

 Open access • Journal Article • DOI:10.3102/00346543049004517

Differential Diagnosis—Prescriptive Teaching: A Critical Appraisal — [Source link](#)

Judith A. Arter, Joseph R. Jenkins

Published on: 01 Jan 1979 - Review of Educational Research (Sage PublicationsSage CA: Thousand Oaks, CA)

Topics: Remedial Teaching, Critical appraisal, Remedial education, Teaching method and Special education

Related papers:

- [Substance over Style: Assessing the Efficacy of Modality Testing and Teaching:](#)
- [Examining the Benefits and Prevalence of Modality Considerations in Special Education](#)
- [The Effectiveness of Psycholinguistic Training](#)
- [Anaphora and logical form: on formal meaning representations for natural language](#)
- [The Learning-Disabilities Test Battery: Empirical and Social Issues](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/differential-diagnosis-prescriptive-teaching-a-critical-cv6m6wmdx3>



I L L I N O I S

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

-

PRODUCTION NOTE

University of Illinois at
Urbana-Champaign Library
Large-scale Digitization Project, 2007.

370.152
T2261

311-90

**T
E
C
H
N
I
C
A
L**
**R
E
P
O
R
T
S**

Technical Report No. 80

**DIFFERENTIAL DIAGNOSIS-PRESCRIPTIVE TEACHING:
A CRITICAL APPRAISAL**

Judith A. Arter
Phoenix Union High School System

Joseph R. Jenkins
University of Illinois at Urbana-Champaign


January 1978

Center for the Study of Reading

THE LIBRARY OF THE
OCT 7 1981
UNIVERSITY OF ILLINOIS
AT URBANA-CHAMPAIGN

UNIVERSITY OF ILLINOIS
AT URBANA-CHAMPAIGN
51 Gerty Drive
Champaign, Illinois 61820

The National
Institute of
Education
U.S. Department of
Health, Education and Welfare
Washington, D.C. 20208



BOLT BERANEK AND NEWMAN INC.
50 Moulton Street
Cambridge, Massachusetts 02138

CENTER FOR THE STUDY OF READING

Technical Report No. 80

DIFFERENTIAL DIAGNOSIS-PRESCRIPTIVE TEACHING:
A CRITICAL APPRAISAL

Judith A. Arter
Phoenix Union High School System

Joseph R. Jenkins
University of Illinois at Urbana-Champaign

January 1978

University of Illinois
at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Massachusetts 02138

The research reported herein was supported by the National Institute of Education under Contract No. US-NIE-C-400-76-0116. The authors are grateful to Barbara Wilcox for her editorial comments on an earlier version of this manuscript.

This page is intentionally blank.

Abstract

The dominant instructional model within special education, Differential Diagnosis-Prescriptive Teaching, involves the assessment of psycholinguistic and perceptual motor abilities that are presumed necessary for learning basic academic skills. Based on the differential pattern of ability strengths and weaknesses resulting from this assessment, individual remedial prescriptions are prescribed. In this article five assumptions underlying this model are identified. Also presented is a comprehensive review of research related to each assumption. The findings seriously challenge the model's validity and suggest that continued advocacy of the model cannot be justified. Children do not appear to profit from current applications of Differential Diagnosis-Prescriptive Teaching.

Differential Diagnosis-Prescriptive Teaching:
A Critical Appraisal

The term "differential diagnosis" refers to the process of assessing the learning characteristics of a child so that instruction can be matched to individual learning needs (Kirk & Kirk, 1971, p. 12; Kirk, 1972, p. 7). Although, in theory, this could include any procedure that attempts to delineate a child's specific strengths and weaknesses (Ysseldyke & Salvia, 1974), it has traditionally referred to those practices that attempt to diagnose abilities that presumably are prerequisite for or underlie academic learning (Mann, 1971; Ysseldyke, 1973). Such general psychological processes include auditory abilities (e.g., auditory discrimination and memory), visual abilities (e.g., visual discrimination and spatial relationships), cross sensory perceptual abilities (e.g., auditory-visual integration) and psycholinguistic abilities (e.g., auditory sequential memory and verbal expression). According to this model, failure to master basic academic skills, such as reading, may be traced to impairments in one or more of these underlying processes or abilities. To illustrate, a child who experiences failure in school tasks such as spelling phonetically irregular words, answering sequence questions based on material read, or copying sentences, may be found to suffer from an impairment in a basic process such as visual sequential memory (the ability to order a series of items so as to match a previously given model).

The term "diagnostic-prescriptive teaching," often used in conjunction with Differential Diagnosis, refers to the practice of formulating instructional prescriptions on the basis of differential diagnostic results. These prescriptions generally take one of two forms. In one form differential diagnostic information is used to generate a program to directly remediate an underlying ability weakness. In a second form weak abilities are not remediated; rather, the focus is on academic targets, such as reading or math, for which instructional programs are devised that capitalize upon the child's pattern of underlying strengths and weaknesses, as identified in the course of diagnosis. An example of the former approach would be provision of visual discrimination and visual memory exercises for the child whose diagnosis indicated weaknesses in these areas. An example of the latter strategy would be identification of an auditory or visual learning pattern so that reading instruction could be geared to the stronger modality.

