothesis that the "verbal behavior of adult subjects

---

[73] F. J. Loeffler, "An Extension and Partial Replication of Meyers, et. al. Primary Abilities at Mental Age Six," Cited by Bateman in The Illinois Test of Psycholinguistic Abilities in Current Research (Urbana: Institute for Research on Exceptional Children, 1964), p. 3.

[74] Idem., p. 4.

[75] G. M. Siegel, "Language Behavior of Adults and Retarded Children in Interpersonal Assemblies, I. Adult Verbal Behavior in 'Play Therapy' Sessions with Retarded Children," J. Speech Hear. Disord., Monogr. Suppl. 10, 1963), pp. 34-38.

would vary as a function of the linguistic level of a child with whom they were assembled in a series of permissive "play-therapy like sessions.'"[76] The hypothesis was not confirmed but the verbal output of the highs was significantly greater than that of the lows indicating the value of the PLS in predicting the verbal behavior of retardates.

A second study by Siegel also used the PLS to classify subjects as "high" or "low" in verbal ability with highs scoring in the upper 25% on the vocal section of the PLS and lows scoring between the 25th and 50th percentiles on the entire scale. Horowitz used a similar technique in a study concerning reinforcement of vocal responses[78] as did Copeland in studying the effects of feedback on verbal behavior.[79]

Other studies by Siegel and Harkins, Spradlin and Rosenburg, and Rosenburg, Spradlin and Mabel again used the PLS and are summarized below in order to show more completely the reasons for classifying subjects as "high" or "low" in verbal ability.

---

[76] Idem., p. 34.

[77] G. M. Siegel, "Language Behavior of Adults and Retarded Children in Interpersonal Assemblies, 3. Verbal Behavior of Retarded Children Assembled with Pre-Instructed Adults," J. Speech Hear. Disord., Monogr. Suppl. 10, 1963, pp. 47-53.

[78] Frances D. Horowitz, "Effects of Consequences on Vocal Behavior, I. Partial and Continuous Reinforcement of Vocal Responses Using Candy, Vocal and Smiling Reinforcers Among Retardates," J. Speech Hear. Disord., Monogr. Suppl. 10, 1963, pp. 55-69.

[79] R. H. Copeland, "Effects of Consequences on Vocal Behavior, 2. The Effects of Feedback Modification on Verbal Behavior," J. Speech Hear. Disord., Monogr. Suppl. 10, 1963, pp. 70-75.

Siegel and Harkins[80] and Harkins[81] felt it was reasonable to expect differences in the verbal behavior of adults when placed in interpersonal assemblies with children of different language level. They tested this hypothesis in two experimental conditions: Unstructured, which allowed the adult considerable freedom of verbal behavior; and Structured, which imposed restrictions on his verbal behavior. This permitted a comparison between two verbal situations independent of the verbal level of the children. The children themselves were classified as either "high" or "low" in language ability by their scores on the Parsons Language Sample test. The "high" group consisted of twenty-one males who scored within the first quartile on the test. The "low" group was composed of twenty-one retardates whose scores fell in the quartile below the median. No child with verbal ability less than this was included in the study because "in many cases they are not ambulatory and because they often present extreme management problems."[82] In addition to this no child who presented a markedly deviant appearance was included in the sample. It was felt this might cause an adult to react to appearance rather than to language level.

---

[80] G. M. Siegel and J. P. Harkins, "Language Behavior of Adults and Retarded Children in Interpersonal Assemblies, 2. Verbal Behavior of Adults in Two Conditions with Institutionalized Retarded Children," J. Speech Hear. Disord., Monogr. Suppl. 10, 1963, pp. 55-69.

[81] J. P. Harkins, "A Study of the Verbal Behavior of Adults Assembled in Two Conditions with Institutionalized Mentally Deficient Children of Two Verbal Levels," (Unpublished Master's Thesis, The University of Kansas, 1961).

[82] Idem., p. 19.

All subjects saw one high and one low child. The children were randomly assigned to each adult subject. The normal adult subjects were 21 male students from Parsons (Kansas) Junior College who volunteered to participate and receive one dollar per hour for their part in the research. They were not informed of the hypothesis or the procedures. The subjects were told the experiment was to test the effects of individual instruction on the learning behavior of retarded children and that they were to try to teach them to assemble a formboard depicting a three-dimensional farm scene. The adult subject was settled in the experimental room and the first of the two children was brought into the room. The adult was informed the experiment would begin in several minutes and to listen for the buzzer which would indicate the equipment was set. The experimenter timed and recorded this unstructured session without the subjects knowing this was part of the experiment. After five minutes the buzzer sounded and the subject started trying to instruct the child in the proper assembly of the formboard. Five minutes later the buzzer sounded again signaling the end of the session and the child was returned to his cottage. The entire ten minute session was recorded. The same procedure was then repeated for the second child.

The recordings were transcribed by two typists. They randomly typed the various tapes and randomly determined which condition of each tape would be typed first. Both typists transcribed ten of the tapes and were found to do so reliably. From these typed records of the sessions six measures were

extracted, these were: (1) total number of responses, (2) total number of questions, (3) ratio of questions to responses, (4) total number of words, (5) mean length of response, and (6) Type-token ratio (the ratio of the number of different words to the total words in a sample of language).

The first step taken in the analysis of the results was to establish that a difference did exist between the "high" and "low" children in verbal behavior. The "high" children were found to use significantly more words than the "low" children. The hypothesis of the experiment was confirmed. The verbal behavior of adults did vary as a function of the language level of the mentally deficient children with whom they were assembled. The findings supporting this conclusion are: (1) the subjects made significantly more responses when assembled with low level children than when assembled with highs (the total verbal behavior of the adult remains constant but is broken up into shorter units); (2) the type-token ratio being significantly different for both "high" and "low" levels and structured and unstructured conditions (a more diverse vocabulary was used when subjects were assembled with high level children than when assembled with lows and also when in the unstructured condition than in the structured condition); and (3) no significant difference between levels in the total number of questions asked by the adults nor the ratio of questions to responses. The ratio of questions to responses for both levels was approximately twice as great in the unstructured condition as in the structured. Harkins felt this supported "the notion that

adults did not require extensive verbal feedback from the children when teaching them the task."[83]

The results of the study suggest: (1)"verbal behavior is subject to social or interpersonal modification and that one of the principal controlling variables is the verbal feedback received from the interactors in the environment," (2) "that the child, himself, through his own vocal behavior, may partially control the type of language stimulation he receives, (3) "that the low language levels of some children may elicit verbal behavior from the environment which inhibits the learning process. Thus, it may be that the child's behavior acts to depress the quantity and quality of the language stimulation he receives, thus increasing the degree of retardation," and (4) "adult verbal behavior may be modified by such other factors as the social situation and the diagnosis or label attached to the child." This being a perpetuating process with the child's behavior influencing the social situation and diagnosis.[84]

Spradlin and Rosenberg conducted a study concerned with the verbal behavior of normal adults when placed in an interaction situation with a retarded child.[85] Informal observations revel that normal adults who try to engage in conversation with

---

[83] Idem., p. 44.

[84] Idem., p. 84.

[85] J. E. Spradlin and S. Rosenberg, "Complexity of Adult Verbal Behavior in a Dyadic Situation with Retarded Children," (Parsons Research Project Working Paper 18, 1959, Parsons State Hospital, Parsons, Kansas), (Mimeographed).

retarded children of limited speech ask binary rather than multiple questions (a binary question being one that has to be answered by one of two responses such as yes or no, where a multiple response question requires a more complex response). If an adult's questions are dependent upon reinforcement then it seems likely multiple response questions would tend to undergo extinction and binary questions would maintain or increase their frequency of occurrence when conversing with a retarded child whose speech is inadequate. This study represents "an attempt to determine whether normal adults increase the percentage of binary questions asked as a function of interaction with children of varying degrees of verbal skills in an interview situation."[86]

The first step in testing this hypothesis was to give ninety-six mentally retarded children between the ages of six and eighteen two subtests of the Parsons Language Sample. The two subtests were Intraverbal which is also included in the revised version of the PLS with modifications, and the Gestures subtests which has also been revised and renamed Intraverbal Gestural. Children who received a score of at least seventeen were classified as high verbal, those below seventeen were classified as low verbal. Children who failed to give at least one appropriate verbal response or three appropriate gestural responses were not considered further as candidates for the study.

---

[86] Idem., p. 2.

The subjects in this study were sixteen junior college students, five males and eleven females, selected by a Dean of the Parsons (Kansas) Junior College from "among the more competent students enrolled in the Junior College."[87] They were told they would be interviewing retarded children and would receive $5 for the six interviews.

The experiment was conducted in the following steps. The experimenter introduced himself to the examiner, took him to the interview room, gave him a list of topics on which to interview the child and some candy, then told him to listen to the instructions for interviewing as they were played into the room via a tape recorder. The experimenter then left and allowed the subject to read over the topics at least five minutes prior to the arrival of the first child. When the child arrived he was introduced to the subject then left alone with him for a twenty-minute interview. After the session the experimenter asked the subject to write his impressions of the child in order to give the subject the impression the interview was real and for a purpose. Essentially the same procedure occurred for the other children; all interviews being tape recorded.

The sixteen subjects were randomly assigned to the four schedules until there were four subjects for each schedule. Each schedule consisted of six interviews and were composed as follows: one with six high verbal children, one with three

---

[87] Idem., p. 3.

highs followed by three lows, one with three lows followed by three highs, and one with six low verbal children. This permitted study of subjects reaction to different verbal levels and any transfer of effect from low to high and high to low in the series.

The tape recorded interviews were heard by two raters who classified the subject's questions as either binary, mixed or multiple. The experimenters tried to free rater bias through: (1) selecting raters that were in highest agreement with the experimenters on a small random sample of the recorded interviews, (2) extensive training on the correct procedure for rating, (3) maintaining naiveness concerning the experimental design and hypotheses under study, (4) scrambling the presentation of recorded interviews to the raters, (5) checks on inter-rater agreement, and (6) separation of the two raters.

The results of the experiment can be summarized as follows: (1) mixed responses occurred with such low frequency that an analysis of them was not made, (2) low verbal children elicited proportionally more binary questions than the high verbal children in the initial interviews, (3) there was no consistent difference in type of questions elicited during later interviews.

Some possible reasons for the essentially negative results were suggested in the discussion: (1) a simple dividing point was made to separate high from low verbal level thus "in some cases high verbal children may have presented very similar behavior to those classified as low verbal," (2) the unlikely possibility that "perhaps there was no difference in the

percentage of multiple and binary questions answered by children with high and low verbal levels," and (3) questions as to "whether the verbal behavior of an adult can be affected in so brief a period as twenty minutes."[88]

A study by Rosenberg, Spradlin, and Mabel employed two subtests of the Parsons Language Sample to classify subjects as either "high" or "low" according to linguistic skill.[89] The study was designed to determine differences in interpersonal behavior among retarded children of different language ability. Pairs of children with different levels of linguistic skill were assembled for observation and records made of their interpersonal behavior as they engaged in free play. Children were placed in pairs on the basis of their scores on the Intraverbal and Tact subtests of the Parsons Language Sample. If a score of at least one was not obtained on both subtests combined the child was excluded from the study. The scores on the Intraverbal subtest for the remainder of the children were used to classify their level of linguistic skill. Those whose score was greater than twenty were classed as "high" with those whose score was less than nine being classed as "low." Children in the intermediate range (nine to twenty) were eliminated. "High" and "low" children also differed in IQ. The "high" subjects ranged from forty-five to eighty-two on the verbal scale of the WISC.

---

[88] Idem., p. 11.

[89] S. Rosenberg, J. E. Spradlin, and S. Mabel, "Interaction Among Retarded Children As a Function of Their Relative Language Skills," J. Abnorm. soc. Psychol., 1961, 63, pp. 402-410.

The "low" subjects were all below the lower limits of the IQ range of the WISC verbal scale.

Subject selection was accomplished by first selecting 64 children at random from a total of 159 children in the hospital between the ages of 12 years, 11 months and 15 years, 0 months. These sixty-four children were tested as described above resulting in eight males and thirteen females classified as "high"; six males and eleven females classified as "low." Sixteen girls and twelve boys were randomly selected as subjects from this pool. Eight of these had to be eliminated because one child in each of two groups of four began to refuse to cooperate. Initially, however, seven subgroups were formed each with four children. Two "high" and two "low" subjects of the same sex were randomly assigned to each group from the pool of twenty-eight randomly selected subjects. This created three male and four female subgroups.

