USES AND ABUSES OF THE ITPA

Samuel A. Kirk and Winifred D. Kirk
University of Arizona, Tucson

This paper describes the circumstances leading to the development of the Illinois Test of Psycholinguistic Abilities (ITPA) and discusses the malpractices in administration, scoring, and interpretation of the test. Certain relevant research endeavors are summarized and evaluated. These include (1) equivalence of the experimental and research editions, (2) outcomes and pitfalls of factor analysis, (3) the relationship of ITPA scores to those of the WISC and Stanford-Binet, (4) ethnic differences shown on the ITPA, (5) differences among groups of atypical children, and (6) ITPA correlates of reading disabilities. A critique is presented regarding the misuse of the ITPA with subjects beyond the range of the norms and possible uses of the test in areas for which it was not originally intended. Emphasis is placed on the major use of the test as being a clinical indicator in determining differential abilities and disabilities (intraindividual differences) among cognitive, perceptual, and memory functions in young children.

This paper (1) describes the circumstances leading to the development of the Illinois Test of Psycholinguistic Abilities (ITPA), (2) evaluates practices in administering, scoring, and interpreting the results, (3) evaluates some current relevant research, and (4) points out the proper uses of the instrument.

THE ORIGINS OF THE ITPA

The ITPA originated out of necessity when we were conducting an experiment on the effects of preschool education on the mental and social development of young mentally retarded children. At that time, 1949, we had organized a preschool for three-, four-, and five-year-old children who had been diagnosed as mentally retarded. We organized a class in a state institution and another in a school in the community, each with an appropriate contrast group. These children were classified as mentally retarded primarily on the basis of medical diagnoses and tests such as the Stanford-Binet, the Kuhlmann-Binet, and the Vineland Social Maturity Scale. These were the preschool tests then available for the diagnosis and classification of mental retardation.

In trying to organize an educational program for these children according to their needs, we found little if any help from the psychological tests men-

tioned that resulted in educational remediation. It was necessary for us to observe each child, find out by observation what he could and could not do, and then organize an educational program based on the child's successful and unsuccessful behaviors as we observed them.

We soon discovered that the diagnosis, "mental retardation," had very little meaning as far as a preschool educational program was concerned. Each child was different, and each required a different educational emphasis. One child had cerebral palsy, an IQ of 55, with a major disability in quantitative thinking. Another child was diagnosed as having Down's Syndrome and had very little receptive or expressive verbal language. Another child, at age five, could not talk, and another child with a diagnosis of complications following rubella had an extreme visual disability and could not respond to visual objects efficiently. It became necessary for us to organize a different remedial program for each child. One child required training in visual reception and speed of perception; another child required auditory training to develop auditory reception and verbal expression; another, special training in quantitative thinking. Our remedial programs were based on our clinical judgment relating to the special abilities found in each child.

Having become disenchanted with IQ's, global scores, and classifications, we concluded that what we needed for the educational assessment of these children was an intraindividual test; a test that would tell us in what areas a child was successful or unsuccessful. An IQ or any global score is a composite of several abilities and disabilities, but does not pinpoint the special needs of the child. If a child is generally low in all areas, his needs are quite different from those of a child who is low in some areas but average or better in other abilities. We needed tests that taxed the use of certain psychological abilities and disabilities if we were to organize an adequate educational program.

To save time and arduous intuitive clinical diagnosis we began searching for procedures that would aid us in determining those areas in which a child could function well and those in which his behavior indicated he had difficulty. Four areas were delineated (1) verbal communication, namely receptive and expressive language, (2) the reception of visual symbolic material and parallel motor expression by gestures and writing, (3) auditory and visual memory and auditory and visual closure, and (4) responses that required the mediating process of seeing relationships and integrating concepts.

In search of an informational processing model, we adapted the theoretical constructs of Osgood (1957), and the clinical model of Wepman et al. (1960) into an ITPA model (Kirk and McCarthy, 1961). It contained (1) two channels of communication: (a) a visual-motor channel, and (b) an auditory-vocal channel; (2) three processes of communication: (a) the receptive process, including auditory and visual reception, (b) an association or mediating process, and (c) an expressive process: vocal and motor; (3) two levels of organization: (a) a representational level involving the use of visual or auditory symbols,

and (b) the less voluntary, automatic level which short-circuits the use of symbols. Over a period of years tests were constructed to tap these two channels. three processes, and two levels. (See Kirk and Kirk, 1971, for a complete description of the model.)

Work on the ITPA commenced in 1951 and continued for 10 years. At that time, 1961, an experimental edition was standardized so that the test could be used for research. After seven years of research and clinical use, the instrument was revised and published in 1968.

Before we discuss the uses and abuses of the ITPA, we will summarize the original major purposes of the ITPA:

1. The ITPA was designed as a diagnostic test to delineate intraindividual variations in functioning in those areas involved in language and other forms of communication. We called it a psycholinguistic test because it was concerned with psychological functions of information processing, perception, and memory as well as the use of linguistic codes. The term may not now be an adequate designation of the test, but it was suitable in 1961 before linguists developed their own use of the term.

2. The test is designed to measure some areas in which a child succeeds or fails, not to obtain an IQ or its equivalent for classification purposes (Kirk and McCarthy, 1961).

3. It is a test for young children. The norms range from two and one-half to 10 years, but its greatest usefulness is between four and eight, since the norms at both extremes were obtained to increase the ceiling for eight-year-olds, or to show a deficit for three- and four-year-olds.