In this paper the terms differential diagnosis and diagnostic-prescriptive teaching are combined as Differential Diagnosis-Prescriptive Teaching (DD-PT) and refer to the psychometric practice of assessing underlying abilities and devising subsequent instruction in accord with ability strengths and weaknesses. Haring and Bateman (1977, p. 130) have described this approach as the "majority position within the field of learning disabilities over the past 20 or 30 years." The DD-PT label encompasses a number of "process" models which are fundamentally equivalent but which have gone by a variety of names. According to Haring

and Bateman, "This conceptualization [the DD-PT model] has been known as the diagnostic-remedial approach (Bateman, 1967), prescriptive teaching (Peter, 1965), ability and process training (Ysseldyke & Salvia, 1974), psychometric phrenology (Mann, 1971) and even task analysis (Johnson, 1967)" (Haring & Bateman, 1977, p. 130).

The Proliferation of DD-PT

In recent years special education has witnessed a proliferation of tests and training programs designed for DD-PT. Sabatino (1973) listed 17 assessment protocols that contain one or more subtests for evaluating auditory perception, along with 16 programs that have as a major goal the remediation of auditory perception. Ysseldyke (1973) and Goodman and Hammill (1973) identified 11 tests designed to assess visual perception, psycholinguistic processes, and motor skills. Keogh (1974) identified 16 authors who have developed visual perceptual training programs. Several factors may account for this proliferation of tests and programs.

Resemblance to task analysis. First, in the DD-PT model, the practice of analyzing academic skills into their components bears a strong resemblance to task analysis. Task analytic approaches to instructional programming are themselves quite popular. In the task analysis model "specific behavioral components are identified and prerequisites for each are determined. The strategy is to develop learning objectives such that mastery of objectives in the hierarchy (simple tasks) facilitates learning of higher objectives (more complex tasks)" (Resneck, Wang & Kaplan, 1973, p. 679). Similarly, the DD-PT model holds that academic tasks must

be analyzed into basic components. Here though, the basic components consist of underlying abilities or psychological processes. If weaknesses are discovered at the foundational or ability level, they must be remediated before proceeding to higher order skills.

Although the task analytic and DD-PT models appear to be similar, we believe that the similarities are quite superficial and that serious differences exist between the two. The differences between the models lie both in their level of analysis and their implications for instruction. With reference to the level of analysis, the task analytic model breaks down larger general tasks into sets of smaller specific tasks. These latter tasks are significant only insofar as they are directly related to the next higher task. In contrast, the DD-PT model analyzes academic tasks into abilities or processes (e.g., visual memory) that are seen as significant for a wide variety of higher level tasks.

With reference to instructional implications from the two models, the task analytic approach maintains that a teacher needs only help the child master specific tasks in the hierarchy that have not been mastered. In the DD-PT model the teacher is faced with a far more serious challenge--to remediate or strengthen an entire process. This requires that the teacher demonstrate improvement or "mastery" of a large number of specific tasks, each of which are thought to depend upon or tap that particular process.

As an illustration of the different instructional implications of these models, suppose that a child encounters difficulty in learning

to count objects. A task analytic teacher might determine that one prerequisite for counting objects is recitation of numerals in order. In contrast, a DD-PT teacher might formulate the same problem more generally as an auditory sequential memory deficit. While the task analytic teacher can satisfy the immediate teaching objective by helping the student learn rote counting, the DD-PT teacher, to satisfy the immediate teaching objective, must improve the child's ability to recite lists of spoken events which are arbitrarily ordered (e.g., color names, animal names, articles of clothing, and perhaps numbers). Thus, in the DD-PT model the teacher is viewed as teaching general abilities; in the task analytic model the teacher is seen as teaching specific components of academic tasks. Clearly, these two models, although they appear to be similar, lead to very different types of instruction.

Pressure to develop effective remediation techniques. Besides DD-PT's resemblance to task analysis, another factor accounting for the proliferation of DD-PT tests and teaching materials has been the pressure felt by special educators to develop effective and innovative remediation techniques. This pressure is due in part to the unflattering outcomes of special education efficacy studies, to the financial expenditures associated with special educational services, and now to regular education's expectations of effective special education contributions in the context of mainstreaming. Older, more global assessment instruments (e.g., IQ tests), although sometimes useful for administrative actions (e.g., placement of children in categorical programs or procurement of state reimbursement for special programs), appear to be inadequate for planning

individualized programs of instruction. Effective individual remediation requires more specific assessment information (Kirk & Kirk, 1971) and differential diagnostic instruments appeared to meet this need.

Needs of early childhood education programs. The DD-PT approach, with its emphasis on psychological abilities, appears also to have benefited from the growth of early childhood education programs. Instruments were needed to structure curriculum and evaluate program effects of such federally funded ventures as Head Start. Since academic skills themselves were not to be taught in these preschool programs, some worthwhile preacademic goals had to be identified. Linguistic and perceptual processes became prime training targets since they were hypothesized to be essential to the future acquisition of academic skills (Sedlack & Weener, 1973).

Differentiation of special and regular education. The DD-PT model appealed to special educators because it served to differentiate their effort from that of regular education. While regular educators concentrated on reading, arithmetic, etc., special educators focused on more basic, underlying processes. This division helped to clarify the respective roles of regular and special educators, and to reduce potential territorial disputes between the parties.