The four subjects in a subgroup were brought to the two playrooms in the research area of the hospital. Two dyads were formed and each was placed in a playroom for a fifteen-minute period. Then two new dyads were formed from the same four subjects for the second fifteen-minute period. This was followed by two more play periods to complete the matrix of: one "high" with the second "high," one "low" with the second "low," one "high" with the second "low," and the second "high" with the first "low." Two completions of this matrix occurred during each session. Two dyads of a subgroup were assembled at the same time one in each of two playrooms. Each dyad spent

one play period in each of the two playrooms at each session. The playrooms each contained one table, two chairs, one beach ball, one teddy bear, and one package of colored clay.

Thirteen sessions were conducted for each subgroup. The first session began three weeks after the administration of the two Parsons Language Sample subtests and subsequent classification of subjects. The first eight sessions were conducted at two to four day intervals with the schedule altered to once a week for sessions nine through thirteen. If a child was absent for a session the session was postponed to the next one on the schedule. Illness, vacations and other miscellaneous causes extended the total time to complete all sessions to twenty-seven weeks in some subgroups. As mentioned above two subgroups were dissolved because of one of the members of each refusing to stay in the playroom for the fifteen-minute period.

Two observers and an alternate were trained to judge the occurrence of three behaviors: vocalization, gesture, and physical contact. They judged the last five sessions. Data from the last four sessions was used in the analysis. The first nine sessions were to allow the two person relationships to realize some degree of stability and in this way partially neutralize the extent of unequal contact among the subjects prior to the study. The observers were two junior college students and one housewife selected from a group of six housewives and twelve students on the basis of availability, interest, and the accuracy of their observations of the three behavior

measures when viewing a twenty-minute film titled, One, Two, Three, Go, produced by Metro-Goldwyn-Mayer in 1946.

The measures obtained by the observers included: "total frequency of vocal responses during a fifteen-minute play period, total frequency of gestural responses during the play session, and square root transformation . . . of the total frequency . . . of physical contacts during the play session. The transformation of frequency of physical contacts was necessary because of extreme skewness in the original frequency counts . . ."[90]

A separate analysis of variance was performed on each response measure. The findings can be summarized as follows: (1) there was no significant difference in the production of vocal and gestural responses between subjects of different linguistic skills, but (2) when children of different linguistic levels are combined in a dyad there are highly significant differences in interaction behavior. There are (a) the vocal and gestural responses of a "low" in play with a second "low" are as great in frequency as that of a "high" with a second "high," (b) the vocal and gestural responses of a "high" and "low" drop almost to zero, and (c) the combination of a "low" with a "low" produced the highest frequency of physical contact responses with the frequency of other combinations being low and essentially equal.

---

[90] Idem., p. 405.

These results strongly suggest the advantages of grouping mentally retarded children according to level of linguistic skill for optimal development of social interaction and the incidental learning that occurs as a result of such interaction.

There is an unpublished paper by Edson and Rolland where several language measures including the Parsons Language Sample, the Peabody Picture Vocabulary Test and the Illinois Test of Psycholinguistic Abilities were compared via correlation coefficients.[91] For the PLS and PPVT comparison fifty subjects were selected randomly from the population of the Parsons State Hospital and Training Center. One examiner administered all tests. The subjects received both tests within a period of one week and usually received both on the same day. The data derived from this testing was used to compute two coefficients; the PLS Verbal Section with the Total Score of the PPVT which yielded .865 as the coefficient of correlation and PLS Total Score with PPVT total which yielded .869 as the coefficient of correlation.

The procedure for comparing the PLS with the ITPA was less well controlled. Four different examiners tested forty different children with not all subjects even receiving both tests from the same examiner and with sometimes a month or more interval between tests.[92] The ITPA total score and PLS total score were found to correlate .852. The Vocal Section of the PLS

---

[91]Sandra K. Edson and J. C. Rolland, "Correlations Between Certain Tests in the Evaluation Battery," (Parsons Demonstration Project Report No. 26, March, 1965, Parsons State Hospital, Parsons, Kansas) (Mimeographed).

[92]Idem., p. 1.

correlated with the ITPA total .90. There was no computation of correlation between the Nonvocal section of the PLS and the ITPA total. Each subtest of the ITPA was correlated with each subtest of the PLS with the coefficients ranging from a positive .743 (PLS Intraverbal subtest with ITPA Auditory Decoding) to a negative .417 (PLS Intraverbal Gesture with ITPA Auditory Vocal Automatic).

There was no attempt to interpret the meaning of these correlations and it was pointed out "the distribution of the PLS scores proved to be skewed to the left, indicating the sample was loaded in favor of higher verbal level children" and the correlations were presented "keeping in mind the limitations of the raw data and thus the interpretations which can be made . . ."[93]

The only published account of a comparison of the ITPA and the PLS that could be found in the literature was made by Spradlin who stated although

> both (are) based on learning models . . . there are considerable differences between the rationales for the two tests. The rationale for the ITPA assumes that the test items are measuring implicit processes within the person and that the language responses are merely effects of these processes. Since the implicit processes are not observable independently of the language responses, the rationale involves problems of dualism and circularity. The authors often slip outside their 'theoretical' system as when they say, 'Encoding is the sum of those abilities required to express ideas in words or gestures.' In this statement the term 'idea' is extraneous to the system. Moreover while the authors claim to be operational

---

[93] Idem., p. 2.

in their definitions of constructs, such constructs as 'representational level' and 'association' cannot be reduced to operations. In contrast, the PLS is based on a descriptive model which assumes only that observable language responses are being sampled and evaluated in various situations. Its importance as a measure does not rest on an inference concerning implied internal processes but on an empirical relationship between scores based on language sampled in a test situation and language external to the test situation.[94]

---

[94]Spradlin, (1963b) op. cit., p. 522.

CHAPTER III

METHOD

Purpose and Hypotheses. This research was designed to present a comparison of two methods of assessing language development in institutionalized retardates. The first and second steps in this comparison were to examine and present the rationale or learning model underlying each test then presenting the actual test allowing a comparison between the test itself and the model upon which it was based. This step also allows a comparison between the methods employed in the construction and empirical evaluation of the two tests. The third step was reviewing the reliability data of the two instruments used in this assessment to see if an experimental comparison of the two instruments could be justified. The fourth step compared the two tests experimentally on a sample of a certain population to determine the degree of relationship between the various subtests and scales of the two instruments.

The questions posed in the fourth step were approached through a null hypothesis stated as follows: The linguistic functioning of institutionalized retardates as measured by the Illinois Test of Psycholinguistic Abilities and the Parsons Language Sample has no factors present common to both instruments. The testing of this hypothesis was accomplished by correlating all of the raw scores of each subtest with each of the remaining subtests of the two instruments and the interpretation of the resulting correlation matrix through factor analysis. This was

done to isolate any factors common to the two tests thus determining if both tests are providing the same information concerning linguistic functioning or if they are providing different information.

Experimental Design. The experimental design used in the fourth step of this research was the "counterbalancing" technique; a design to balance the effects when all subjects are to receive two experimental treatments and there is the possibility of transfer between them.[1] The subjects were divided in half with the first half receiving treatment one first and the second half receiving treatment two first. The experimental treatments (administration of the ITPA and PLS) were counterbalanced to equally distribute any transfer that might have occurred between them.

The testing occurred over a nine-week period and was done entirely by the author of this research. The testing started October 20, 1965, and ended December 16, 1965. Six children were scheduled to be tested on Wednesday and again on Friday of each week for eight weeks. The testing was scheduled at ten, eleven, one, two, three and four o'clock. The addition of a Tuesday and Thursday session was necessary for the rescheduling of those who failed for various reasons to appear for their first regularly scheduled session. Those who failed to appear for their second session after having appeared for

---

[1]F. J. McGuigan, Experimental Psychology--A Methodological Approach, (Englewood Cliffs: Prentice-Hall, Inc., 1960), pp. 113-115.

their first were to be dropped from the study and another subject chosen at random from their age and sex population. This was necessary to keep the one day interval between the tests constant. Fortunately, this occurred in only one instance where during the one day interval a subject was confined to the hospital and could not appear for the second day of testing. Another subject was picked at random from her age group to replace her in the study. An additional reason for the rescheduling period was to replace those who failed to respond to the tests sufficiently to score one above the minimum. This occurred in four cases, two males and two females, each time a replacement subject was picked from the appropriate age and sex group and placed on the testing schedule.

At the start of the first day of testing for a group of subjects a coin was tossed to determine whether the ITPA or PLS would be the first test administered. After the first test was decided in this manner the tests were given alternately throughout the remainder of the day. After a one day interval as an additional control for transfer and fatigue the testing sequence was switched with each subject receiving his second test. Those who had taken the PLS on the first day now were administered the ITPA and those who were administered the ITPA on the first day now received the PLS. This procedure was repeated beginning at the start of each of the days that marked the first testing for a particular group of subjects.

Sample. The sample consisted of fifty mentally retarded children who were patients of the Parsons (Kansas) State Hospital

and Training Center, a training oriented institution for mentally retarded children between the ages of six and twenty-one who are unable to be educated in the public schools. The sex, chronological age, length of institutionalization, adaptive behavior level and measured intelligence level of each subject is presented in Table I. The population of the hospital at the time of the study was 627 children, 363 males and 264 females. The program at the Parsons Hospital and Training Center is one of the most progressive in the nation providing medical, psychiatric, psychological, speech and hearing, and social work services as well as adjunctive music therapy, occupational therapy, recreational therapy, religious education, special education, and vocational training. Resident of the hospital live in cottages supervised by psychiatric aides. Placement in a particular cottage is determined by sex, age, and adaptive behavior level.

The procedure for selecting these fifty subjects was first to enumerate from hospital records all children whose age was either ten, eleven, twelve, thirteen, or fourteen by the start of the study; divide each age range into male and female groups; and randomly select from each of these ten groups five subjects for participation in the study. This provided a stratified random sample composed of an equal number of males and females from each of the five age levels that constitute the middle range of the ages of children included in the PLS standardization group. The sample was stratified in order to produce a gain in the degree the sample is characteristic of the population of the

TABLE I

SAMPLE DATA--MALES

Subj. No.	Chron. Age*	Length of Institution-alization*	Adaptive Behavior Level	Measured Intell. Level**
1	10- 4	0- 9	-4	-2
2	10- 4	0- 6	-1	-1
3	10- 5	0- 9	-3	-3
4	10- 7	3- 6	-2	-2
5	10-11	1- 1	-3	-3
6	11- 3	3- 4	-4	-3
7	11- 4	4- 3	-3	-4
8	11- 4	0- 6	-2	-1
9	11-10	3- 3	-3	-3
10	11-11	3- 7	-3	-3
11	12- 1	2- 8	-2	-1
12	12- 2	5- 4	-3	-3
13	12- 9	3- 1	-4	-5
14	12-11	3- 7	-3	-2
15	12-11	2-11	-2	-2
16	13-10	0- 6	-3	-3
17	13-11	4- 3	-3	-3
18	13-11	5-11	-3	-3
19	13-11	1- 7	-2	-2
20	13-11	0- 8	-1	-1
21	14- 4	8- 3	-3	-3
22	14- 5	0- 7	-4	-5
23	14- 6	5- 2	-2	-2
24	14-10	5-11	-3	-4
25	14-11	2- 3	-2	-2

* In years and months.

** From hospital records, approximate IQ range for each level is -1, 70-84; -2, 55-64; -3, 40-54; -4, 25-39; and -5, 24 and below.

TABLE I (continued)

SAMPLE DATA--FEMALES

Subj. No.	Chron. Age*	Length of Institution-alization*	Adaptive Behavior Level	Measured Intell. Level**
26	10- 2	1- 0	-2	-3
27	10- 5	0-11	-3	-3
28	10- 8	0- 9	-3	-3
29	10- 9	0- 9	-3	-3
30	10-11	4-11	-3	-3
31	11- 7	1-10	-3	-3
32	11- 7	0- 7	-4	-5
33	11- 9	3- 1	-3	-3
34	11-10	3- 5	-2	-2
35	11-11	3- 9	-3	-3
36	12- 2	4- 3	-3	-3
37	12- 2	2- 2	-2	-2
38	12- 6	2- 0	-3	-3
39	12-10	1- 5	-3	-3
40	12-10	1- 5	-3	-4
41	13- 4	2- 2	-3	-3
42	13- 7	6- 9	-2	-3
43	13- 7	5- 1	-3	-3
44	13- 9	7- 3	-1	-1
45	13- 9	4- 8	-3	-4
46	14- 3	1-10	-2	-2
47	14- 7	0-10	-3	-2
48	14-10	6- 2	-3	-3
49	14- 9	3- 4	-2	-2
50	14-10	2- 4	-2	-1

* In years and months.