4. The functions it tests are designed to denote deficits that require remediation. If the child has difficulty in communication because he cannot translate the code used in spoken language, that is, if he does not understand auditory language (receptive aphasia), then a remedial program should be organized to teach him to understand language. If he is delayed in talking (expressive aphasia), he should be taught techniques for expressing himself verbally. If he has a deficiency on a number of auditory vocal tests and is average or above on visual motor tests, remediation should be organized to improve the whole range of auditory vocal abilities. Furthermore, if the child has superior abilities in some functions, they should be used to develop parallel abilities in the deficient areas.

EVALUATION OF PRACTICES

After describing the purpose of the ITPA and its possible uses as a diagnostic and research instrument, we will discuss its misuse and some pitfalls in administration, scoring, and interpretation. It should be emphasized that the ITPA is not all things to all children. It is not the solution to all children's problems. It is one step beyond global tests that may lead to a program of remediation. One of its greatest abuses is to consider it a solution to all prob-

lems and to use scores as a final diagnosis instead of another possible aid to clinical judgment.

Five specific problems should be noted:

1. The ITPA is not an easy test to learn, although conscientious professionals can learn to administer it efficiently if they are willing to put in the time and practice. It does require careful study, but the material for such study is available in the *Examiner's Manual of the ITPA* (Kirk, McCarthy, and Kirk, 1968) and in the later booklet *Aids and Precautions in Administering the ITPA* (Kirk, 1974).

2. Many individuals are misusing the test by not following directions in the manual. We have found many using the test inappropriately. In one workshop designed for so-called experienced examiners, we found only three out of the 17 enrolled who could adequately administer the test. One was a school psychologist, one a speech-language pathologist, and one a learning-disabilities teacher. In another workshop we found that the college instructor teaching the ITPA had not taken the time to learn the test.

3. Many universities are now offering courses on the administration and interpretation of the ITPA. Unfortunately, however, many of these courses are being taught by people who do not know how to administer and interpret the test adequately. Hopefully, as more and more people are adequately trained to use and understand the test, this situation will diminish.

4. Lack of understanding of the basic concepts that underlie the test often creates misinterpretation of the test results. Even if the test has been well administered, the test results are of little value if not adequately interpreted. Using the overall score to evaluate a child or to compare him with other children rather than to compare his own subtest scores with each other are common misconceptions of the use of the test. Mechanical use of scores and labels without their relationship to other information is also a common mistake. The concept of intraindividual differences is difficult to understand since we have for many years dealt only with interindividual differences. Kirk and Kirk (1971) have published a volume on interpretation and remediation.

5. Another error that is made is the use of the test with the wrong subjects. A differentiation must be made between the clinical use of the test and its use for research purposes. Some clinical use of the test can be made even at the top and bottom range of the norms, which extend from two and one-half to 10 years. Scores above and below these ages of course do not define the child's abilities; and, as with any test, scores at the upper and lower limits of the norms are somewhat tenuous. Clinical prudence is necessary in interpretation of such scores. Some clinical use can also be made with atypical subjects outside the normative ages, for example, with aphasic adults. Such use, however, is usually made to supplement other clinical evidence.

For research purposes, though, the test is practical only for children of about four to eight years of age, since many four year olds score lower and

many eight year olds score higher than the test norms. A buffer zone above and below these limits is essential for reputable research. It is most disconcerting to peruse research reports where the ITPA has been used with third- and fourth-grade children, many of whom undoubtedly score at or above the normative limits.

EVALUATION OF RESEARCH

In discussing some of the research presented herein, we may appear to be somewhat critical. For that reason, we are presenting some criteria we use in evaluating research projects.

1. First, research using experimental and control groups is very difficult and, in many cases, impossible to conduct because of the difficulty in finding comparable cases. Populations of children with learning disabilities are so heterogeneous that it is practically impossible to form experimental and control groups of children with the same problems. Some of the children's cases are so rare that their prevalence is probably one in 5000 children, and such cases cannot be used in nomothetic research.

2. A second criterion for research is that of using appropriately aged children for the study. It must be remembered that research in contrast to clinical work with the ITPA is applicable only to young children between the ages of four and eight. When research conclusions are based on data from nine- and 10-year-old average children, the results are of dubious import. Some of the children in such groups reach the ceiling on most of the subtests and many reach the ceiling on one or more of the subtests, thus invalidating the research results. Such instances will be pointed out in the review of research studies.

3. A third criterion is the qualification of those who administer the tests for the research project. Many studies are made, and many children are tested by people inadequately trained and prepared in the administration of the ITPA.

These criteria should be borne in mind in evaluating the voluminous research using the ITPA. Much solid information on the meaning and interpretation of the ITPA and its concepts can be derived from the research, although some research conclusions are contradictory because of variations in methods, subjects, designs, and administration. Several excellent reviews of research using the ITPA have been made by Bateman (1965), Sedlack and Weener (1973), Proger, Cross, and Burger (1973), and Buros (1972).

We would like to make note, however, of the following topics of research: (1) the equivalence of the experimental and revised editions of the ITPA, (2) factor analysis on the ITPA, (3) relationship of the ITPA to other tests, (4) ethnic differences as measured by the ITPA, (5) studies on various clinical types, and (6) ITPA functions related to reading. In the following, some relevant research projects are noted and evaluated.

Equivalence of Experimental and Revised ITPA Editions

One question asked frequently relates to the equivalence between the experimental and the revised edition of the ITPA. Can results from the two be safely compared? Two studies have been conducted to answer this question. One study (Hubschman, Polizzotto, and Kaliski, 1970) found a correlation of 0.95 between two administrations of the experimental edition and 0.93 between the second administration of the experimental edition with the 1968 revised edition. These correlations are much higher than those found in the test-retest stability measures conducted by the authors. In examining these studies we found that the researchers did not partial out age or IQ, thus producing spuriously high correlations. Another study by Waugh (1973) that controlled for age obtained correlations of 0.65. The latter is probably a more accurate correlation.