Support from authorities. Support for DD-PT from special education authorities has been strong. The following quotations are illustrative.

The visual dyslexic rarely is able to learn by an ideo-visual approach since he cannot associate words with their meanings.

He cannot retain the visual image of a whole word and consequently needs a more phonetic or elemental approach to reading.

(Johnson & Myklebust, 1967, p. 156)

Many children are coming into our schools lacking in basic perceptual-motor skill....We [need] to help the child to build up the sensory-motor skills which are required by the more complex activities of reading, writing and arithmetic. (Kephart, 1960, p. 16)

The major emphasis [of the concept of learning disabilities] is the use of psychological tests and/or observation for the purpose of organizing...a remedial educational program. Such a program...is very dependent upon the determination of psychological abilities and disabilities. (Kirk & Kirk, 1971, p. 13)

A child's learning type--his maximum modality--needs to be understood before a particular approach to reading can be determined for him....Today this determination can be made with reasonable accuracy. (Wepman, 1967, p. 355)

Encouragement from publishers. Publishers of educational materials have found it lucrative to develop and market an array of ability assessment instruments and related instructional materials. New tests and training materials appear on the market almost daily. School systems invest heavily to purchase DD-PT materials for their special education programs. Special education teacher training programs devote considerable resources to instruction in the philosophy and implementation of DD-PT assessment devices and instructional materials.

We believe that the widespread adoption of DD-PT warrants a critical appraisal of the model's efficacy. To date, experimental studies and reviews of literature in this area have focused on isolated aspects of DD-PT: the present paper is an attempt to consolidate information in order to present a comprehensive picture of the support for and efficacy of the DD-PT model. We will examine the basic assumptions of DD-PT and evaluate the extent to which data from diverse studies support those assumptions. In producing a review of the DD-PT literature we have found it both necessary and valuable to draw extensively on other published reviews of particular aspects of the DD-PT model. There are several reasons for this reliance on secondary sources. First, the literature on various aspects of the DD-PT is so extensive that a comprehensive review of primary sources would be prohibitive. In addition, it would be unwise to ignore the unique and valuable contributions of a number of scholars to the analysis of the DD-PT literature. Finally, many of the existing reviews lend themselves well to our purpose because they present summary data in tabular form, permitting individual reanalysis as warranted. Primary sources were consulted only when the review articles were unclear or contradictory, or when studies were reported after the most recent review.

ASSUMPTIONS BASIC TO THE DD-PT MODEL

Underlying the DD-PT model are several assumptions regarding psychological abilities and their relationship to academic skills, the measurement of these abilities, and their susceptibility to modification through

training. The five basic assumptions of DD-PT that are presented below have been synthesized from various authors: Hammill (1972), Larsen and Hammill (1975), Mann (1971), Proger, Cross, and Burger (1973), and Ysseldyke (1973). Along with each assumption we have attempted to specify the kinds of empirical evidence which would be useful in evaluating their validity.

Assumption 1:

Educationally Relevant Psychological Abilities Exist

and Can Be Measured

The psychological abilities referred to here are not themselves directly observable but must be inferred from behaviors which presumably require the hypothesized ability. These abilities are referred to as underlying in the sense that they are foundational to academic skills. Each ability is general in that it is important for a number of academic behaviors. In another sense, each ability is specific in that its strength may be independent of the strength of other abilities. Thus, for an individual student, it is theoretically possible for visual reception to be strong but for visual memory to be weak.

Assumption 1 would be supported by data documenting the success of the DD-PT model (either in remediating ability deficits or in accelerating academic performance by capitalizing on ability strengths) or by data supporting the reliability or validity claims of the instruments frequently employed by the model. In the absence of such data Assumption 1 could be questioned. However, even without supporting data this assumption

cannot be disproven. It is impossible to prove that educationally relevant underlying abilities do not exist, since if failures to measure or remediate abilities were documented, they could be attributed to the inadequacies of present day measurement and instructional technology. These previous failures would not necessarily reflect upon what could occur in the future through improvements in technology. Nevertheless, such failures as those mentioned above should raise doubts about the validity of Assumption 1. In contrast to Assumption 1 the remaining assumptions have been expressed in terms of existing technology.

Assumption 2:

Existing Tests Used in Differential Diagnosis are Reliable

The DD-PT model assumes that abilities which underlie academic learning are stable, non-ephemeral individual traits. Thus, an instrument which purports to measure these abilities should result in relatively constant scores on repeated testings or on different parts of the same test. If such consistency is not observed, it would naturally raise doubt about the capability of the test to produce reliable information on the target ability. For there to be confidence in the results of differential ability tests, or for that matter, for the DD-PT model to be employed effectively, it is essential that these tests produce a picture of performance that is relatively stable over time.

The evidence relating to the reliability of DD-PT instruments is straightforward, coming from studies of test-retest reliability and studies of internal consistency. The reliabilities associated with a particular

instrument must be high enough to warrant that instrument's use in making educationally significant decisions. Otherwise, efforts to generate precise instructional prescriptions would be misspent.

Assumption 3:

Existing Tests Used for Differential Diagnosis are Valid

Validity refers to the extent to which a test measures what it is supposed to measure. The validity data generated with regard to DD-PT is classified into concurrent and predictive criterion validity, diagnostic validity, and construct validity. Although the information generated by examining one type of validity is corroborative evidence for the others, each type of validity will be discussed separately.