** From hospital records, approximate IQ range for each level is -1, 70-84; -2, 55-69; -3, 40-54; -4, 25-39; and -5, 24 and below.

hospital.[2] It is possible through selecting a stratified random sample to include an equal representation of age and sex which would not be possible through unrestricted random sampling. Drawing the sample to be in any way representative of the ITPA standardization group was possible only in terms of chronological age as this test was standardized on normals rather than on those classified as mentally retarded. This was the basis for deciding to have the sample represent the middle range of the PLS reference group.

Procedure and circumstances of administration. The room in which the tests were administered was an office in the basement level of a new building housing the special education and clinical psychology departments of the hospital and the national adaptive behavior project. It was a large well illuminated office equipped with attractive furniture of contemporary design. The names of the children selected for the sample were sent to the coordination center and were scheduled for the appointment time that conflicted least with their regular daily schedule. They reported to the testing room at the proper time, were helped off with their coats and seated at the desk by the examiner. A simple conversation followed in which they were asked their name, the name of the cottage where they lived, and what activities they had experienced so far that day. The tests were administered in the exact manner called for in the directions

---

[2]For a discussion of the gains possible through stratified random sampling see W. G. Cochran, Sampling Techniques, (New York: John Wiley & Sons, Inc., 1963, pp. 87-88 and 98-102.

of the Examiners Manual[3] of the ITPA and the test booklet of the PLS.[4] The only breaks in testing occurred at the subject's request. This was usually a request for a drink of water or a trip to the bathroom. The time of administration for the PLS ranged from twenty to forty minutes with the ITPA ranging from thirty-five to sixty minutes. At the end of the testing the child was helped to put on his coat, thanked for being a good boy or girl, and told that the examiner would soon see him again.

Controls. The control in this study was provided by the counter-balancing design which distributes any transfer equally between the two tests. In addition to the counter-balancing design a one day rest period was provided which separated the testing periods by forty-eight hours. This was to minimize the transfer and fatigue that might have occurred if both instruments had been administered to each subject at the same testing session. The advantage of the counter-balancing technique is that it allows each subject to receive both experimental treatments thus eliminating subject variables.

---

[3]McCarthy and Kirk (1961), op. cit., pp. 31-59.

[4]Edson and Rolland (1964b), op. cit., pp. 1-15.

CHAPTER IV

RESULTS AND DISCUSSION

This study was designed to compare the Illinois Test of Psycholinguistic Abilities and the Parsons Language Sample in assessing the linguistic functioning of institutionalized mentally retarded children. The data obtained was analyzed by the Pearson product-moment correlation coefficient[1] and a principle-components factor analysis with a varimax rotation of the first five factors to redistribute the variance.[2]

Linearity and homoscedasticity of the data. Many authors have presented criteria which the distribution of data must meet in order to justify the use of the product-moment correlation coefficient. The two main requirements listed by these authors are: (1) that the relationship between the variables be rectilinear, and (2) that the two distributions be similar or more technically demonstrate homoscedasticity. The distributions

---

[1] Q. McNemar, Psychological Statistics, (New York: John Wiley and Sons, Inc., 1949), pp. 94-96.

[2] The factor analysis used in this research was a principle-components (BMD03M) factor analysis version of September 23, 1965, performed at the Health Science Computing Facility of the University of California at Los Angeles. For a general presentation of the factor analytic method refer to C. J. Adcock, Factorial Analysis for Non-mathematicians, (Carlton, Victoria: Melbourne University Press, 1954). For a detailed presentation of the primciple-components method refer to B. Fruchter, Introduction to Factor Analysis, (New York: Van Nostrand, 1954), pp. 99-104; and G. H. Thompson, The Factorial Analysis of Human Ability, (Boston: Houghton Mifflin, 1950), pp. 66-80.

need not be normally distributed (provided the deviation is not excessive) if these two requirements are met.[3]

The assumptions underlying the Pearson product-moment correlation coefficient as mentioned previously were tested to determine the extent the data of this research meets these assumptions. A comparison of the second, third and fourth moments for skewness and kurtosis was performed. This revealed the data (total scores) of both the ITPA and PLS was slightly negatively skewed (for the ITPA g_1 equaled -.254 and for the PLS g_1 equaled -.692) and kurtosis was platykurtic (for the ITPA g_2 equaled -1.155 and for the PLS g_2 equaled -.406).[4] Table II shows the distribution of the ITPA and the PLS total scores revealing the negative skewness of the data for both the ITPA and PLS. Table III shows a scattergram of the ITPA and PLS total scores for each of the fifty subjects which from a visual analysis shows a rectilinear relationship. Tables IV and V show the means and standard deviations of the raw scores of each subtest, scale and total score of both the ITPA and PLS. The fact that both distributions are slightly negatively skewed and essentially platykurtic leads to the conclusion that the data is homoscedastic, or has homogeneity of variance.

---

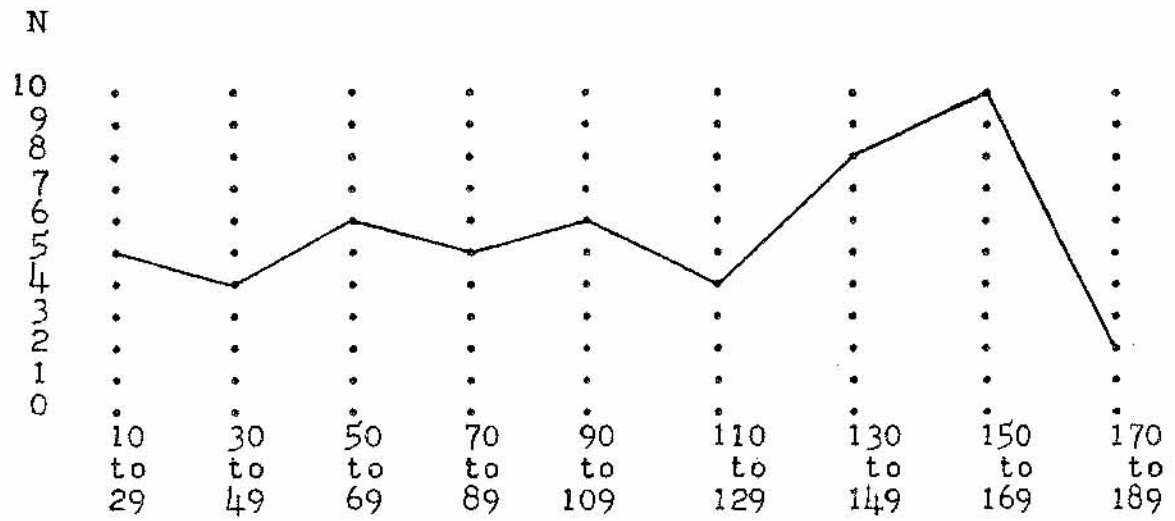
[3]G. A. Ferguson, Statistical Analysis in Psychology and Education, (New York: McGraw-Hill, Inc., 1959), pp. 109-110; J. P. Guilford, Fundamental Statistics in Psychology and Education, (New York: McGraw-Hill, Inc., 1942), pp. 223-224; and McNemar, op. cit., pp. 113-117.

[4]Ferguson, op. cit., p. 64f.

TABLE II

DISTRIBUTION OF THE ITPA AND PLS TOTAL SCORES

ITPA



PLS

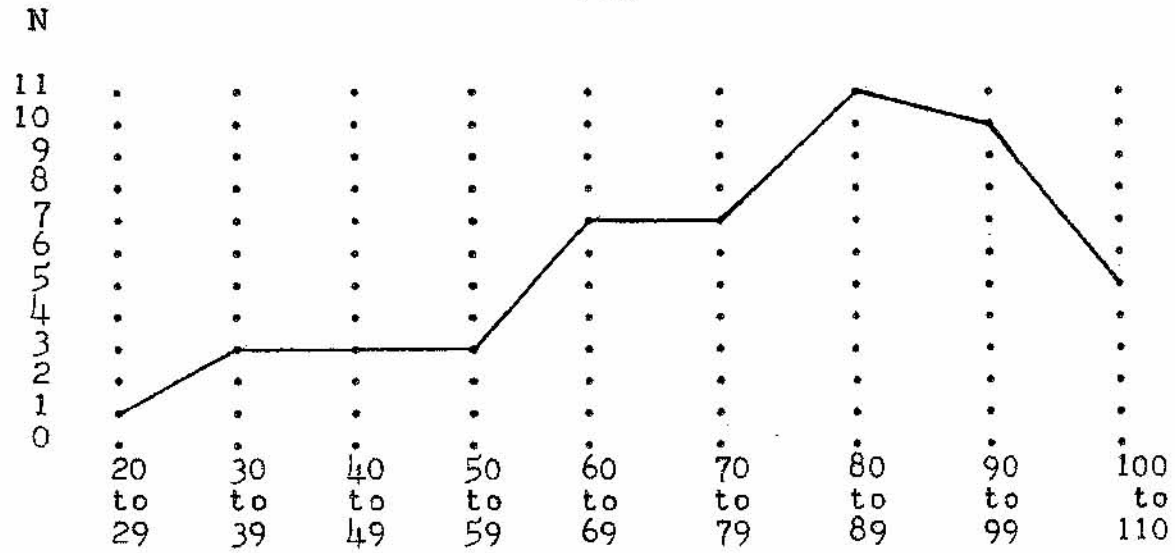
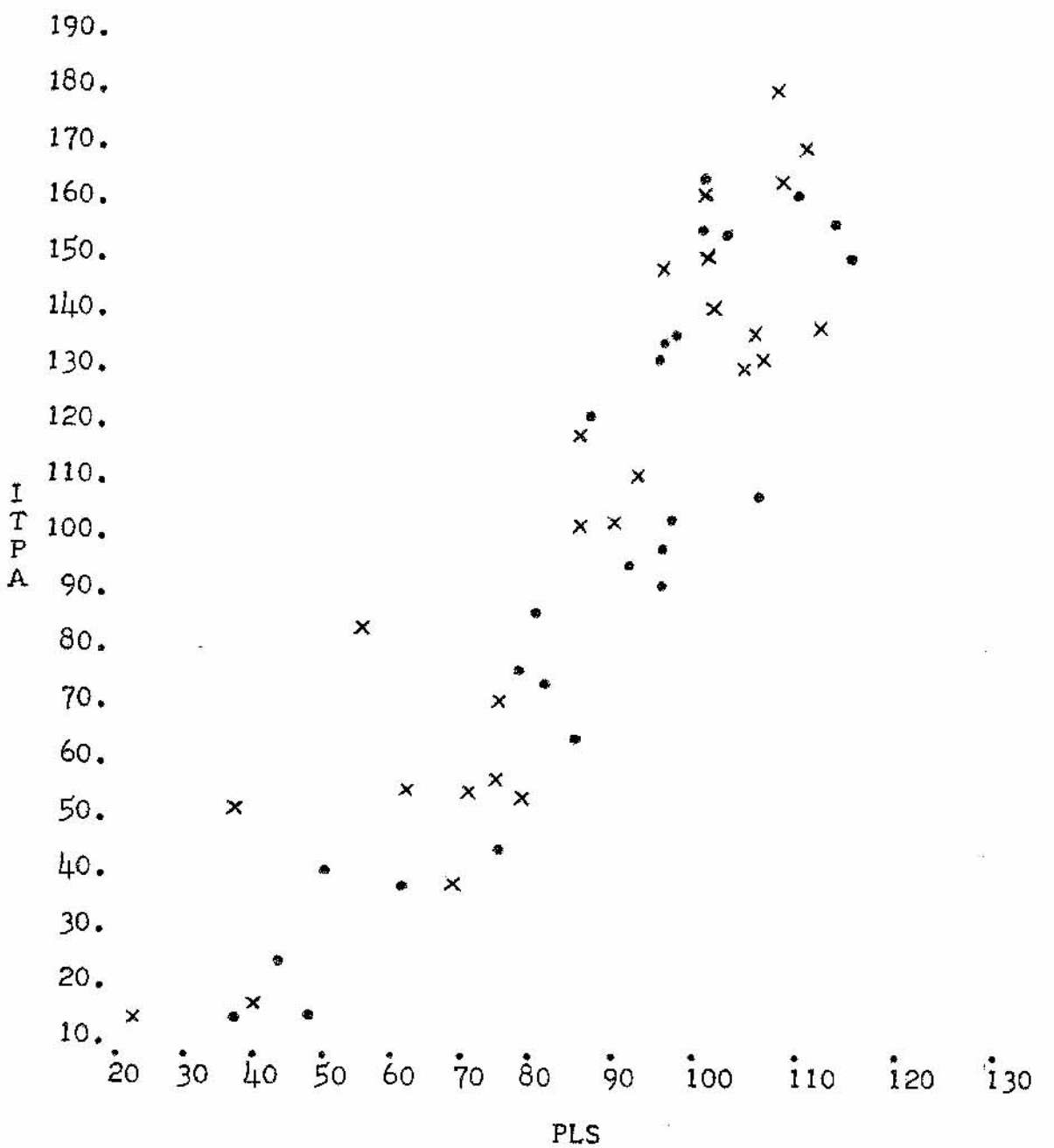


TABLE III

SCATTERGRAM OF THE DISTRIBUTION OF THE ITPA AND PLS TOTAL SCORES OF EACH SUBJECT*



* Dots denote female subjects, x's denote male subjects.