Factor Analysis of the ITPA

In the book on the *Development and Psychometric Characteristics of the Revised ITPA,* Paraskevopoulos and Kirk (1969) stated:

> The intercorrelations among ITPA subtests are too complex to achieve simple structures. . . . Attempts to factor analyze ITPA data would probably generate results yielding only confusion, rather than simplicity and parsimony. (p. 184)

In spite of our warning, many factor analytic studies have been made. The availability of computer programs for factor analysis has resulted in report after report of factor analysis of ITPA data. After reviewing the results of 20 factor-analytic studies on the ITPA, Sedlack and Weener (1973, pp. 123-125) concluded that the factor-analytic studies yielded "confusing and contradictory results" as we had predicted. They stated:

> It is safe to say that no more factor analyses with small *N*s using only ITPA subtests are needed. The availability of computer routines for factor-analytic procedures has resulted in a tendency for researchers to factor-analyze whenever they have more than five variables available on a set of subjects. This tendency has risen to promiscuous levels in the case of the ITPA. Future factor-analytic work should proceed from a careful a priori theoretical framework, should use a large number of subjects from different age and ability subpopulations, and should be done by those with a thorough grasp of factor-analytic procedures.

Since Sedlack and Weener's review, Hare, Hammill, and Bartel (1973) conducted a further analytic study using 126 third-grade children and included a very important technique of factor-analyzing ITPA data along with other reference tests. Seven factors emerged and accounted for 66% of the variance. They concluded that the ITPA does have construct validity. Unfortunately, the authors of this study used eight- and nine-year-old children—children at the upper levels of the norms. The study, however, using reference tests, is in the right direction.

Taken all together, we feel that the numerous factor-analytic studies of the ITPA have neither proved nor disproved the construct validity of the ITPA.

A different technique of studying the construct validity of the ITPA has been postulated by both Cohen (1973) and Elkins (1973[1]). They used a Guttman-Lingoes nonmetric space-analysis technique. Both found that the process and channel dimensions were supported, but the levels (representational and automatic) were not clear. Apparently, the Grammatic Closure Test factors out as a representational rather than an automatic level test. Aside from this, the levels appear to have some validity.

Relationship of the ITPA to Other Tests

There have been many studies relating the ITPA to other psychometric tests. Most of the studies relate the global scores of the ITPA to IQ's or mental ages.

By administering the Stanford Binet, WISC, and ITPA to 100 first-grade children, Huizinga (1971) found that the ITPA correlated 0.90 with the Binet and 0.80 with the WISC full scale. The WISC correlated 0.84 with the Binet. The auditory-vocal scaled scores correlated 0.84 with the Binet and 0.75 with the verbal scale of the WISC. The visual-motor subtests correlated 0.68 with the Binet and 0.58 with the Performance WISC. Other correlational studies relating the ITPA to the full scale WISC were Polley's (1971) showing a correlation of 0.49, Guest's (1970) with a correlation of 0.67, Humphrey and Rice's (1973) with a correlation of 0.88, and Bartin's (1971) with a correlation of 0.61.

The conclusion from these studies is that the ITPA correlates quite highly with the Binet, less with the WISC, and still less with the performance scale of the WISC.

Ethnic Differences as Measured by the ITPA

Another question that has been raised relates to the applicability of the ITPA to minority groups. Consequently, ethnic differences have been studied with both the experimental and the revised edition.

The first study was conducted by Ryckman (1966) using the experimental edition of the ITPA. He compared 50 middle-class 5-year-old black children with 50 lower-class 5-year-old black children. He found a significant difference between the middle-class and lower-class black children on all tests. The middle-class blacks, however, resembled the normative sample of predominantly white children except for the superiority of the blacks on auditory sequential memory. The lower-class blacks also showed a superiority in auditory sequential memory, but were depressed in all functions and in comparison to the middle-class blacks.

[1]J. Elkins, A Guttman-Lingoes non-metric representation of the subtests of the Revised ITPA. Personal communication, University of Queensland (1973).

Lombardi (1970) found that Papago Indian first- and third-grade children were below the white norms on most tests and especially on the auditory vocal tests. One outstanding result was that the Indian children scored significantly above the white norms on visual sequential memory.

Jorstad (1971) studied the ITPA scores of 20 Mexican-American children diagnosed as learning disabled. She found major deficiencies on the auditory vocal tests and a slight superiority, as with the Papago Indians, on visual sequential memory.

The Westinghouse Study (1969, p. 214) and Cicirelli et al. (1971) showed similar tendencies for Head Start children. In this study blacks are superior in auditory sequential memory, the Mexican Americans are superior in visual sequential memory, and both groups are below the whites on the other auditory vocal tests (auditory reception, auditory association, verbal expression, and grammatic closure) (Kirk, 1972). This is not surprising in view of the difficulties arising from bilingualism.

Studies on Various Clinical Types

In addition to studies of the relationship of the ITPA and other tests to academic achievement, there have been a series of research projects on various types of handicapped children. These deal with psycholinguistic characteristics of children with speech disorders, Down's Syndrome, cerebral palsy, mental retardation, and epilepsy.

Speech Disorders. Several studies have been conducted on children with speech disorders. Ferrier (1962), Hallom (1964), and Foster (1963) investigated the ITPA performance of school children with articulation disorders. These studies showed that children with articulation disorders are defective at the automatic level, rather than at the representational level.