Concurrent Criterion Validity

To assess the concurrent criterion validity of an instrument used in Differential Diagnosis, one determines the extent to which the results obtained with this instrument correlate with measures of academic achievement taken at the same time. Since weak abilities are assumed to impede academic achievement, children who obtain low scores on an ability measure should obtain similarly low scores on measures of academic achievement. Likewise, children obtaining scores indicating an ability strength should, on the average, score high on achievement measures.

Predictive Criterion Validity

For a DD-PT test instrument to have predictive criterion validity, children's scores on it should predict their later academic achievement.

Evidence consists of correlations between an ability measure given at one time and a measure of academic achievement given at a later time. Again, since weak abilities should hinder academic success, children receiving low ability scores should perform poorly on subsequent achievement tests, relative to children receiving high ability scores.

Diagnostic Validity

This type of validity is similar to the concurrent criterion type, except that the procedures for estimating validity differ. Whereas investigations of concurrent validity correlate performance on ability and achievement measures, studies of diagnostic validity group students according to their performance on one measure (ability or achievement) and then examine their performance on the other measure. To illustrate, an ability assessment device is said to have diagnostic validity if children who are independently identified as poor readers via an achievement test also perform significantly worse on that ability measure than children identified as good readers.

Construct Validity

In assessing the construct validity of DD tests, attention is directed to the theoretical model upon which the tests are based. Construct validity is the degree to which the test measures an hypothetical variable. Thus, "construct validation requires the gradual accumulation of information from a variety of sources. Any data throwing light on the nature of the trait under consideration and the conditions affecting its development and manifestation are grist for

this validity mill" (Anatasi, 1968, p. 115). In the broadest sense, all of the previously discussed types of validity can be viewed as evidence relating to construct validity. In our examination of construct validity we focus on factor analytic studies that investigate the independence of various subtests used in differential diagnosis, (e.g., the 11 subtests of the Illinois Test of Psycholinguistic Abilities [ITPA] and the five subtests of the Developmental Test of Visual Perception [DTVP]). Since the DD-PT model assumes that subtests of a larger assessment device provide information that is crucial for instructional programming, it is important to demonstrate that these subtests tap different abilities.

Assumptions 4A and 4B:

4A: Prescriptions Can Be Generated from Differential Diagnosis
to Remediate Weak Abilities

4B: Remediation of Weak Abilities Improves Academic Achievement

According to the DD-PT model, failure to acquire academic skills is the result of one or more underlying ability deficits. Applications of the model may take two forms. Assumptions 4A and B relate to the first form which involves the direct training of weak abilities with the intention of strengthening them. This accomplished, the impediment to academic achievement is removed, and the child's progress can be expected to accelerate. This assumption would be supported by evidence that ability training strengthens weak abilities and enhances academic performance.

Assumption 5:Prescriptions Can Be Generated from Ability Profiles to Improve Academic Achievement, With No Direct Training of Weak Abilities

This assumption, which describes the second application of the DD-PT model, is really a combination of two assumptions: Not all children learn best under a single instructional approach. And, secondly, the approach which will maximize the child's educational progress is best identified from the child's profile of ability strengths and weaknesses, as determined by differential diagnosis. The most common implementation of the DD-PT model in this form involves the matching of instructional materials and methodologies to children's modality strengths (visual, auditory or kinesthetic). Evidence supporting this assumption would consist of research which indicates that designing instruction in accord with modality strength and weakness leads to more significant educational gains than does instruction which does not incorporate such modality-program matching.

EVIDENCE RELATING TO THE FIVE ASSUMPTIONS OF DD-PT

In this section the evidence relating to Assumptions 2-5 is presented. As mentioned earlier, Assumption 1 can best be evaluated by considering the empirical support for the remaining four assumptions.

Assumption 2:Reliabilities of DD-PT Assessment Instruments

A number of tests have been used for differential diagnosis. Among those most frequently used are the Bender-Gestalt (Bender, 1938), the

DTVP (Frostig, 1963), the ITPA (Kirk, McCarthy, & Kirk, 1968) and the Purdue Perceptual Motor Survey (Roach & Kephart, 1966). Other less frequently used tests include the Auditory Discrimination Test (Wepman, 1973), the Benton Revised Visual Retention Test (Benton, 1955), the Dennis Visual-Perceptual Scale (Dennis & Dennis, 1969), the Developmental Test of Visual Motor Integration (Beery and Buktenica, 1967), the Goldman-Fristoe-Woodcock Test of Auditory Discrimination (Goldman, Fristoe, & Woodcock, 1970), Memory for Designs (Graham & Dendall, 1960), Primary Visual Motor Test (Haworth, 1970) and the Screening Test for Auditory Perception (Kimmell & Wahl, 1969).

In summarizing the reliabilities of these instruments we relied on reviews by Hammill and Wiederholt (1973), Sedlack and Weener (1973), Waugh (1975), and Ysseldyke (1973). In addition, test manuals were consulted in an effort to obtain reliability information on DD-PT instruments that were not included in previous reviews.