TABLE IV

RAW SCORE MEANS AND STANDARD DEVIATIONS OF THE ITPA SUBTESTS AND TOTAL SCORES

	Mean	SD	N
Primarily Verbal			
Auditory Vocal Automatic	6.36	7.0240	50
Auditory Vocal Association	11.78	7.4294	50
Vocal Encoding	9.34	5.3590	50
Auditory Vocal Sequencing	12.40	7.6692	50
Primarily Verbal Total	40.38	24.2293	50
Primarily Nonverbal			
Visual Decoding	11.58	5.8067	50
Motor Encoding	10.96	5.4134	50
Visual Motor Sequencing	9.08	4.8691	50
Visual Motor Association	14.22	6.6587	50
Auditory Decoding	16.46	10.1141	50
Primarily Nonverbal Total	61.70	28.2411	50
ITPA Total	102.48	50.1578	50

TABLE V

RAW SCORE MEANS AND STANDARD DEVIATIONS OF THE PLS SUBTESTS AND TOTAL SCORES

	Mean	SD	N
Vocal Scale			
Tact	19.30	5.0639	50
Echoic	12.44	4.1559	50
Intraverbal	13.84	8.9633	50
Vocal Total	44.98	16.2298	50
Nonvocal Scale			
Echoic Gesture	8.34	1.6734	50
Comprehension	13.36	2.9190	50
Intraverbal Gesture	9.06	6.3709	50
Nonvocal Total	30.76	7.9323	50
PLS Total	75.74	20.9088	50

The data of each subtest of the ITPA and PLS was plotted to determine the extent of deviation from the slight negative skewness and flattening of the curve demonstrated by the distribution of the total scores. A visual analysis revealed only the Auditory-Vocal Automatic subtest of the ITPA and Intraverbal Gesture subtest of the PLS were different. These two subtests both had positive skewness and leptokurtic kurtosis thus the correlations of these measures with the others may not reflect the degree of accuracy present in the remaining comparisons. Peters and VanVoorhis discuss the quesion of the validity of a relationship established through comparison of a positively and negatively skewed distribution.[5] They state the Pearson product-moment formula and the Spearman formula (which has as its only criteria for use that both variables be measured in at least an ordinal scale of measurement and thus can be compared regardless of the shape of the distribution)[6] "are merely different algebraic forms of the same thing, so that, whenever the one is applicable, the other is also applicable."[7] The Pearson coefficients are accepted as more often providing the truer relationship because of the loss of accuracy in converting to ranks using Spearman's method. The results of using either method are "always well within the probable error of the correlation coefficient."[8]

---

[5]C. Peters and W. R. VanVoorhis, Statistical Procedures and Their Mathematical Bases, (New York: McGraw-Hill, Inc., 1940).

[6]S. Sigel, Nonparametric Statistics for the Behavioral Sciences, (New York: McGraw-Hill, Inc., 1956), p. 202.

[7]Peters and VanVoorhis, op. cit., p. 109.

[8]Ibid.

It should be pointed out that moderate departures from homogeneity of variance will not seriously affect the inferences that can be drawn from the data.[9] With most sets of data the assumptions underlying the mathematical model are only roughly satisfied, the raw data from experimentation rarely falls into the proper theoretical form as exactly as would be desired. The characteristics of the data demonstrated above show the Pearson product-moment correlation coefficient is the appropriate statistic for this data (with two exceptions if one demands more criteria for its use than do Peters and VanVoorhis).

Results of the intercorrelations. The results of the intercorrelations of each subtest and scale with each of the other subtests and scales of the ITPA are presented in Table VI. These intercorrelations were computed to compare the extent of interrelatedness of the data from this research with the intercorrelations derived from the standardization data of the ITPA. Table VI presents the intercorrelations of this research from asterisks down to the lower left corner and the intercorrelations from the standardization data[10] from the asterisks up to the upper right corner. The intercorrelations may be compared by starting at the asterisks for a particular subtest or scale and reading to the right or up for the intercorrelations from the standardization data and to the left or down for the intercorrelations of this research.

---

[9]Guilford, op. cit., p. 150.

[10]McCarthy and Kirk (1963), op. cit., p. 42.

TABLE VI

THE ILLINOIS TEST OF PSYCHOLINGUISTIC ABILITIES
INTERCORRELATIONS

	Primarily Verbal					Primarily Nonverbal						
	1	2	3	4	5	6	7	8	9	10	11	12
Primarily Verbal												
1. Auditory Vocal Automatic	***	.89	.78	.80		.78	.74	.80	.79	.82		.90
2. Auditory Vocal Association	.74	***	.79	.84		.83	.78	.85	.84	.87		.95
3. Vocal Encoding	.59	.72	***	.68		.71	.75	.75	.72	.73		.82
4. Auditory Vocal Sequencing	.55	.67	.51	***		.70	.70	.76	.73	.76		.83
5. Verbal Total	.85	.92	.79	.81	***							
Primarily Nonverbal												
6. Visual Decoding	.53	.76	.60	.48	.70	***	.74	.78	.74	.77		.84
7. Motor Encoding	.40	.61	.69	.33	.58	.76	***	.73	.72	.73		.82
8. Visual Motor Sequencing	.66	.76	.61	.62	.78	.76	.66	***	.80	.79		.88
9. Visual Motor Association	.56	.72	.63	.45	.69	.76	.56	.74	***	.77		.85
10. Auditory Decoding	.67	.86	.67	.62	.82	.73	.61	.80	.71	***		.87
11. Nonverbal Total	.64	.85	.73	.57	.80	.90	.81	.89	.85	.90	***	.87
12. ITPA Total	.77	.93	.80	.71	.94	.85	.74	.88	.81	.91	.96	***

The gaps in the table comprise correlations that were not computed for the original data. For the purpose of this research the ITPA was divided into primarily verbal and primarily nonverbal sections in much the same manner as the PLS is divided into vocal and nonvocal sections. This was done mainly to determine if there is perhaps some merit in classifying the tests of the ITPA in this manner. This was not attempted by the authors of the ITPA thus no correlations of this type are available for the standardization population. The tests considered primarily verbal in nature are Auditory-Vocal Automatic, Auditory Vocal Association, Vocal Encoding and Auditory Vocal Sequencing. The tests considered primarily nonverbal in nature are Visual Decoding, Motor Encoding, Visual Motor Sequencing, Visual Motor Association and Auditory Decoding. The type of response required by the subject served as the basis for the verbal or nonverbal classification. If a test could be successfully completed without a verbal response (even though a subject might respond verbally if he chooses) the test was considered primarily nonverbal. If the test could not be successfully completed without a verbal response the test was considered primarily verbal.

The ITPA intercorrelations were generally smaller for this research than for the standardization group with the mean variance being just over 2% (.021). The exception to this was the Visual Decoding subtest in which the intercorrelations were nearly equivalent to those of the standardization group. Thirty-eight intercorrelations were lower than those of the standardization group ranging from .37 lower to .01 lower with a mean variance

of 2½% (.0250). Six intercorrelations of this research were higher than those of the standardization data ranging from .01 higher to .04 higher with the mean variance being less than 1% (.00058). One ITPA intercorrelation was exactly the same in both studies.

The formula for determining the significance of difference between two correlation coefficients for independent samples revealed the difference between the ITPA intercorrelations presented in these two studies must be .285 or greater to be significantly different at the .05 level of confidence.[11] Only two intercorrelations exceed this value thus the two sets of data would not be considered significantly different. Table VII shows the extent of difference between the two sets of data and the level of confidence at which they would be considered significantly different.

The two intercorrelations that were significantly different from those obtained from the original data were the intercorrelations of Motor Encoding with Auditory-Vocal Automatic and Motor Encoding with Auditory-Vocal Sequencing. The Motor Encoding test is evidently somewhat unreliable. This is supported by additional data from the original standardization which found the Motor Encoding subtest to have the lowest internal consistency and test-retest stability.[12]

---

[11] Ferguson, op. cit., pp. 153-154.

[12] McCarthy and Kirk (1963), op. cit., pp. 28-33.

TABLE VII

EXTENT OF DIFFERENCE* BETWEEN ITPA INTERCORRELATIONS OF THIS RESEARCH AND INTERCORRELATIONS DERIVED FROM THE STANDARDIZATION DATA

	1	2	3	4	5	6	7	8	9	10
Primarily Verbal										
1. Auditory Vocal Automatic	***									
2. Auditory Vocal Association	-.15	***								
3. Vocal Encoding	-.19	-.07	***							
4. Auditory Vocal Sequencing	-.25	-.17	-.17	***						
Primarily Nonverbal										
5. Visual Decoding	-.25	-.07	-.11	-.22	***					
6. Motor Encoding	-.34[a]	-.17	-.06	-.37[a]	.02	***				
7. Visual Motor Sequencing	-.14	-.09	-.14	-.14	-.02	-.07	***			
8. Visual Motor Association	-.23	-.12	-.09	-.28	.02	-.16	-.06	***		
9. Auditory Decoding	-.15	-.01	-.06	-.14	-.04	-.12	.01	-.06	***	
10. ITPA Total	-.13	-.02	-.02	-.12	.01	.03	.00	-.04	.04	***

* A negative sign preceding the value indicates the value is that much below the intercorrelation derived from the standardization data, absence of sign indicates the value exceeds that of the standardization data.

[a]Significantly different at the .05 level.

This same procedure was used to compare the intercorrelations obtained from the Parsons Language Sample data of this research with the intercorrelations presented by Spradlin in the original monograph.[13] Table VIII presents the two sets of intercorrelations. The intercorrelations obtained from this research are presented from the asterisks down to the lower left corner and the intercorrelations obtained from Spradlin's original data from the asterisks up to the upper right corner. The intercorrelations may be compared by starting at the asterisks for a particular subtest or scale and reading to the right or up for the intercorrelations presented by Spradlin and to the left or down for the intercorrelations as determined by this research.

The PLS intercorrelations were generally smaller for this research than for the standardization group. The exception to this is Intraverbal Gesture subtest with each of the other variables. The intercorrelations were larger in this research but not significantly larger. Twenty-eight intercorrelations of this research were lower than those of the standardization group ranging from .49 lower to .02 lower with the mean variance being approximately 3% (.0341). Seven intercorrelations of this research were higher than those of the standardization data ranging from .01 higher to .20 higher with the mean variance being just over 1% (.0103). One PLS intercorrelation was exactly the same in both studies. The extent of variance for all the

---

[13]Spradlin (1963a), op. cit., p. 24.

TABLE VIII

PARSONS LANGUAGE SAMPLE
INTERCORRELATIONS

	1	2	3	4	5	6	7	8	9
Vocal Scale									
1. Tact	***	.77	.82	.93	.64	.79	.12	.67	.91
2. Echoic	.69	***	.81	.93	.63	.69	.01	.56	.85
3. Intraverbal	.78	.79	***	.94	.65	.74	.02	.58	.88
4. Vocal Total	.88	.88	.95	***	.69	.80	.04	.64	.94
Nonvocal Scale									
5. Echoic Gesture	.39	.56	.48	.49	***	.62	.16	.75	.78
6. Comprehension	.57	.56	.69	.67	.34	***	.32	.85	.90
7. Intraverbal Gesture	.10	.09	.08	.10	.09	.26	***	.68	.32
8. Nonvocal Total	.37	.39	.42	.43	.26	.65	.88	***	.86
9. PLS Total	.82	.83	.90	.94	.48	.76	.41	.71	***

PLS intercorrelations of this research from those of the standardization data averaged approximately 3% lower (.0292).