Smith and McWilliams (1968) tested a large group of children with cleft palates and cleft lips. They found a significant expressive deficiency in these children in both vocal and motor expression.

Down's Syndrome. Several studies have been conducted on children with Down's Syndrome. McCarthy's (1965) results are typical. In this study McCarthy compared mongoloid and brain-injured children. She found that the automatic level is more defective than the representational level for both Down's Syndrome and brain injured, trainable mentally retarded children, but the difference between these two groups is that the mongoloid children have superior motor expression in relation to their other abilities, and to the motor expression of comparable brain-injured children. This marked ability to motor expression should have some implications for education.

Cerebral Palsy. Another clinical group of handicapped children that has been researched is the cerebral-palsied group. McCarthy (1957) and Myers (1963) conducted studies on spastic and athetoid cerebral-palsied children. Both studies showed the athetoid children to be superior to the spastic children

at the representational level, but inferior to the spastic children at the automatic level.

Mentally Retarded. It had been anticipated that the mentally retarded were primarily defective at the representational level. The profile of McCarthy for trainable children showed a defect at the automatic level. Wiseman (1965) showed similar deficits with educable mentally retarded children.

Epileptics. A recent study by von Isser (1977) found that epileptics have no special disabilities. The composite profile was remarkably normal for the total ITPA and for every subtest.

ITPA Functions Related to Reading

Many researchers have been interested in determining the differences between good and poor readers on some subtests of the ITPA. The research in this area shows contradictory results. The earlier studies were consistent in their results, while the more recent studies are inconsistent.

The first study in this area was conducted with the Experimental Edition by Kass (1962). She used the ITPA and other tests to determine basic differences between good and poor readers in the second grade. Her results showed that poor readers were deficient in tests at the automatic level and in auditory association. The result was confirmed by other studies including those of Ragland (1964) and McLeod (1965).

A study by Macione (1969) using the revised ITPA showed similar results. The ITPA subtests for the second-grade poor readers showed significant deficits at the automatic level. These included grammatic closure, visual closure, visual sequential memory, and sound blending.

An Australian study by Elkins (1972) compared good and poor readers in Grades 1 and 2 on the ITPA. He found the disabled readers to be significantly lower than the good readers in verbal expression, grammatic closure, auditory sequential memory, and sound blending, but not in visual sequential memory.

Recent studies, however, are not showing the same results as earlier studies, nor are they consistent among themselves. Many of the recent studies used third- and fourth-grade children, instead of first- and second-grade children. Results might also be influenced by changes in the emphasis schools are placing on certain learning techniques.

Bartin (1971) studied three groups of third graders—normal-progress readers, mildly disabled readers, and severely disabled readers. He found that the auditory closure and sound blending tests differentiated the groups. He did not obtain results similar to those of Kass (1962), Macione (1969), or McLeod (1965), all of whom found that visual sequential memory differentiated the subjects.

Celebre (1971) also studied a group of older children ranging in age from eight to 10. Fifty-two of them were classified as learning disabled and 52 as normal children. He found significant differences in grammatic closure, audi-

tory association, auditory closure, and sound blending. No differences were found in visual sequential memory. Could the difference between earlier studies and this one be due to the older age of these children or to the changing emphasis in teaching reading?

Deese (1971) studied normal children aged seven to 10 who had been referred to a reading clinic. She found significant differences between successful and unsuccessful readers in grammatic closure, auditory sequential memory, visual sequential memory, and sound blending. One may ask if these clinic children were the severe cases compared to the groups found in schools.

Ikeda (1970) also studied the relationship of reading to the ITPA for third-grade children. No differentiation was found between good, average, or poor readers on the ITPA subtests. Here again we find only slight differences in reading abilities between groups and a possible ceiling effect with eight- and nine-year-old children.

One study by Kirby, Lyle, and Hurbel (1972) dealt with the relationship of reading to the ITPA scores of adult prison inmates. They report a correlation of 0.65 with reading comprehension and 0.70 with reading vocabulary. It seems highly irregular to use the ITPA with adults, even if raw scores are used.

Lagerman (1970) reported another study with third- and fourth-grade children. He used two groups of children—one reading one year below grade level, and one reading at grade level. These two groups were given the ITPA, the WISC, and the California Achievement Test. No differences were found in this group of narrowly differentiated subjects.

In an attempt to determine the psycholinguistic correlates of academic achievement, Hammill, Parker, and Newcomer (1975) tested 137 children in the fourth grade with an average age of nine years and four months. A short form of the ITPA was given, using alternate items from eight of the subtests—basically the even-numbered items, although some arbitrary selection was used. They then gave the California Achievement Test (CAT) and calculated correlations between the subtests of the shortened form of the ITPA and the CAT subtests of reading, language, spelling, and arithmetic. They found (1) 45 out of the 60 correlations were significant, (2) Using a correlation of 0.35 and above (since they were concerned only with the predictive indicators), grammatic closure was the only subtest to show a satisfactory correlation, (3) There was little difference between the correlations of the low, high, and average performers on the CAT except in grammatic closure. They concluded that the ITPA does not predict academic achievement.

In evaluating this study as an example of other similar studies, a number of questions can be asked.

1. *Who administered the tests?* One prerequisite of a reported study is that a clear statement be made of the qualifications of the examiners. No mention is made in this article concerning the qualifications of the examiner or examiners, but in another article by the same authors the testing was done by 67 students (Newcomer and Hammill, 1974). Were these student-practice tests?