Generally, two kinds of test reliability are reported. Test-retest reliability measures the stability of scores over time; it is obtained by administering the same form of the test on two occasions and correlating the scores from each testing. Split-half reliability, a measure of the internal consistency of a test, is determined by dividing the items in the test into two groups and then correlating the scores obtained on each half.

Various authors have discussed the reliability levels considered necessary for a test to be useful. Anastasi (1968) proposed that test

reliabilities need be above .80. Nunnally (1967), on the other hand, suggests that the minimum reliability level should be determined according to the purpose for which the test is employed. For instruments used in basic research, minimum reliability is .80, but for instruments upon which important educational decisions are based, reliabilities should be greater than .90 and preferably above .95. Since DD-PT tests are employed essentially for educational decision making, we have adopted .85 as a minimum reliability level and .90 as a desired level--a compromise between the two recommendations. (Hammill & Wiederholt, 1973).

ITPA

Sedlack and Weener (1973), Waugh (1975), and Ysseldyke (1973) review studies which report reliabilities for the ITPA. The test-retest reliability for ITPA Total Score ranges from .66 to .95. The number of coefficients was not reported in all the reviews, so the median value is not calculable. Subtest reliabilities are even more variable, with Visual Sequential Memory yielding the lowest coefficients (.12-.71) and Auditory Association the highest (.62-.90). The overall median of subtest reliabilities is .71 (Sedlack & Weener, 1973). While ITPA Total Score reliability is acceptable, the subtest reliabilities are not, especially with test-retest intervals of greater than six months (Waugh, 1975). In contrast, split-half reliabilities for the ITPA are generally satisfactory. All are above .85, except Visual Closure which ranges from .67 to .83.

DTVP

Overall, test-retest reliabilities for the DTVP have ranged from .29 to .98 depending upon the subtest examined and grade level of the children tested (Hammill & Wiederholt, 1973; Ysseldyke, 1973). Total Score reliability ranges from .69-.98 (median = .79), and subtest reliabilities range from .29 to .80. The most reliable subtest is Form Constancy (.67-.80), and the least reliable is Eye-Hand Coordination (.29-.42).

As with the ITPA, split-half reliabilities are higher than test-retest reliabilities. Overall, the split-half reliability for Total Score has ranged from .78 to .89. Individual subtests range from .35 to .96 depending on the subtest and the age of the children tested. The most consistent subtest is Figure-Ground (.91-.96), and the least consistent is Spatial Relationships (.52-.67).

Other Visual Tests

Of the seven other predominantly visual perception tests identified by Ysseldyke (1973) or by us, the Purdue yielded the highest test-retest reliability coefficient (.95). The Memory for Designs (median = .87) and the Benton (.85) also met the minimum standard. The remaining instruments either fell below the minimum reliability level: Chicago Test of Visual Discrimination (.35-.68), Developmental Test of Visual-Motor Integration (median .83) and the Primary Visual Motor Test (.82), or failed to report any reliability level (Dennis Visual-Perceptual Scale). Again, split-half reliabilities tended to be higher than test-retest.

Other Auditory Tests

Since we could locate no reviews of the reliability of auditory perceptual tests, we examined selected test manuals. Two of three tests which focus on auditory perception yielded acceptable reliabilities: the Auditory Discrimination Test (.91) and the Goldman (.81-.87). The third instrument, the Screening Test for Auditory Perception, reported no reliability estimates.

Summary of Reliability Evidence

Although split-half reliabilities are generally within the acceptable range, we must concur with Hammill and Wiederholt (1973) that the test-retest reliabilities of most DD-PT instruments are unacceptable. Possible exceptions include the Auditory Discrimination Test, Goldman, Memory for Designs and Benton. Only the Auditory Discrimination Test and the Purdue Perceptual Motor Survey report test-retest reliabilities above .90. The latter estimate is based on only one study which employed 30 children tested at a one week interval.

Low test reliabilities cannot be taken casually. Sedlack and Weener (1973) dramatize the consequences of the " $r = .71$ " coefficient of the ITPA subtests:

Suppose that the bottom 30 percent of the first graders in a school is selected for a special remediation program based on their September score on a particular ITPA subtest; how many of this group would be selected for the program based on retesting five months later in February? Sixty-three percent of the group selected

in September would also be selected in the February testing, but 37 percent diagnosed as "special" in September would be classified as "regular" in the February testing. More than one out of three of the judgments made of the first testing would be considered errors, on the basis of the retest which correlates .70 (p. 117)

The reliabilities associated with many of the popular DD-PT instruments are too low to justify confidence in these measures. While it is possible that isolated subtests may be more reliable for specific populations, it has yet to be demonstrated.

Assumption 3:

Validity of Test Results

Validity involves the extent to which a test measures what it is intended to measure. Research related to DD-PT for each type of validity described under Assumption 3 is discussed below.

Concurrent Criterion Validity

Concurrent criterion validity is studied by correlating performance on two or more tests which were given at approximately the same time. A test is considered to be concurrently valid if it is highly related to other criterion measures to which, in theory, it should be related. Since DD-PT tests are assumed to measure abilities that are crucial for academic success, the "other criterion measures" used to determine their concurrent criterion validity are measures of academic achievement.