The formula for determining the significance of difference between two correlation coefficients for independent samples revealed the difference between the PLS intercorrelations presented in these two studies must be .30 or greater to be significantly different at the .05 level of confidence.[14] Only three intercorrelations exceed this value thus the two sets of data would not be considered significantly different. Table IX shows the extent of difference between the two sets of data and the level of confidence at which they would be considered significantly different.

Two of the three intercorrelations that were significantly different from those obtained from the original data were the intercorrelations of Echoic Gesture with Nonvocal total and Echoic Gesture with PLS total, thus the Echoic Gesture subtest appears somewhat unreliable. This is supported by the test-retest data presented in the original monograph which showed a significant difference at the .05 level when the retest was administered by a different examiner.[15] The split-half reliability coefficient for Echoic Gesture was also below the customary .90[16] being .86[17] which indicates the possibility of

---

[14]Ferguson, op. cit., pp. 153-154.

[15]Spradlin (1963a), op. cit., p. 20.

[16]E. R. Hilgard, Introduction to Psychology, (New York: Harcourt, Brace and World, Inc., 1962), p. 397.

[17]Spradlin (1963a), op. cit., p. 19.

TABLE IX

EXTENT OF DIFFERENCE* BETWEEN PLS INTERCORRELATIONS THIS RESEARCH AND INTERCORRELATIONS DERIVED FROM THE STANDARDIZATION DATA

	1	2	3	4	5	6	7	8	9
Vocal Scale									
1. Tact	***								
2. Echoic	-.08	***							
3. Intraverbal	-.04	-.02	***						
4. Vocal Total	-.05	-.05	.01	***					
Nonvocal Scale									
5. Echoic Gesture	-.25	-.07	-.17	-.20	***				
6. Comprehension	-.22	-.13	-.05	-.13	-.28	***			
7. Intraverbal Gesture	-.02	.08	.06	.06	-.25	-.06	***		
8. Nonvocal Total	-.30[a]	-.17	-.16	-.21	-.49[a]	-.20	.20	***	
9. PLS Total	-.09	-.02	.02	.00	-.30[a]	-.14	.09	-.15	***

* A negative sign preceding the value indicates the value is that much below the intercorrelation derived from the standardization data, absence of sign indicates the value exceeds that of the standardization data.

[a] Significantly different at the .05 level.

variance due to less than the desired internal consistency. Echoic Gesture also had the lowest test-retest reliability coefficients of the six tests for children nine to twelve years of age retested after twenty-nine months[18] and on male new admissions retested after fifteen months.[19]

The only other significant difference, Tact with the Nonvocal total, is possibly due to either the sample of this research varying somewhat from the overall characteristics of the standardization group or the difference being one of those five times in a hundred that could be attributed to external variables.

Correlation of the ITPA and PLS subtests, scales and totals. Table X presents the correlations of each subtest, scale and total score of the PLS with each of the subtests, primarily verbal and nonverbal subtotals, and total score of the ITPA. The correlations computed from the data of this research are presented in the lower row of each group of figures. The upper figures are the ITPA and PLS correlations computed from the data gathered in the research of Edson and Rolland.[20] These two sets of data were compared using the formula for determining the significance of difference between two correlation coefficients for independent samples.[21] The difference between the correlations

---

[18] Idem., pp. 21-22.

[19] Idem., pp. 22-23.

[20] Edson and Rolland (1965), op. cit., p. 5.

[21] Ferguson, op. cit., pp. 153-154.

TABLE X

COMPARISON OF ITPA AND PLS CORRELATIONS OF TWO STUDIES

	Vocal Scale				Nonvocal Scale				
	Tact	Echoic	Intraverbal	Vocal Total	Echoic Gesture	Comprehension	Intraverbal Gesture	Nonvocal Total	PLS Total
Primarily Verbal									
Auditory Vocal Automatic	.59[a]	.64	.63		.52	.25	÷42		
	.66	.61	.74	.72	.40	.37	÷13	.12	.61
Auditory Vocal Association	.71	.70	.84		.70	.46	÷29		
	.76	.79	.94	.91	.45	.73	.07	.42	.87
Vocal Encoding	.62	.38	.74		.49	.47	.91		
	.64	.59	.74	.70	.38	.70	.10	.42	.70
Auditory Vocal Sequencing	.47	.86	.65		.70	.35	÷28		
	.50	.90	.67	.73	.61	.46	.05	.34	.70
Verbal Total	.74	.85	.89	.90	.53	.66	.01	.36	.83
Primarily Nonverbal									
Visual Decoding	.82	.43	.74		.40	.53	÷03		
	.72	.56	.76	.74	.45	.66	.24	.53	.78
Motor Encoding	.65	.33	.64		.34	.48	÷03		
	.56	.45	.64	.60	.34	.64	.35	.59	.69
Visual Motor Sequencing	.72	.51	.70		.60	.56	÷21		
	.69	.69	.78	.81	.55	.57	.13	.43	.79
Visual Motor Association	.65	.41	.67		.53	.58	÷16		
	.66	.56	.77	.74	.41	.64	.11	.41	.73
Auditory Decoding	.73	.51	.74		.54	.51	÷23		
	.76	.70	.89	.88	.43	.59	.05	.35	.81
Nonverbal Total	.78	.67	.89	.87	.49	.70	.20	.52	.87
ITPA Total	.81	.79	.94	.93	.53	.72	.14	.49	.91

[a]This data is that reported by Edson and Rolland, op. cit., p. 5, none of their values (presented in the first line) differ from those of this research (presented in the second line) at either the .01 or .05 levels of confidence.

presented in these two studies must be .515 or greater to be significantly different at the .01 level of confidence and .405 or greater to be significantly different at the .05 level of confidence. None of the correlations reported in the research of Edson and Rolland differ from those of this research at either the .01 or .05 level thus the correlations derived from the two sets of data are statistically equivalent.

A second statistical analysis was performed on the correlations obtained from this research to see if they significantly differ from zero, or if some or all of the correlations could be this large due to variables external to this study. With at least 45 degrees of freedom a correlation coefficient must be at least .288 to be significantly different from zero at the .05 level and must be at least .372 to be significantly different from zero at the .01 level of confidence.[22] The correlations that are not significantly different from zero at the .01 level of confidence are marked with an asterisk in Table XI.

Only 17 of the 108 correlations were not significant at the .01 level of confidence. Twelve of these were the correlations of the Intraverbal Gesture subtest of the PLS with each of the measures of the ITPA. Four additional nonsignificant correlations were the result of the Intraverbal Gesture subtest lowering the correlations between the Nonvocal total of the PLS and each of the measures of the ITPA. The Intraverbal Gesture subtest of the PLS did not correlate significantly different from

---

[22] Idem., pp. 152-153.

TABLE XI

ITPA AND PLS CORRELATIONS

	Vocal Scale				Nonvocal Scale				
	Tact	Echoic	Intraverbal	Vocal Total	Echoic Gesture	Comprehension	Intraverbal Gesture	Nonvocal Total	PLS Total
Primarily Verbal									
Auditory Vocal Automatic	.66	.61	.74	.72	.40	.37	.13*	.12*	.61
Auditory Vocal Association	.76	.79	.94	.91	.45	.73	.07*	.42	.87
Vocal Encoding	.64	.59	.74	.70	.38	.70	.10*	.42	.70
Auditory Vocal Sequencing	.50	.90	.67	.73	.61	.46	.05*	.34*	.70
Verbal Total	.74	.85	.89	.90	.53	.66	.01*	.36*	.83
Primarily Nonverbal									
Visual Decoding	.72	.56	.76	.74	.45	.66	.24	.53	.78
Motor Encoding	.56	.45	.64	.60	.34*	.64	.35*	.59	.69
Visual Motor Sequencing	.69	.69	.78	.81	.55	.57	.13*	.43	.79
Visual Motor Association	.66	.56	.77	.74	.41	.64	.11	.41	.73
Auditory Decoding	.76	.70	.89	.88	.43	.59	.05*	.35*	.81
Nonverbal Total	.78	.67	.89	.87	.49	.70	.20*	.52	.87
ITPA Total	.81	.79	.94	.93	.53	.72	.14*	.49	.91

* Not significantly different from zero at the .01 or .05 levels of confidence.

zero with any of the variables of the ITPA nor any of the other subtests or scales of the PLS. It has been demonstrated to have internal consistency, (although lower than other subtests of the PLS) test-retest stability, and examiner equivalence.[23] There are two possibilities that may account for the lack of correlation with the other variables. Either the Intraverbal Gesture subtest is measuring an aspect of language completely independent of the other variables or is not consistent in the aspect of language measured. Support of this latter possibility has been furnished by experience with the test. Only gestural responses to statements made by the examiner are scored positively; yet the subtest provides no directions to indicate to the subject a gestural response is expected. The nature of most of the statements are such that either a gestural or verbal response could be given. Some children use gestures while others give verbal answers with the choice being inconsistent with the scores on the other subtests. Thus while children may consistently give verbal or gestural responses to this subtest (thus demonstrating internal consistency and test-retest stability), the intercorrelations with other PLS subtests and correlations with other language measures do not significantly differ from zero.

The one remaining nonsignificant correlation was between the Motor Encoding subtest of the ITPA and the Echoic Gesture subtest of the PLS. Motor Encoding was the subtest that produced two intercorrelation coefficients in this research that were

---

[23] Spradlin (1963a), op. cit., pp. 19-20.

significantly different from the intercorrelations of the standardization data. The Motor Encoding subtest has been demonstrated to have less internal consistency and test-retest stability than the other ITPA subtests. The Echoic Gesture subtest produced two intercorrelations that were significantly different from the intercorrelations of the standardization data. This difference is attributed to less than adequate reliability as the test has been demonstrated to have lower internal consistency, test-retest stability and examiner equivalence than the other PLS subtests. Motor Encoding had lower correlation coefficients with the variables of the Parsons Language Sample than any other ITPA subtest. Echoic Gesture had lower correlation coefficients with the variables of the Illinois Test of Psycholinguistic Abilities than did any other PLS subtest except Intraverbal Gesture. The correlations of Motor Encoding and Echoic Gesture with each of the other variables of the ITPA and PLS are significantly different from zero. However, when compared with each other they do not significantly differ from a zero correlation.

The subtests of the Illinois Test of Psycholinguistic Abilities and the Parsons Language Sample show a high degree of interrelatedness with the exception of one subtest (Intraverbal Gesture) showing no significant relationship with any of the other variables and two other subtests (Motor Encoding and Echoic Gesture) that are significantly related to all the other variables but not to each other.

Tables XII and XIII show the correlations of the subtests, scales and total scores of the ITPA and PLS with the Adaptive Behavior levels,[24] chronological ages and sex of the subjects. Table XII reveals the ITPA is highly related to the adaptive behavior level of the subjects. The numbering of the levels was reversed so the higher the positive correlation the greater the relationship of the variable to adaptive capacity. Table XII also reveals the correlations of the variables of the ITPA with chronological age and sex did not significantly differ from zero. Thus neither age nor sex is related to a subject's performance on the ITPA.

A look at Table XIII shows the PLS also to be significantly related to Adaptive Behavior level with two exceptions; Intraverbal Gesture subtest, which has been demonstrated to be an inconsistent measure, and the Nonvocal total of which the Intraverbal Gesture subtest is a part reducing its relatedness to Adaptive Behavior. Table XIII also shows the Echoic Gesture subtest and the Intraverbal Gesture subtest to be significantly different from zero in their relationship to chronological age. Echoic Gesture is evidently an easier test for older subjects than younger ones. Intraverbal Gesture subtest scores were higher for younger subjects than older ones. This is possibly

---

[24] For a discussion on Adaptive Behavior and its assessment see H. Leland, "Conference on Measurement of Adaptive Behavior," (Rehabilitation Center, Parsons State Hospital, Parsons, Kansas, 1964) pp. 1-43, (Mimeographed); and H. Leland, "Some Thoughts on the Current Status of Adaptive Behavior," Mental Retardation, 1964, Vol. 2, No. 3, pp. 171-177.

TABLE XII

ITPA CORRELATIONS WITH ADAPTIVE BEHAVIOR LEVEL CHRONOLOGICAL AGE AND SEX

	Adaptive Behavior Level	Chronological Age	Sex
Primarily Verbal			
Auditory Vocal Automatic	.56	.34*	-.02*
Auditory Vocal Association	.74	.21*	-.04*
Vocal Encoding	.51	.15*	-.10*
Auditory Vocal Sequencing	.71	.06*	.13*
Verbal Total	.77	.22*	-.01*
Primarily Nonverbal			
Visual Decoding	.64	.27*	-.27*
Motor Encoding	.49	.14*	-.13*
Visual Motor Sequencing	.68	.13*	-.11*
Visual Motor Association	.53	.08*	-.15*
Auditory Decoding	.66	.24*	.03*
Nonverbal Total	.74	.21*	-.12*
ITPA Total	.74	.21*	-.06*

*Not significantly different from zero at the .01 or .05 levels of confidence.