With so many examiners of questionable experience, and with the well-known factor of examiner variability, how valid are such results?

2. *How valid are the results of any correlational study using subjects at the upper limits of the norms of the tests used?* In several of the studies the subjects were nine-year-old children in the fourth grade, some of whom inevitably reached the top of the norms, thus depressing the measure of the ability of some. If fourth-grade children are tested on a reading test with a ceiling at fourth grade, of course you would narrow the range of scores since many fourth-grade children can read above that level.

Hammill et al. (1975) recognized the problem of a ceiling, since they state that "it is necessary to demonstrate that ceiling effects are not present" with fourth-grade children. But they brush aside this caution by saying, ". . . no more than the expected number of children reached the test's ceiling." If the expected number of children reached the test's ceiling, that in itself should invalidate the correlations.

3. *How valid are assumptions that correlations based on the short form of the ITPA apply also to the long form?* By using a very questionable short form of the test, Hammill et al. concluded that there is no relationship between academic achievement and most of the subtests of the Illinois Test of Psycholinguistic Abilities. Newcomer and Hammill (1974) published a study in which they tried to correlate the full form of the ITPA with a short form that they had devised. The correlations were surprisingly high, but it was pointed out by Kirk (1975) in the *Journal of Learning Disabilities* that the correlations were spurious since Newcomer and Hammill had not partialled out chronological age. This was an obvious flaw in the study and a criticism with which the authors agreed. It is a little difficult to understand how the authors drew conclusions about the ITPA when the ITPA as standardized was not administered.

4. *Why are these correlations labelled "predictive" instead of "concurrent"?* Both the CAT and the ITPA were given in the fourth grade. Even if they had correlated highly, they would not have been "predictive." This is a concurrent correlation, not a predictive correlation. Such a study would require testing the children in kindergarten, finding the children who show discrepancies in development, and then testing them in the second grade to see if the children with deviations failed in academic subjects. Even this approach would not completely answer the question since intervention could change the picture.

Newcomer and Hammill recognized that few decisive predictive studies have been conducted. They state, "It is particularly unfortunate that we could locate no studies which evaluated the psycholinguistic competencies of preschool children and followed them up through the second or third grade" (Newcomer and Hammill, 1975, p. 736). On the same page, despite their own statement, they mention the Hirshoren (1969) study, which did exactly what they said they could not find. In Hirshoren's study, 41 white children were tested in kindergarten with the Stanford-Binet and the experimental edition of

the ITPA. These children were then followed up two years later in the second grade with the CAT. Hirshoren reported (1) The median correlation between the composite scores of the ITPA given in kindergarten and the achievement variables of the CAT given in the second grade was 0.60. The median correlation between Binet IQ's similarly administered was 0.56. (2) Among the correlations of the subtests of the ITPA, the visual sequential memory test was 0.61 for reading vocabulary and 0.51 for reading comprehension. (3) Auditory association correlated 0.53 with vocabulary, and 0.41 with comprehension.

5. *Why shoot down a straw man?* The authors of the ITPA have never claimed that the test has predictive value for reading. The test was devised to analyze various aspects of the children's communicative abilities. It was an effort to diagnose children's behavior by assessing variations in abilities within the child. It was not organized to predict academic achievement in the three R's in third and fourth grades. Hammill, Parker, and Newcomer (1975), as well as other researchers, have set up a straw man in order to shoot it down.

In addition to the above comments on research with third- and fourth-grade children, other speculations can be made to explain the different results correlating reading ability to the ITPA subtests. Few if any of the studies took into consideration the Aptitude Treatment Interaction Factor. In England this is called the Aptitude X Instruction Interaction Factor. This means that the results that we obtain are dependent, not only on the characteristics of the child, but also on the instruction he receives. The teaching of phonics is a point in question. The earlier studies that showed a strong relationship between poor reading ability and visual sequential memory ability were all done in the sixties when less emphasis was placed on the formal use of phonics in teaching reading. Toward the end of that decade, much greater emphasis began to be placed on phonics until today the State Board of Education in Arizona, for example, has prescribed that all schools must use readers that emphasize phonics. The teaching of phonics emphasizes the sequence of letters and their sounds in a word and in so doing teaches the child to use visual sequential memory in learning to read and spell. The more recent studies have been conducted with children who are more sophisticated in this process. As a result of this, children who would have had weak visual sequential memory may have overcome their potential deficit through the training in phonics by the time they reach third or fourth grade. Cronbach and Snow (1969) report that auditory sequencing correlated positively with "look-say" methods of teaching reading, and negatively with phonic methods.

Another explanation of the differences between the earlier studies and the later ones is that the earlier ones used first- and second-grade children whereas the later ones used third- and four-grade children. Sound blending is important at the early stages of learning to read (Grades 1 and 2) but of much less importance in the later stages as in Grades 3 and 4.

In brief, it is illogical for us to study the relationship between certain characteristics or aptitudes of children and their school achievement without also

evaluating and defining the environment and teaching methods used. The child's experience and environment may also be significant in determining his psycholinguistic abilities.

THE USES OF THE ITPA

The ITPA, as described earlier, was developed to assist clinicians and teachers diagnose visual-motor and auditory-vocal abilities and disabilities in linguistic, cognitive, perceptual, and memory functions of young children. It assesses deviations in functioning within a single child so that remedial programs for the use of these functions can be formulated.

As indicated earlier, the ITPA has been used in situations and under conditions that were not intended by the developers of the test. The test was not designed to solve all problems of all children as some users imply. It was not designed to be used as a classification instrument or as a predictor of academic achievement in the third and fourth grades as a number of studies have attempted to assay. It has repeatedly been emphasized that it is not to be administered by untrained examiners—a procedure that is detrimental to clinical practice and to research results.