To evaluate the correlational evidence, a criterion of acceptability must be established. Mere statistical significance is not sufficient

evidence for validity. Any size correlation, no matter how small, can be statistically significant if a large enough sample is employed. The guidelines for determining how large correlations must be to satisfy validity requirements are based upon recommendations by Guilford (1956), and Garrett (1954): a correlation coefficient of .35 is set as the minimum acceptable cut off point as evidence of adequate criterion validity, while a coefficient of .30 is considered marginally useful. In examining the evidence on validity it is important to consider those studies that control for extraneous variables (e.g., intelligence) which spuriously inflate correlations between specific ability measures and achievement. We have tried to report separately studies in which IQ was controlled. A discussion of this problem is presented in the section summarizing the validity research.

The strategy in summarizing validity studies was to consider first the most comprehensive reviews in each area (psycholinguistic, visual perception, and auditory perception), and subsequently to examine studies which were not included in the more extensive reviews. In most studies reading achievement was the criterion measure, with arithmetic, spelling, and science achievement examined less frequently. Correlations between DD-PT assessments and intelligence and other perceptual motor tests are considered in the section on construct validity.

Psycholinguistic abilities. Five reviewers focus on the relationship between academic achievement and the ITPA (Haring & Bateman, 1977; Larsen & Hammill, 1975; Newcomer & Hammill, 1975; Proger et al., 1973;

and Sedlack & Weener, 1973). Newcomer and Hammill's review is the most comprehensive, reporting 1152 separate correlation coefficients taken from 24 studies (see Table 1). They located 820 correlations between ITPA scores and reading performance. Considering the 12 ITPA subtests and Total Score, only Auditory Association, Grammatic Closure, Sound Blending and Total Score showed median correlations equal to or in excess of .35. "The other subtests, including all those measuring visual processing skills, yielded coefficients which are either not statistically significant or are so low as to have little practical value" (Newcomer & Hammill, 1975, p. 734). In the five studies which partialled

Insert Table 1 about here

out intelligence, only Grammatic Closure ($r = .38$) survived as a useful predictor of reading achievement. A total of 178 correlations were reported between ITPA performance and spelling. Again, only Grammatic Closure yielded an adequate correlation (.41), and even it failed to meet criterion in the one study which controlled for intelligence. Of the 154 correlations between ITPA scores and arithmetic performance, only Grammatic Closure, Auditory Association and Total Score achieved correlations above .35. In the one study which controlled for intelligence no correlations reached the minimum criterion.

In examining the four other reviews of the ITPA only two additional studies relating to concurrent validity were found. Primary sources were consulted in both of those cases. Lovell and Gorton (1968) reported

correlations of .48 and .21 between ITPA Total Score and reading age for "backward" and normal readers, respectively. They did not report results for ITPA subtests and reading achievement. Cicirelli, Granger, Schemmel, Cooper, and Holthouse (1971) found that Auditory Reception, Auditory Association, Grammatic Closure, and Total Score were most highly correlated with scores on the Metropolitan Readiness Test (MRT), $r = .60$.

Results of investigations on the concurrent criterion validity of the ITPA indicate that while individual investigators sometimes reported satisfactory validity coefficients the preponderant finding (Newcomer & Hammill, 1975) is that only Grammatic Closure, Sound Blending, Auditory Association and the Total Score are useful concurrent correlates of achievement. Of these, only Grammatic Closure achieved concurrent validity when investigations controlled for IQ.

Visual perception. In this section evidence is presented on the concurrent validity of several tests of visual perception (e.g., the DTVP), as well as on specific visual perceptual abilities as measured by different instruments.

Developmental test of visual perception. Nine studies investigating the relation between performance on the DTVP and academic performance were reviewed by Hammill and Wiederholt (1973) and Larsen and Hammill (1975). These studies reported a total of 204 correlation coefficients. A summary of those results, reported by subtest, appears in Table 2.

Insert Table 2 about here

All subtests except Eye-Hand Coordination are useful estimates of arithmetic performance. With respect to reading, however, no DTVP subtest is concurrently valid. The DTVP Total Score may be more highly correlated with reading achievement at lower grade levels. Based on the correlations from five studies summarized in Hammill and Wiederholt's earlier review (1973), the average correlation between reading achievement and DTVP Total Score was .39 for first graders, .34 for the second graders, and .19 for third graders. Thus, concurrent validity of the DTVP may be dependent upon the ages and experiences of the children tested.

Other visual perceptual tests. Larsen and Hammill (1975) report the results of 11 studies using the Bender-Gestalt, 12 studies using Birch-Belmont-like assessments, six studies using subtests of the Wechsler Intelligence Scales for Children (WISC), and five using the matching subtest of the MRT (see Table 2). These tests appear to hold little or no promise as concurrent estimates of either reading or arithmetic achievement.

Specific visual perceptual abilities. Larsen and Hammill (1975) summarized the relationships between academic performance (collapsed across reading, spelling, arithmetic, and readiness) and four specific visual perceptual abilities as measured by a variety of instruments (Chicago Test of Visual Discrimination, MRT, Perceptual Achievement Forms Test, WISC, Memory for Designs, Birch-Belmont-like instruments, Bender-Gestalt, ITPA & the DTVP). A number of standardized tests served as

achievement measures (e.g., Metropolitan Achievement Test and the Wide Range Achievement Test). Specific abilities considered were visual discrimination, visual memory, spatial relations, and auditory-visual integration. They reviewed 60 studies that included 600 individual correlation coefficients (see Table 3). Inspection of the correlations shows rather clearly that none of these visual perceptual abilities were a valid indicator of academic achievement.