TABLE XIII

PLS CORRELATIONS WITH ADAPTIVE BEHAVIOR LEVEL CHRONOLOGICAL AGE AND SEX

	Adaptive Behavior Level	Chronological Age	Sex
Vocal Scale			
Tact	.65	.22*	-.10*
Echoic	.73	.13*	.04*
Intraverbal	.53	.15*	.10*
Vocal Total	.73	.22*	.02*
Nonvocal Scale			
Echoic Gesture	.69	.29	.01*
Comprehension	.61	.20*	.21*
Intraverbal Gesture	.06*	-.33	-.03*
Nonvocal Total	.67	.11*	.04*
PLS Total	.67	.11*	.04*

* Not significantly different from zero at the .01 or .05 levels of confidence.

due to the fact that subjects who communicate with gestures more than words are likely to have greater success on the Intraverbal Gesture subtest. Younger retarded children evidently communicate more with gestures than with words and when asked a question that can be answered with either gestures or words choose the former. As a result they score higher than the older subjects who answer more with words than with gestures. None of the correlations of the variables of the PLS with the sex of the subjects were significantly different from zero at the .01 or .05 levels of confidence. Thus the sex of the subject in this sample is not related to performance on the PLS.

Factor Analysis of the Correlation Matrix. The overall correlation matrix was subjected to a principle-components factor analysis and the factor matrix derived from this process is presented in Table XIV. The first five factors derived from this factor matrix were rotated through nineteen iteration cycles to maximize the amount of variance extracted. The rotated factor matrix is presented in Table XV. These five factors were found to account for 85% of the variance as an average for the fifteen variables. The percentage of variance extracted by each factor for each variable is presented in Table XVI.

Table XVII presents a description of the factors isolated through this research. Factor I (labeled Immediate Recall of Auditory Symbols) accounted for 17% of the variance. Two subtests (ITPA Auditory-Vocal Sequencing and the PLS Echoic) received their highest loadings on this factor. The factor

TABLE XIV

MATRIX OF EXTRACTED FACTOR LOADINGS
FOR ITPA AND PLS SUBTESTS

	Factors				
	I	II	III	IV	V
Auditory Vocal Automatic	.73440	.35685	-.26382	-.15722	.15123
Visual Decoding	.84496	-.25879	-.07045	.20868	.17915
Motor Encoding	.73415	-.46563	-.04075	.22879	-.03616
Auditory Vocal Association	.93384	.05443	-.08706	-.15196	-.06049
Visual Motor Sequencing	.87382	.01177	.01005	.14410	.25374
Vocal Encoding	.79974	-.12280	-.13415	-.01552	-.39420
Auditory Vocal Sequential	.73326	.36018	.48124	-.17077	-.06648
Visual Motor Association	.81368	-.11214	-.19506	.14817	.08965
Auditory Decoding	.88975	.05509	-.13679	-.08939	.11696
Tact	.83256	-.00333	-.15437	-.12921	.15361
Echoic	.83124	.26023	.33916	-.24209	-.04282
Intraverbal	.94627	.04674	-.08765	-.11688	-.00647
Echoic Gesture	.57685	.36191	.40655	.55403	-.00361
Comprehension	.75531	-.32429	.02503	.02084	-.43610
Intraverbal Gesture	.13937	-.77445	.47295	-.23067	.22358

TABLE XV

ITPA AND PLS ROTATED FACTOR MATRIX

	Factors				
	I	II	III	IV	V
Auditory Vocal Automatic	.40276	-.72860	.01516	.27958	-.11141
Visual Decoding	.07935	-.72542	.32062	-.24852	-.40552
Motor Encoding	-.02840	-.51881	.26287	-.33912	-.59752
Auditory Vocal Association	.47521	-.67965	.06898	-.01236	-.46550
Visual Motor Sequencing	.28507	-.74986	.36777	-.12146	-.23533
Vocal Encoding	.28667	-.42468	.06244	.04483	-.74823
Auditory Vocal Sequential	.85416	-.27384	.31238	-.03716	-.17039
Visual Motor Association	.10942	-.71359	.22586	-.04941	-.40993
Auditory Decoding	.37059	-.76283	.10617	-.00630	-.32328
Tact	.32381	-.75051	.04407	-.05674	-.28938
Echoic	.80505	-.42378	.19760	-.06630	-.25104
Intraverbal	.45091	-.71635	.10503	-.00621	-.43755
Echoic Gesture	.36555	-.20607	.85281	.11452	-.13964
Comprehension	.24835	-.29802	.11024	-.16906	-.82200
Intraverbal Gesture	.04515	-.01384	-.07240	-.96175	-.11679

TABLE XVI

PERCENTAGE OF VARIANCE EXTRACTED BY EACH FACTOR FOR EACH SUBTEST OF THE ITPA AND PLS

	Factors				
	I	II	III	IV	V
Auditory Vocal Automatic	16%	52%	0%	8%	1%
Visual Decoding	1%	52%	10%	6%	16%
Motor Encoding	0%	26%	7%	11%	36%
Auditory Vocal Association	22%	46%	0%	0%	22%
Visual Motor Sequencing	8%	56%	13%	1%	5%
Vocal Encoding	8%	18%	0%	0%	56%
Auditory Vocal Sequential	72%	7%	10%	0%	3%
Visual Motor Association	0%	50%	5%	0%	16%
Auditory Decoding	14%	58%	1%	0%	10%
Tact	10%	56%	0%	0%	8%
Echoic	65%	18%	4%	0%	6%
Intraverbal	20%	51%	1%	0%	19%
Echoic Gesture	13%	4%	72%	1%	2%
Comprehension	6%	8%	1%	3%	67%
Intraverbal Gesture	0%	0%	0%	92%	1%

TABLE XVII

LIST OF FACTORS

Factor	Description
I	Immediate Recall of Auditory Symbols
II	General Linguistic Ability
III	Imitation of a Motor Act
IV	Ability to Make a Gestural Response to a Verbal Stimulus
V	Comprehension and Expression

analysis shows these two subtests to be sampling the same ability--immediate recall of auditory symbols. An inspection of the content of these two subtests gives additional support to this analysis. The Auditory Vocal Sequencing subtest is purported to measure the ability to reproduce a series of symbols presented auditorily. It consists of twenty different series of numbers two to seven digits in length. These are presented verbally to the subject who in turn attempts to repeat them to the examiner. The Echoic subtest of the PLS also consists of a digit span section (part B) as well as the repetition of words and sentences (part A) and measures the ability to repeat a verbal response made by another person.

Factor II (labeled General Linguistic Ability) accounted for 34% of the variance with six subtests of the ITPA and two of the PLS receiving their highest loadings on this factor. The subtests are presented according to test and in decreasing order of factor loading. They are, for the ITPA: Auditory Decoding, Visual Motor Sequencing, Auditory-Vocal Automatic, Visual Decoding, Visual-Motor Association and Auditory Vocal Association which are all the ITPA subtests except the two Encoding tests and the Auditory-Vocal Sequencing test of Factor I. The two subtests of the PLS loading highest on this factor, again in decreasing order of factor loading, are Tact and Intraverbal. They both measure the ability to produce a verbal response to environmental stimuli; in one case visual, through identification of objects and pictures and in the other verbal, through answering questions. These are essentially the same

abilities measured by the ITPA Auditory Decoding and Visual Decoding which also had high loadings on this factor and respectively measure the ability to understand what is heard and the ability to understand what is seen. Also related to the visual and verbal component are the high loadings received by Visual-Motor Association and Auditory Vocal Association which measure the ability to draw relationships from what is seen and heard respectively. The remaining two subtests with high loadings on this factor also have a visual or verbal component. They are Visual Motor Sequencing, which measures the ability to reproduce a series of symbols presented visually; and Auditory Vocal Automatic, which measures the ability to use the structure of language automatically. The fact that these subtests require verbal responses, motor responses, recognition of visual stimuli, drawing relationships, understanding, and automatic use of the structure of language leads to the interpretation of this factor as one of general linguistic ability.

Factor III (labeled Imitation of a Motor Act) accounted for 8% of the total variance extracted and only one subtest had a high loading on this factor. This was the Echoic Gesture subtest of the PLS which requires the imitation of a motor response, thus this factor may be labeled immediate imitation of a motor act.

Factor IV (labeled Ability to Make a Gestural Response to a Verbal Stimulus) accounted for another 8% of the total variance extracted. Only one subtest had a high loading on this factor. This was the Intraverbal Gesture subtest

which requires a gestural response to a verbal stimulus, thus this factor was labeled the ability to make a gestural response to a verbal stimulus.

It is believed that the inconsistency of the Intraverbal Gesture subtest is responsible for its specific loading on an individual factor rather than the alternate possibility that it is measuring an aspect of language that differs completely from the rest of the subtests of the two instruments. This is not felt to be the case for Echoic Gesture, however, as all but one of the correlations of this measure with those of the ITPA were significantly different from zero. The exception to this being a correlation with a test demonstrating lower reliability than that of the remaining ITPA subtests. The Echoic Gesture correlations with ITPA variables were lower than those for the remaining subtests of the PLS yet significantly different from zero. This supports the conclusion that the Echoic Gesture of the PLS measures an aspect of language behavior different than that measured by any of the ITPA subtests.

Factor V (labeled Comprehension and Expression) accounted for 18% of the total variance extracted. One subtest of the PLS and two from the ITPA received their highest loadings on this factor. The Comprehension subtest of the PLS had the highest loading. This subtest requires the child to respond to three types of directions; vocal, gestural, and a combination of the two. The two ITPA subtests that were highly loaded on this factor were Vocal Encoding and Motor Encoding. These

require the ability to express "ideas" verbally and manually, thus this factor involves comprehension and expression. It may be that the assumption underlying the Encoding tests (that the subjects cannot fail from lack of recognition due to the simple nature of the items) may not be a valid one in the case of mentally retarded children. The Encoding tests may require comprehension (decoding) of the stimuli provided by the tests that are beyond the capacity of many mentally retarded children. The test then becomes more a measure of their capacity to first decode or comprehend the nature of the test item then demonstrate their capacity for encoding (expression). Another possibility is the Comprehension subtest of the PLS may not be measuring just comprehension but also the expression of this comprehension in a motor act. The child may be comprehending but lacks the ability to consistently transform this comprehension into a motor act, A definitive answer to this will dependent upon additional research.

Interpretation of variance percentages. Table XIV shows several subtests shared a large percentage of their variance with more than one factor though their loading on a specific factor was one of the highest on that factor. The Auditory-Vocal Automatic subtest of the ITPA while receiving its highest loading on the general linguistic ability factor shared 16% of its variance with the immediate recall of auditory symbols factor. A look at the content of the Auditory-Vocal Automatic subtest reveals that it too deals in part with auditory symbols

but not to the extent required by the ITPA Auditory-Vocal Sequential and PLS Echoic subtests which had the highest loadings on this factor.

The Visual Decoding subtest while also receiving its highest loading on the general linguistic factor shared 16% of its variance with the comprehension factor. Visual Decoding was designed to test the ability to comprehend. It is surprising that it received less of a loading on this factor than the two Encoding tests of the ITPA which were designed to eliminate comprehension as a factor in producing a successful response. Auditory Decoding while also loading highest on the general linguistic factor shared 10% of its variance with comprehension--much less than that of the Visual Decoding subtest. Auditory Decoding also shared 14% of its variance with the immediate recall of verbal symbols factor. A look at the content of the test shows that it is designed to measure the ability to understand (comprehension) the spoken word (auditory symbols) so the sharing of variance with these factors is easily understood.

Motor Encoding while loading highest on the comprehension factor shared 26% of its variance with the general linguistic factor and another 11% with the factor labeled ability to make a gestural response to a verbal stimulus. A look at the Motor Encoding subtest shows it requires making a gestural response but to a visual rather than verbal stimulus thus the reduced loading on this factor. Vocal Encoding also had its highest loading on the comprehension factor but shared 18% of its

variance with the general linguistic factor. The test was designed to describe symbols in spoken words and 8% of the variance was shared by the immediate recall of auditory symbols factor.