The test was designed to be used primarily with young children to obtain clinical insights into those who have communicative problems. Its main function is to help assess discrepancies in cognitive and perceptual functioning, and in some aspects of language and memory performance. Since the latter is the main purpose of the test, a report on one child will be made.

A four-year-old child was diagnosed as having Down's Syndrome since a chromosomal translocation of the type found in children with Down's Syndrome was present. The child's features did not have the usual appearance of mongolism. He did not talk, and on the Stanford-Binet he was said to be untestable, with an IQ below 50. He was considered to be functioning at the trainable retarded level and was assigned to a preschool class for trainable mentally retarded children.

Figure 1, Profile A, shows the profile of this four-year-old on the ITPA. It will be noted that this child scored at or near the five- and six-year levels on some of the visual motor tests (visual reception, visual association, manual expression, and visual closure) but was unable to score on the auditory and verbal tests. He was obviously not mentally retarded, but he was unable to use the auditory-vocal code that others used. His basic need was to learn to talk and to understand speech. It should be noted that this boy's visual-motor channel was superior to his chronological age, whereas his auditory-vocal channel was inferior.

Profile B on the same graph shows a child of five whose IQ was similar to the child with a chromosome aberration. This child's abilities were depressed below his chronological age on all linguistic, perceptual, and memory tests. While the child purported to have Down's Syndrome shows significant dis-

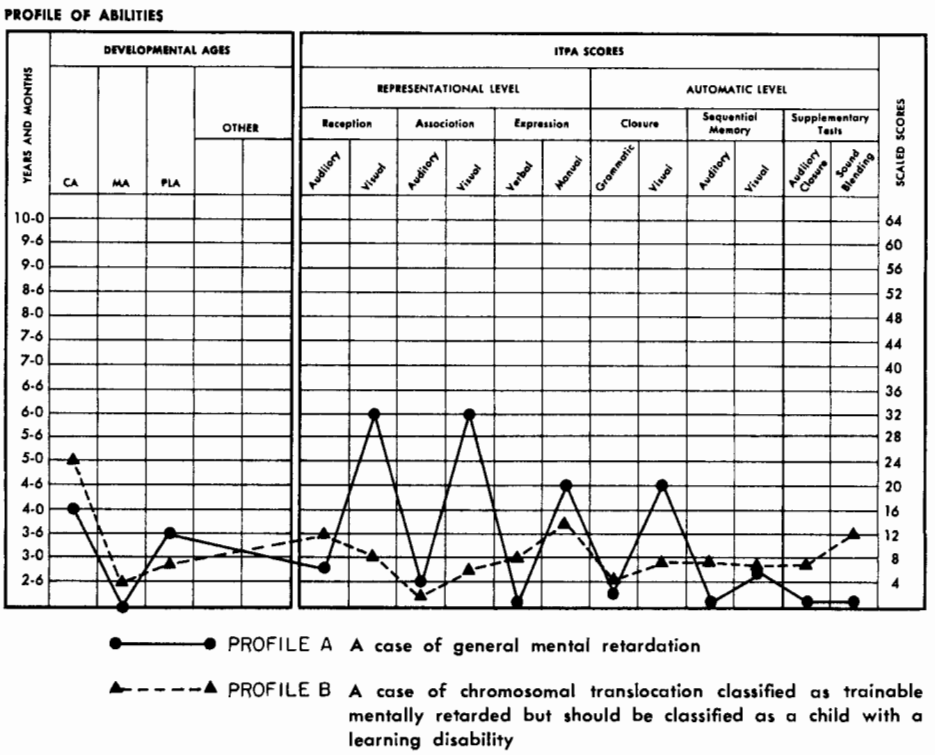

Figure 1. ITPA profiles for two children.

crepancies in abilities, the second child is more evenly retarded and requires a more general curriculum rather than a program of remediation of specific deficits.

What does Profile A tell us? We knew beforehand that the child did not talk. What did the ITPA profile contribute to our knowledge of this child?

1. First, this four-year-old child was not simply mentally retarded as diagnosed by physicians and psychologists on the basis of the chromosomal translocation and the lack of verbal communication. He scored significantly above his chronological age on the visual tests and on manual expression, in spite of low scores on the auditory-vocal subtests and on the intelligence tests. This is one major use of the ITPA, to differentiate between mental retardation, where all the abilities are deficient, and a language-disordered child (learning disability) where some of the abilities are average or superior while the others are depressed.

2. Second, the ITPA profile tells us what strengths can be used to overcome deficits. The visual ability may be used as a bridge to understanding the verbal code, just as motor ability may serve to spark the use of words. We can use the abilities to train the disabilities by organizing a program that will associate doing with saying, and seeing with understanding the words.

CONCLUSION

In conclusion, a few generalizations can be made.

1. The ITPA, with all its limitations (and it has many) has become popular nationally and internationally. Why? Because it is a diagnostic rather than a classification test. If properly administered, it helps pinpoint some abilities and disabilities in children between the ages of two and 10 (or four and eight). This kind of assessment gives badly needed cues to a remedial teacher or clinician in designing a program to help the child develop more effective use of his abilities. The test profile approach helps the teacher in organizing a remedial program by understanding why the child has difficulty in some accomplishments.