Insert Table 3 about here

Thirteen studies not appearing in Larsen and Hammill's review were located in reviews by Silverston and Deichman (1975), Ysseldyke (1973), Hammill (1972), and Sabatino (1973), and by an additional library search. We obtained primary sources for seven of these studies.¹ For these seven, median correlation coefficients were computed between visual abilities (visual discrimination, audio-visual integration, copying and visual memory) and achievement (comprehension, vocabulary, arithmetic, spelling, and writing). Using the .35 cutoff, visual discrimination (42 coefficients) evidenced concurrent validity with regard to reading comprehension (but not word recognition), spelling and writing; copying subtests (5 coefficients) appeared to be valid estimates of reading comprehension and writing achievement; and audio-visual integration subtests (16 coefficients) were valid estimates for reading comprehension. While these studies reported correlations somewhat higher than those reviewed by Larsen and Hammill

(1975), it is clear that their addition to the 60 studies contained in that review would not raise the median correlations enough to satisfy validity requirements. Moreover, none of the seven studies partialled out IQ.

We were unable to procure primary sources for the remaining six studies.² According to other reviewers, however, these studies tended to report coefficients that would not satisfy minimum validity standards.

Auditory perception. Hammill (1972), Hammill and Larsen (1974b), Haring and Bateman (1977), Sabatino (1973), and Silverston and Deichman (1975) have reviewed studies of the relationship between auditory abilities and academic achievement. Of these, Hammill and Larsen provided the most comprehensive account, reviewing 30 studies that contained a total of 279 correlation coefficients. Since four of these studies were longitudinal, they will be considered in the section on predictive validity.

Auditory perception tests. Of the frequently used auditory perception tests (Auditory Discrimination Test, the Birch-Belmont, Detroit: Attention for Related Syllables, ITPA: Auditory Sequential Memory, Roswell-Chall: Auditory Blending Test, Seashore, and WISC: Digit Span) only two were correlated greater than .35 with reading: Roswell-Chall with general reading and Birch-Belmont with comprehension (Hammill & Larsen, 1974b). In neither case was intelligence partialled out.

Specific auditory perceptual abilities. Median correlations between five auditory abilities (auditory-visual integration, sound

blending, auditory memory, phonetic auditory discrimination and non-phonetic auditory discrimination) and general reading ability were reported by Hammill and Larsen (1974b). When correlations were corrected for intelligence (as was done in 63 of the coefficients) no auditory abilities provided useful concurrent estimates of general reading skill. This was also true when the correlations were broken down by grade level. When intelligence was not partialled out, sound blending was correlated above the .35 cutoff with overall reading achievement. Table 4 reports a breakdown of the relationship between the five auditory abilities and two reading components: word recognition and comprehension. As can be seen, only auditory-visual integration achieved a median correlation of greater than .35 with a reading subskill.

Insert Table 4 about here

Only three studies were located which did not appear in Hammill and Larsen (1974b). Lovell and Gorton (1968) reported correlations of .38 and .13 between auditory discrimination and reading for "backward" and normal readers, respectively. In 1970, Zunif and Carsen (described in Haring & Bateman, 1977) reported correlations of .54 and .58 between two rhythm tests and reading. These same tests correlated only .03 and .07 with arithmetic. In neither of these studies was intelligence partialled out. Finally, Sabatino reported that auditory integration (blending) has been shown by Alshan (1965) to "predict reading achievement in restricted populations" (1973, p. 65).

Taken together, studies investigating the relationship between measures of auditory perceptual abilities and academic achievement fail to demonstrate the concurrent validity of the auditory measures. Lack of evidence for concurrent validity is particularly striking within those studies that control for IQ.

Summary of concurrent validity. Results of this review indicate that the ITPA has not proven to be concurrently valid with respect to academic achievement; only Grammatic Closure and the Total Score correlate satisfactorily with academic skills when IQ is controlled. Visual perception as measured by the DTVP may adequately predict arithmetic but not reading performance. Other visual perceptual tests produce similarly disappointing results. Likewise, instruments which assess auditory perception tend not to serve as concurrently valid estimates of reading or arithmetic achievement. This picture remains consistent when one considers specific visual and auditory abilities across tests. Except for sound blending ability, which correlated with reading achievement at the secondary level, and audio-visual integration, which correlated with reading comprehension, the nine specific abilities studied are not valid estimators of academic achievement. The exceptions may also be suspect since IQ was not partialled out of the correlations. It is difficult to escape the general conclusion that measurement devices traditionally used in differential diagnosis lack concurrent criterion validity with respect to academic skills.

Predictive Criterion Validity

While concurrent validity is examined by correlating ability and achievement tests given at the same point in time, predictive validity is determined by correlating psychological abilities measured at one point in time (often kindergarten) with later academic achievement. Correlations above the .35 criterion would raise confidence in the predictive validity of ability measures. Barrett (1965a), Hammill and Larsen (1974b), Larsen and Hammill (1975), Newcomer and Hammill (1975), Sedlack and Weener (1973), and Silverston and Deichman (1975) together examined 29 studies of predictive validity. The interval between ability and achievement testing in these studies ranged from three months to three years.