Auditory-Vocal Association while loading highest on the general linguistic ability factor shared 22% of its variance with the immediate recall of auditory symbols factor and another 22% with the comprehension factor. Auditory Vocal Association is designed to measure the ability to relate spoken words (auditory symbols) in a meaningful (comprehension) way. Visual-Motor Association loaded highest on the general linguistic ability factor but shared 16% of its variance with the comprehension factor. Visual Motor Association is designed to measure the ability to relate visual symbols in a meaningful way and evidently comprehension of the visual symbols represents a part in the scores obtained on this test in the case of institutionalized mentally retarded children.

Visual Motor Sequencing also loaded highest on the general linguistic ability factor but shared 13% of its variance with the factor labeled immediate imitation of a motor act. Visual Motor Sequencing was designed to measure the ability to reproduce from memory (imitation of a motor act) a sequence of visual symbols. This subtest also requires to a certain degree the immediate recall of symbol arrangements and comprehension of what is to be done. This accounts for this subtest sharing 8% of its variance with the immediate recall of verbal symbols

factor and 5% of its variance with the comprehension factor. Auditory Vocal Sequencing subtest loaded highest on the immediate recall of verbal symbols factor which is what it is designed to measure. The test also shared 10% of its variance with the immediate imitation of a motor act. This is probably due to the fact the subtest has an element of immediate imitation although of verbal symbols rather than a motor act.

Several of the subtests of the PLS also split their variance among the factors. The Tact subtest while loading highest on the general linguistic ability factor shared 10% of its variance with the immediate recall of auditory symbols factor and 8% of its variance with the comprehension factor. The test requires the ability to verbally identify objects and pictures; thus it has an element of recall of verbal symbols as well as comprehension of the objects, pictures and the task at hand. The Echoic subtest while loading highest on immediate recall of verbal symbols which is the task required by the test also shared 18% of its variance with the general linguistic ability factor. The Intraverbal subtest while receiving its highest loading on general linguistic ability shared 20% of its variance with the immediate recall of verbal symbols factor and 19% of its variance with the comprehension factor. The test requires the ability to produce the correct verbal response (recall of verbal symbols) from the cues in verbal stimuli (comprehension) such as in answering questions. Echoic Gesture loaded highest on a factor labeled immediate imitation of a motor act which is exactly what

the test requires but 13% of the variance was shared by the factor labeled immediate recall of verbal symbols. The reason for this is probably the test requires immediate recall; but it is recall of a motor act rather than verbal symbols thus the variance extracted by this factor is fairly small. Comprehension and Intraverbal Gesture were both so highly loaded on their first factors very little variance was shared with the other factors.

CHAPTER V

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

Summary. This study presented a comparison of the Illinois Test of Psycholinguistic Abilities and the Parsons Language Sample in assessing the linguistic functioning of institutionalized mentally retarded children. Both tests are based on learning models, the ITPA on a revision of Osgood's model and the PLS on a model derived from Skinner's system of verbal behavior. These two tests, both based on learning models, purport to measure linguistic functioning and were compared to determine their similarities and/or differences. The method of determining what factors were common to the two instruments was an experimental comparison of the two tests via the correlation coefficient and a factor analysis of the results.

The sample was composed of fifty mentally retarded children: five males and five females from each of five age levels (ten, eleven, twelve, thirteen and fourteen) selected randomly from all the patients in residence at the Parsons (Kansas) State Hospital and Training Center who fell within these age brackets. The experimental design employed the "counterbalancing" technique with half the subjects receiving the ITPA first and PLS second and the other half the PLS first and ITPA second. A one day interval separated test administrations as an additional control for transfer and fatigue. Testing was done entirely by the author of this research and took place over a nine week period.

The results of the administrations of the two tests were compared through the Pearson Product-Moment Correlation Coefficient. The Assumptions underlying the use of this statistic were tested and found to be satisfied. The ITPA and PLS intercorrelations derived by this process were compared statistically with the intercorrelations presented in the standardization data of each test. The intercorrelations for the ITPA and PLS derived from this research were generally lower than those presented in the standardization data, possibly because of the large differences in sample size. For the ITPA only two of the intercorrelations were significantly different at the .05 level. These two were Motor Encoding with Auditory-Vocal Automatic and Motor Encoding with Auditory-Vocal Sequencing. The original reliability data revealed that of the nine ITPA subtests Motor Encoding had the lowest internal consistency and test-retest stability. Three of the PLS intercorrelations exceeded the .05 level of confidence. These were Echoic Gesture with Nonvocal Total, Echoic Gesture with PLS Total, and Tact with the Nonvocal Total. This difference was attributed to lower reliability in the case of Echoic Gesture, as supported by the low examiner equivalence, internal consistency and test-retest stability data presented in the original monograph. The Tact and Nonvocal Total discrepancy was considered a result of sample variation.

The correlations between the variables of the ITPA and PLS were tested to determine if they differed significantly from zero

correlation or if the difference could be explained more expediently as due to outside variables. Seventeen of the 108 correlations were not significantly different from zero. Twelve of these were the Intraverbal Gesture subtest with each of the variables of the ITPA and four additional nonsignificantly correlations were due to the influence of the Intraverbal Gesture subtest lowering the Nonvocal total to the point the correlations were not significantly different from zero. The remaining nonsignificant correlation was Motor Encoding with Echoic Gesture. Apparently, they lacked the needed reliability to demonstrate a significant relation to each other. These ITPA and PLS correlation coefficients were also compared with one other unpublished research. The correlation coefficients of the two studies did not differ at either the .01 or .05 levels of confidence.

The correlation matrix obtained from this research was subjected to a factor analysis and the first five factors derived from the factor matrix were rotated to extract the maximum amount of variance. These factors accounted for a total of 85% of the variance. Factor I (Immediate Recall of Auditory Symbols) accounted for 17% of the total variance. The two subtests loading highest on this factor were the ITPA Auditory-Vocal Sequencing and the PLS Echoic subtest. Factor II (General Linguistic Ability) accounted for 34% of the total variance. Six subtests of the ITPA (Auditory Decoding, Visual Motor Sequencing, Auditory Vocal Automatic, Visual Decoding, Visual-Motor Association and Auditory Vocal Association) all

received their highest loadings on this factor. The two subtests of the PLS receiving their highest loading on this factor were Tact and Intraverbal. These subtests requiring verbal responses, motor responses, recognition of visual stimuli, drawing relationships, understanding, and automatic use of the structure of language leads to the interpretation of this factor as one of general linguistic ability. The only subtest with a high loading on Factor III (Imitation of a Motor Act) was the PLS Echoic Gesture and on Factor IV (Ability to Make a Gestural Response to a Verbal Stimulus) the PLS Intraverbal Gesture. Each factor accounted for 8% of the total variance extracted. The appearance of Factor IV was felt to be due to inconsistency of measurement as a result of the nature of the task not being made clear to the subject in the directions of the Intraverbal Gesture subtest. Factor III, however, reveals that the Echoic Gesture subtest of the PLS is somewhat independent of the variables of the ITPA. Echoic Gesture had the lowest correlations with the variables of the ITPA yet all but one correlation was significantly different from zero. This indicates it is sampling an ability related to linguistic behavior but the ability is different from those measured by the ITPA. Factor V (Comprehension and Expression) accounted for 18% of the total variance extracted. The Comprehension subtest of the PLS and the Vocal Encoding and Motor Encoding subtests of the ITPA received their highest loadings on this factor. At least two possibilities exist

for the explanation of this factor. It may be mentally retarded children do not consistently comprehend the stimuli provided by the Encoding tests thus they become more a measure of comprehension than encoding ability or the children are comprehending the stimuli provided by the Comprehension but are unable to consistently express this (encoding) in a motor act. The determination of this point requires further research.

Several of the ITPA and PLS subtests shared a portion of their total variance with factors other than the one upon which their highest loading occurred. This points out that both tests are composed of subtests that require more than just one ability. Interpretation of the variance percentages revealed the extent the fifteen subtests of the two instruments are dependent upon factors other than the ones which received the highest loadings.

Conclusions. The main conclusion to be derived from this research is that the Illinois Test of Psycholinguistic Abilities and the Parsons Language Sample tap highly related behavior. The only subtests that did not show an extremely high degree of interrelatedness were the Motor Encoding subtest of the ITPA with the Echoic Gesture subtest of the PLS and the Intraverbal Gesture subtest of the PLS with all the subtests of the ITPA.

The factor analysis revealed: (1) the null hypothesis of no factors common to both instruments cannot be accepted; (2) there are at least three factors common to both instruments- immediate recall of auditory symbols, general linguistic ability and comprehension; (3) the ITPA Auditory-Vocal Sequencing appears

to tap essentially the same ability as the PLS Echoic; (4) the ITPA Vocal Encoding and Motor Encoding subtests appear to tap the same ability as the Comprehension subtest of the PLS; (5) the Auditory Decoding, Visual Decoding, Visual-Motor Sequencing, Visual-Motor Association, and Auditory-Vocal Automatic subtests of the ITPA, and PLS Tact and Intraverbal appear to measure general linguistic ability rather than a number of specific abilities, ITPA Auditory Vocal Association and PLS Intraverbal being the most closely related; (6) the Intraverbal Gesture subtest of the PLS is sampling behavior unrelated to the other language measures and perhaps unrelated to language; (7) the Echoic Gesture subtest is sampling behavior different from that of the other subtests of the ITPA and PLS but is sampling behavior related to language.

Additional conclusions are: (1) The division of the ITPA into primarily verbal tests and primarily nonverbal tests is of little value with institutionalized mentally retarded children as the two divisions correlate quite high (.80) and are apparently measuring similar abilities. (2) The results of the ITPA when administered to institutionalized mentally retarded children are not dependent upon either the chronological age or sex of the child. The correlations between the subtests of the ITPA and chronological age and sex of the subjects were not significantly different from zero. (3) The results of the PLS when administered to institutionalized mentally retarded children are not dependent upon the sex of the child for the same reason as stated above. The correlations of the subtests of the PLS with the chronological

age of the subjects revealed that the older the child the higher the score on the Echoic Gesture due to the fact that it is significantly related to chronological age in a positive direction (.29). The nature of the Echoic Gesture subtest evidently requires certain ability that comes with advances in chronological age. The correlation of the Intraverbal Gesture subtest with chronological age revealed that the younger the child the higher the score on the Intraverbal Gesture subtest. This subtest was significantly related to chronological age in a negative direction (-.33). Younger subjects evidently respond to questions through gestures more than do older ones. (4) All of the subtests of the ITPA are significantly related to the Adaptive Behavior Level of the subject. The closer the Adaptive Behavior level is to that of a normal child the higher the scores on the subtests of the instrument. (5) The Intraverbal Gesture subtest of the PLS as well as the Nonvocal total are not significantly related to Adaptive Behavior level but the remaining subtests do demonstrate such a relationship. Again the closer the Adaptive Behavior level is to that of a normal child the higher the scores on the subtests of the instrument.

Skinner's and Osgood's models have both led to the development of highly useful tests for assessing the linguistic deficiencies of the mentally retarded. Of the two Skinner's model fit into the PLS more readily than Osgood's into the ITPA. The model developed by Osgood had to be revised considerably before subtests could be developed that adequately assessed

linguistic ability in exceptional children. The model developed by Osgood while not lending itself as readily to test construction has nevertheless led to the development of an extremely useful test. Both tests deviate considerably from the language models upon which they are based. Several parts of Skinner's theory were not included in the Parsons Language Sample because they were concerned with aspects of language not readily measurable or perhaps even present in the mentally retarded. Experience with the ITPA and PLS has led this author to conclude Osgood's model provides a more thorough basis for understanding linguistic deficiencies in the mentally retarded. The theory attempts to fractionate linguistic ability into processes and reveal the processes that are weak or inoperative. Although some additional research is needed along these lines (especially in the encoding area) this type of approach appears extremely useful. Skinner's model seems limited to pointing out that deficiencies do exist but provides limited insight into the meaning this deficit has for the individual and little recourse to overcoming the deficit.

Recommendations. It is hoped this study has provided some meaningful information concerning the Illinois Test of Psycholinguistic Abilities and the Parsons Language Sample. There is a continually increasing interest in the ITPA as it becomes increasingly well known and used. While the Parsons Language Sample is not as widely known or used, those who do know of it and have used it feel it is the nucleus of a

potentially valuable instrument. It is the hope of this author that this research will stimulate increased interest in bringing the PLS to the point of development currently enjoyed by the ITPA.

One of the first things that needs to be considered is the revision of the directions for the PLS Intraverbal Gesture subtest. As stated in the discussion the nature of the response required for successful performance on this test is not made clear to the subject. The initial aim of the subtest that of measuring both gestural and nongestural behavior was abandoned after certain complications of standardization; but the directions were not revised so subjects could be aware that gestural behavior is the type of behavior expected. A simple SHOW ME stated before each item of the test should accomplish this.