2. Children with developmental disabilities of one kind or another show certain kinds of deficits on the ITPA that can be of value in organizing programs for some groups. From the research that has been reported, we can generalize the following:

(a) Many minority children, Mexican Americans, Indians, some blacks, and others who are bilingual are relatively normal on visual-motor psychological functions but deficient in the auditory-vocal functions. Educationally, this means that they need a curriculum that develops success in auditory-vocal interaction between the teacher and the children. Contrary to this need, we find many such children working on visual-motor workbooks in classes because such tasks require less teacher-pupil interaction and it is something the child can do by himself. The auditory-vocal curriculum requires more work on the part of the teacher and a higher ratio of teachers to children.

(b) The mentally retarded, children with articulatory disorders, athetoid cerebral-palsied children, and some other types are more deficient at the automatic level than at the representational level.

(c) Children with severe reading disabilities show varying deficits, probably related to the reading level studied and the teaching method used. Studies show that in the first or second grade, visual sequential memory and sound blending are correlates of reading difficulty. In the third and fourth grades grammatic closure and auditory association are correlated with reading ability. During the early part of the century when phonics was emphasized in teaching reading, the reading clinics were full of children who were word callers and did not derive meaning from reading. Later, when phonics fell into disuse and sight reading was emphasized, the reading clinics were full of children who needed phonics to learn to read. Have we come full circle in our schools? As with many social and educational issues, we allow the pendulum to swing from one extreme to the other.

3. There is no psychometric instrument that is highly reliable and valid. Mechanical use of psychometric tests sometimes tends to lead to misdiagnosis. Tests cannot substitute for good observations and clinical judgments. They can be used as supplements. This statement applies to the ITPA as well as to other tests.

ACKNOWLEDGMENT

This report is a condensed version of two addresses delivered at the forty-ninth and fiftieth Conventions of the American Speech and Hearing Association, November 5, 1974, (Las Vegas) and November 22, 1975 (Washington, D.C.). Requests for reprints should be directed to Samuel A. Kirk, Department of Special Education, University of Arizona, Tucson, Arizona 85721.

REFERENCES

BARTIN, N. G., The intellectual and psycholinguistic characteristics of three groups of differentiated third grade readers. Doctoral dissertation, State Univ. of New York at Buffalo (1971).

BATEMAN, B., *The Illinois Test of Psycholinguistic Abilities in Current Research: Summaries of Studies*. Urbana, Ill.: Univ. of Illinois Press (1965).

BUROS, O. K. (Ed.), *The Seventh Mental Measurements Yearbook*. Highland Park, N.J.: Gryphon, 814–825 (1972).

CELEBRE, G., Psycholinguistic abilities and oral word recognition associated with relative level of personality adjustment in primary school-age children with minimal brain dysfunction. Doctoral dissertation, Temple Univ. (1971).

CICIRELLI, V. G., GRANGER, R., SCHEMMEL, D., COOPER, W., and HOLTHOUSE, N., Performance of disadvantaged primary-grade children on the Revised Illinois Test of Psycholinguistic Abilities. *Psychology Schools*, **8**, 240–246 (1971).

COHEN, A., Smallest space analysis of the Revised Illinois Test of Psycholinguistic Abilities. *Psychology Schools*, **10**, 107–110 (1973).

CRONBACH, L. J., and SNOW, L. S., Final report: Individual differences in learning ability as a function of instructional variables. Contract No. OEC. 4-6-061269-127. Standard Univ.: U.S. Office of Education, School of Education (1969).

DEESE, C., A study of the discrimination by the subtests of the Revised Illinois Test of Psycholinguistic Abilities between successful and unsuccessful readers of normal intelligence. Doctoral dissertation, Memphis State Univ. (1971).

ELKINS, J., Some psycholinguistic aspects of the differential diagnosis of reading disability in grades I and II. Doctoral dissertation, Univ. of Queensland (1972).

FERRIER, E. E., An investigation of the ITPA performance of children with functional defects in articulation. *Exceptl Child.*, **9**, 32 (1962).

FOSTER, S., Language skills for children with persistent articulation disorders. Doctoral dissertation, Texas Woman's Univ. (1963).

GUEST, K. E., Relationships among the ITPA, receptive expressive language tasks, intelligence and achievement. Doctoral dissertation, Univ. of Wisconsin (1970).

HALLOM, J. J., An exploratory study to determine psycholinguistic abilities of a group of six-year-old children with severe articulation problems. Master's thesis, Sacramento State College (1964).

HAMMILL, D. D., PARKER, R., and NEWCOMER, P., Psycholinguistic correlates of academic achievement. *J. Sch. Psychol.*, **13**, 18 (1975).

HARE, B. A., HAMILL, D., and BARTEL, N. R., Construct validity of selected subtests of the ITPA. *Exceptl Child.*, **40**, 13–21 (1973).

HIRSHOREN, A., A comparison of the predictive validity of the Revised Stanford-Binet Intelligence Scale and the Illinois Test of Psycholinguistic Abilities. *Exceptl Child.*, **35**, 517–522 (1969).

HUBSCHMAN, E., POLIZZOTTO, E. A., and KALISKI, M. S., Performance of the institutionalized retardates on the PPVT and two editions of the ITPA. *Am. J. ment. Defcy.*, **74**, 579–580 (1970).

HUIZINGA, R. J., The relationship of the Illinois Test of Psycholinguistic Abilities to the Stanford-Binet Form L-M and Wechsler Intelligence Scale for Children. Doctoral dissertation, Univ. of Arizona (1971).

HUMPHREY, J., and RICE, A., An evaluation of several methods of predicting Full-Scale IQ from the ITPA. *J. spec. Educ.*, **7**, 133–140 (1973).

IKEDA, M., The relationships between the ITPA, reading performance, and IQ of third grade children. Doctoral dissertation, Univ. of New Mexico (1970).

JORSTAD, D., Psycholinguistic learning disabilities in twenty Mexican-American students. *J. learn. Dis.*, **4,** 143–147 (1971).

KASS, C. E., Some psychological correlates of severe reading disability (dyslexia). Doctoral dissertation, Univ. of Illinois (1962).

KIRBY, E. A., LYLE, W., and HURBLE, B. R., Reading and psycholinguistic processes of innate problem readers. *J. learn. Dis.*, **5,** 295–298 (1972).

KIRK, S. A., Ethnic differences in psycholinguistic abilities. *Exceptl Child.*, **39,** 112–118 (1972).

KIRK, S. A., Letters to the editor. *J. learn. Disabil.*, **8,** 119–120 (1975).

KIRK, S. A., and KIRK, W. D., *Psycholinguistic Learning Disabilities: Diagnosis and Remediation.* Urbana, Ill.: Univ. of Illinois Press (1971).

KIRK, S. A., and MCCARTHY, J. J., The Illinois Test of Psycholinguistic Abilities: An approach to differential diagnosis. *Am. J. ment. Defic.*, **66,** 399–412 (1961).

KIRK, S. A., MCCARTHY, J. J., and KIRK, W. D., *The Illinois Test of Psycholinguistic Abilities.* Urbana, Ill.: Univ. of Illinois Press (1961, rev. 1968).

KIRK, W. D., *Aids and Precautions in the Administration of the Illinois Test of Psycholinguistic Abilities.* Urbana, Ill.: Univ. of Illinois Press (1974).

LAGERMAN, A. P., Psycholinguistic characteristics of children with reading disabilities, and the effects of remediation on psycholinguistic development and reading achievement. Doctoral dissertation, Univ. of South Dakota (1970).

LOMBARDI, T. P., *Psycholinguistic abilities of Papago Indian children. Exceptl Child.*, **36,** 485–494 (1970).

MACIONE, J., Psycholinguistic correlates of reading disabilities as defined by the Illinois Test of Psycholinguistic Abilities. Doctoral dissertation, Univ. of South Dakota (1969).

MCCARTHY, J. J., Qualitative and quantitative differences in the language abilities of young cerebral palsied. Doctoral dissertation, Univ. of Illinois (1957).

MCCARTHY, J. M., Patterns of psycholinguistic development of mongoloid and non-mongoloid severely retarded children. Doctoral dissertation, Univ. of Illinois (1965).

MCLEOD, J., Some psychological and psycholinguistic aspects of severe reading disability in children. Doctoral dissertation, Univ. of Queensland (1965).

MYERS, P. A., Comparison of language disabilities of young spastic and athetoid children. Doctoral dissertation, Univ. of Texas (1963).

NEWCOMER, P., and HAMMILL, D., A short form of the Revised Illinois Test of Psycholinguistic Abilities. *J. learn. Disabil.*, **7,** 570–572 (1974).

NEWCOMER, P., and HAMMILL, D., ITPA and academic achievement: A survey. *Reading Tchr.*, **28,** 731–741 (1975).

OSGOOD, C. E., Motivational dynamics of language behavior. *Nebraska Symposium on Motivation.* Lincoln, Neb.: Univ. of Nebraska Press (1957).

PARASKEVOPOULOS, J., and KIRK, S. A., *The Development and Psychometric Characteristics of the Revised Illinois Test of Psycholinguistic Abilities.* Urbana, Ill.: Univ. of Illinois Press (1969).

POLLEY, D., The relationship of the channels of communication of the ITPA and the WISC. Doctoral dissertation, Univ. of Northern Colorado (1971).

PROGER, B. B., CROSS, L. H., and BURGER, R. M., Construct validation of standardized tests in special education: A framework of reference and application to ITPA research (1967-71). In L. Mann and D. Sabatino (Eds.), *First Review of Special Education.* Philadelphia: Buttonwood Farms, 165–201 (1973).

RAGLAND, G. G., The performance of educable mentally handicapped students of differing reading ability on the ITPA. Doctoral dissertation, Univ. of Virginia (1964).

RYCKMAN, D. B., The psychological processes of disadvantaged children. Doctoral dissertation, Univ. of Illinois (1966).

SEDLACK, R. A., and WEENER, P., Review of research on the Illinois Test of Psycholinguistic Abilities. In L. Mann and D. Sabatino (Eds.), *First Review of Special Education.* Philadelphia: Buttonwood Farms, 113–163 (1973).

SMITH, R. M., and MCWILLIAMS, B. J., Psycholinguistic abilities of children with clefts. *Cleft Palate J.*, **5,** 238–249 (1968).

VON ISSER, A., Psycholinguistic abilities in children with epilepsy. *Exceptl Child.*, **43,** 270–275 (1977).

WAUGH, R., Comparison of revised and experimental editions of the Illinois Test of Psycholinguistic Abilities. *J. learn. Dis.*, **63,** 236–238 (1973).

WEPMAN, J. M., JONES, L. V., BOCK, R. D., VAN PELT, D., Studies in aphasia: Background and theoretical formulations. *J. Speech Hearing Dis.*, **25,** 323–332 (1960).

Westinghouse Study: The Impact of Head Start. Vol. 1. Columbus, Ohio: Westinghouse Learning Corp., Ohio Univ. (1969).

WISEMAN, D. E., The effects of an individualized remedial program on mentally retarded children with psycholinguistic disabilities. Doctoral dissertation, Univ. of Illinois (1965).

Received June 6, 1977.
Accepted June 29, 1977.