The primary consideration for the ITPA according to the results of this study is the revision of the Encoding subtests if the tests are to accurately measure this ability in institutionalized mentally retarded children. As pointed out in the discussion both Motor and Vocal Encoding subtests apparently depend to too great an extent on comprehension or decoding of the test stimuli. Future research should determine if the items of these two subtests can be comprehended by those who demonstrate encoding ability on an equivalent measure. The Motor Encoding subtest also needs additional research to tighten the scoring system and thus its reliability.

A few recommendations are in order to make research similar to this more enlightening. First, the sample should be sufficiently large to enable a factor analysis of the results according to a particular level of linguistic ability. The wide range of talent employed as a result of a random sample of an entire hospital age group tends to increase the correlation coefficients. The factor analysis of levels of linguistic ability would provide increased sensitivity to the factors present at that particular level. Another recommendation concerns the possibility of additional stratification of the sample. If the number of males and females at each of several age levels were increased and further stratified to include the five measured intelligence levels, this would provide a comparison of each test and its scales and subtests with the measured intelligence level of the children in the sample as well as sex and age comparisons.

The last recommendation concerns comparison of the two tests with several other measures of language ability. This will help in establishing the concurrent validity of the tests and add additional knowledge for interpretation and perhaps future modifications.

As for the specific recommendations as to the use of the ITPA and PLS based on the results of this study it would appear the minimum battery of subtests that could be administered and still represent the abilities measured by each test would be composed of the ITPA Auditory Vocal Sequencing to measure

immediate recall of auditory symbols, the ITPA Auditory Decoding test to measure general linguistic ability, the PLS Echoic Gesture to measure the ability to imitate a motor response and the PLS Comprehension subtest to measure comprehension. A more complete battery of subtests from the two instruments would include all but the PLS Intraverbal Gesture which apparently in its present form does not measure ability related to language behavior and the ITPA Vocal and Motor Encoding unless they are considered as possibly more a measure of comprehension ability than expression ability.

BIBLIOGRAPHY

BIBLIOGRAPHY

Books

Adcock, C. J. Factorial Analysis for Non-mathematicians. Carlton, Victoria: Melbourne University Press, 1954.

Bateman, Barbara D. "Reading and Psycholinguistic Process of Partially Seeing Children," Selected Studies on the Illinois Test of Psycholinguistic Abilities. Madison: Photo Press Inc., 1963.

Bateman, Barbara D. The Illinois Test of Psycholinguistic Abilities in Current Research. Urbana: Institute for Research on Exceptional Children, 1964.

Center, W. R. "A Factor Analysis of Three Language and Communication Batteries," Cited by Bateman in The Illinois Test of Psycholinguistic Abilities in Current Research. Urbana: Institute for Research on Exceptional Children, 1964.

Cochran, W. G. Sampling Techniques. New York: John Wiley and Sons, Inc., 1963.

Dunn, L. M. Expanded Manual--Peabody Picture Vocabulary Test. Minneapolis: American Guidance Services, Inc., 1965.

Ferguson, G. A. Statistical Analysis in Psychology and Education. New York: McGraw-Hill, Inc., 1959.

Fruchter, B. Introduction to Factor Analysis. New York: Van Nostrand, 1954.

Guilford, J. P. Fundamental Statistics in Psychology and Education. New York: McGraw-Hill, Inc., 1942.

Hilgard, E. R. Introduction to Psychology. New York: Harcourt, Brace and World, Inc., 1962.

Hull, C. L. Principles of Behavior. New York: Appleton-Century, 1943.

Kass, Corrine E. "Some Psychological Correlates of Severe Reading Disability (Dyslexia)," Selected Studies on the Illinois Test of Psycholinguistic Abilities. Madison: Photo Press, Inc., 1963.

Loeffler, F. J. "An Extension and Partial Replication of Meyers, et. al. 'Primary Abilities at Mental Age Six,'" Cited by Bateman in The Illinois Test of Psycholinguistic Abilities in Current Research. Urbana: Institute for Research on Exceptional Children, 1964.

McCarthy, J. J. "Qualitative and Quantitative Differences in the Language Abilities of Young Cerebral Palsied Children," Selected Studies on the Illinois Test of Psycholinguistic Abilities. Madison: Photo Press, Inc., 1963.

McCarthy, J. J., and S. A. Kirk. Illinois Test of Psycholinguistic Abilities--Examiner's Manual. Urbana: Institute for Research on Exceptional Children, 1961.

McCarthy, J. J., and S. A. Kirk. The Construction, Standardization and Statistical Characteristics of the Illinois Test of Psycholinguistic Abilities. Madison: Photo Press, Inc., 1963.

McCarthy, J. J., and J. L. Olson. Validity Studies on the Illinois Test of Psycholinguistic Abilities. Madison: Photo Press, Inc., 1964.

McGuigan, F. J. Experimental Psychology--A Methodological Approach. Englewood Cliffs: Prentice-Hall, Inc., 1960.

McNemar, Q. Psychological Statistics. New York: John Wiley and Sons, Inc., 1949.

Osgood, C. E. "A Behavioristic Analysis," Contemporary Approaches to Cognition. Cambridge: Harvard University Press, 1957a.

Osgood, C. E. "Motivational Dynamics of Language Behavior," Nebraska Symposium on Motivation. Lincoln: University of Nebraska Press, 1957b.

Osgood, C. E., C. J. Suci, and P. H. Tannebaum. The Measurement of Meaning. Urbana: University of Illinois Press, 1957.

Olson, J. L. "A Comparison of Receptive Aphasic, Expressive Aphasic, and Deaf Children on the Illinois Test of Psycholinguistic Abilities," Selected Studies on the Illinois Test of Psycholinguistic Abilities. Madison: Photo Press, Inc., 1963.

Semmel, M. I., and M. W. Mueller. "A Factor Analysis of the Illinois Test of Psycholinguistic Abilities with Mentally Retarded Children," Cited by Bateman in The Illinois Test of Psycholinguistic Abilities in Current Research. Urbana: Institute for Research on Exceptional Children, 1964.

Sievers, Dorothy J. "Development and Standardization of a Test of Psycholinguistic Growth in Preschool Children," Selected Studies on the Illinois Test of Psycholinguistic Abilities. Madison: Photo Press, Inc., 1963.

Skinner, B. F. Verbal Behavior. New York: Appleton-Century Crofts, Inc., 1957.

Slanner, B. F. Science and Human Behavior. New York: McMillan Co., 1953.

Spradlin, J. E. "Language and Communication of Mental Defectives," Handbook of Mental Deficiency. Edited by Norman R. Ellis. New York: McGraw-Hill, Inc., 1963b.

Thompson, G. H. The Factorial Analysis of Human Ability. Boston: Houghton Mifflin, 1950.

Publications of the Government

Lassers, L., and G. Low. "A Study of the Relative Effectiveness of Different Approaches of Speech Therapy for Mentally Retarded Children." Report to the Office of Education on Contract No. 6904, 1960.

Periodicals

Bateman, Barbara D. "The Role of the ITPA in Differential Diagnosis and Program Planning for Mentally Retarded." Amer. J. Orthopsych., 35:465-472, (April, 1965).

Bateman, Barbara D., and Janis Wetherell. "Psycholinguistic Aspects of Mental Retardation." Mental Retardation. Vol. 3, No. 2:8-13, (April, 1965).

Copeland, R. H. "Effects of Consequences on Vocal Behavior, 2. The Effects of Feedback Modification on Verbal Behavior." J. Speech Hear. Disord., Monogr. Suppl. 10:70-75, (January, 1963).

Horowitz, Frances D. "Effects of Consequences on Vocal Behavior, I. Partial and Continuous Reinforcement of Vocal Responses Using Candy, Vocal and Smiling Reinforcers Among Retardates." J. Speech Hear. Disord., Monogr. Suppl. 10:55-69, (January, 1963).

Leland, H. "Some Thoughts on the Current Status of Adaptive Behavior." Mental Retardation. Vol. 2, No. 3:171-177, (June, 1964).

McCarthy, J. J. "A Test for the Identification of Defects in Language Usage Among Young Cerebral Palsied Children." Cer. Palsy Rev. 21:3-4, 1960.

McCarthy, J. J. "Notes on the Validity of the ITPA." Mental Retardation. Vol. 3, No. 2:25-26, (April, 1965).

Mecham, M. J. "The Development and Application of Procedures for Measuring Speech Improvement in Mentally Defective Children." Amer. J. Ment. Defic., 60:301-306, 1955.

Olson, J. L., H. R. Hahn, and Anita L. Herman, "Psycholinguistic Curriculum." Mental Retardation. Vol. 3, No. 2:14-19, (April, 1965).

Rosenberg, S., J. E. Spradlin, and S. Mabel. "Interaction Among Retarded Children As a Function of their Relative Language Skills." J. Abnorm. soc. Psychol. 63:402-410, 1961.

Siegel, G. M. "Language Behavior of Adults and Retarded Children in Interpersonal Assemblies, I; Adult Verbal Behavior in 'Play Therapy' Sessions with Retarded Children." J. Speech Hear. Disorders, Monogr. Suppl. 10:34-38, (January, 1963).

Siegel, Gerald M. "Language Behavior of Adults and Retarded Children in Interpersonal Assemblies, 3. Verbal Behavior of Retarded Children Assembled with Pre-Instructed Adults." J. Speech Hear. Disord., Monogr. Suppl. 10:47-53, (January, 1963).

Siegel, G. M., and Jerome P. Harkins. "Language Behavior of Adults and Retarded Children in Interpersonal Assemblies, 2. Verbal Behavior of Adults in Two Conditions with Institutionalized Retarded Children." J. Speech Hear. Disord., Monogr. Suppl. 10:55-69, (January, 1963).

Spradlin, J. E. "Assessment of Speech and Language of Retarded Children: The Parsons Language Sample." J. Speech Hear. Disorders, Monogr. Suppl. 10:8-97, (January, 1963).

Wiseman, D. E. "A Classroom Procedure for Identifying and Remediating Language Problems." Mental Retardation. Vol. 3, No. 2:20-24, (April, 1965).

Unpublished Materials

Edson, Sandra K., and J. C. Rolland, "Rationale for Modification of Instruction on PLS and Standardization, Administration and Scoring of the PLS." Parsons Demonstration Project Report No. 14A and 14B, (September, 1964) Parsons State Hospital, Parsons, Kansas (Mimeographed).

Edson, Sandra K., and J. C. Rolland. "Revision of the Parsons Language Sample, September 1964." Parsons Demonstration Project Report No. 15, (September, 1964) Parsons State Hospital, Parsons, Kansas (Mimeographed).

Edson, Sandra K., and J. C. Rolland. "Correlations Between Certain Tests in the Evaluation Battery." Parsons Demonstration Project Report No. 26, (March, 1965) Parsons State Hospital, Parsons, Kansas (Mimeographed).

Harkins, J. P. "A Study of the Verbal Behavior of Adults Assembled in Two Conditions with Institutionalized Mentally Deficient Children of Two Verbal Levels." Unpublished Master's Thesis, The University of Kansas, 1961.

Horowitz, Frances D. "The Relationship Between the Parsons Language Sample Scores and Wechsler Intelligence Quotients of Mentally Retarded Subjects." J. Genet. Psychol. (In Press), p. 2.

Leland, H. "Conference on Measurement of Adaptive Behavior." Rehabilitation Center, Parsons State Hospital, Parsons, Kansas, 1964, (Mimeographed).

McCarthy, Dorothea. "A Preliminary Report on the Verbal Items of a New Psychological Appraisal Test with Institutionalized Mentally Retarded Children." Paper read at the Council for Except. Child., Northeast Assn., Atlantic City, New Jersey, (April, 1959).

Spradlin, J. E., and S. Rosenberg. "Complexity of Adult Verbal Behavior in a Dyadic Situation with Retarded Children." Parsons Research Project Working Paper 18, 1959, Parsons State Hospital, Parsons, Kansas, (Mimeographed.)

Wiseman, D. E. "Program Planning for Retarded Children with Psycholinguistic Disabilities." Unpublished Doctoral dissertation, The University of Illinois, Urbana, 1964.

Tests

Dunn, L. M. The Peabody Picture Vocabulary Test. Minneapolis: American Guidance Services, Inc., 1958.

McCarthy, J. J., and S. A. Kirk. Illinois Test of Psycholinguistic Abilities. Urbana: Institute for Research on Exceptional Children, 1961.

Spradlin, J. E. The Parsons Language Sample. Parsons Research Project, Parsons State Hospital, Parsons, Kansas, 1963.