Psycholinguistic abilities. Four studies³ were reviewed by Newcomer and Hammill (1975) in which ITPA performance was correlated with achievement measures taken at least nine months later. In general, ITPA Total Score was a useful predictor of general achievement.

As for the ITPA subtests, Auditory Association consistently predicted reading achievement across studies. Failure to partial out IQ may account for this finding, given that Auditory Association appeared to be a concurrently valid estimate of reading except in studies which controlled for IQ. The picture of predictive validity is unclear for other ITPA subtests and achievement measures. Results from the various studies are in conflict. Hirshoren (1969) found six of nine subtests predictive of spelling performance, but this was not replicated in the Westinghouse (1969) study. Similar disparities have been noted when arithmetic was

the object of prediction. Hirshoren reported that every ITPA subtest usefully predicted arithmetic achievement; Mueller (1969), on the other hand, found no significant predictors among the subtests. In none of the studies was the contribution of intelligence controlled. This, along with the fact that so few studies have been performed make it extremely difficult to evaluate the predictive validity of various ITPA subtests.

Visual perception. Five studies⁴ which investigated the predictive validity of the Bender for academic achievement appeared in reviews by Larsen and Hammill (1975) and Newcomer and Hammill (1975). Only reading comprehension had five or more reported correlations with the Bender. These ranged from .17 to .51 with a median of .28. The 15 coefficients presented for all types of achievement ranged from .16 to .51, with a median of .33. It appeared that the Bender predicted arithmetic and spelling performances better than it did reading. In only one of these investigations was intelligence partialled out. Keogh (1965) reported that correlations between the Bender and reading achievement became insignificant when this was done. Keogh expressed the problems of predictive validity for the Bender.

Cutoff scores defining good Bender performance at kindergarten or third grade clearly identified successful readers; poor Bender performance at either grade was nondefinitive for individual prediction....although correlations between the Bender and reading criteria were of statistical significance, magnitudes were too small to allow confident individual prediction of reading from

the Bender. Individual design interpretation...for differential diagnosis [was] not supported. (1965; p. 83)

Eighteen studies⁵ involving 112 correlations pertaining to the predictive validity of miscellaneous visual perceptual tests were included in the reviews by Barrett (1965a), Larsen and Hammill (1975), Newcomer and Hammill (1975), and Silverston and Deichman (1975). Table 5 presents our calculation of median correlations from these studies between four visual abilities (visual discrimination, auditory-visual integration, gross and fine motor movement, and laterality and body image) and various achievement areas. The only correlation that meets

Insert Table 5 about here

validity standards is between visual discrimination and reading comprehension. Since there were only three reported correlations between auditory-visual integration and academic performance, it is difficult to evaluate the predictive validity of this ability. Intelligence was not partialled out of these correlations.

Auditory perception. From reviews by Hammill and Larsen (1974b), Larsen and Hammill (1975), and Newcomer and Hammill (1975), four primary sources⁶ were located which dealt with the predictive validity of auditory perceptual abilities for later achievement. A total of 26 separate correlations yielded a median correlation of .38. Only the relationship of auditory blending and auditory discrimination to reading

comprehension were examined in enough analyses to be summarized separately. Median correlations were .50 (8 coefficients) and .37 (5 coefficients) respectively. The highest correlations were reported for reading and the lowest for writing (although the latter is based on only a few coefficients). In general, these studies suggest that auditory ability measures may have satisfactory predictive validity. However, some caution should be exercised since intelligence was not partialled out in any of the studies.

Summary of predictive validity. ITPA Total Score, certain ITPA subtests, visual discrimination, auditory discrimination and auditory blending all appear to be correlated with various academic skills beyond the .35 level. It is difficult, however, to draw firm conclusions regarding the predictive validity of various DD-PT measures since so few longitudinal studies have investigated the relationship between specific abilities and specific academic skills.

Caution is warranted in interpreting the correlation coefficients presented in the concurrent and predictive criterion validity sections. Correlation coefficients between tests can be influenced by many factors, not all of which are related to the true relationship between the measures (Proger et al., 1973; Ysseldyke, 1973). Spuriously low or high correlations could result from several of the following conditions. First, there tends to be much common variation in correlation coefficients. A specific ability may appear to correlate highly with reading achievement, but this correlation could be the result of some other factor that is

being inadvertently measured by the ability test. Since intelligence appears to be correlated substantially, both with reading and with underlying perceptual abilities, the correlation between reading and a perceptual ability could be due to the common component of intelligence. Second, correlations between an ability test and a criterion could be low not because the ability test is invalid, but because the criterion to which it is compared is itself invalid. This should not be a major problem for the studies reported in this section, since the criteria are generally measures of academic achievement. Third, coefficients are less reliable when based on small samples. With small samples reported correlations may be spuriously high, or drastically underestimated. Generally, information on sample size was omitted in the reviews. In addition, correlations tend to be deflated when samples are drawn from highly homogenous groups with a restricted range of ability. Again, the reviews tended not to supply specific information on the characteristics of the research population and sampling procedures. For these reasons we believe that examination of several studies gives a more accurate picture of DD-PT instrument validity than does any single study by itself. Finally, the validity of any test is limited by its reliability. This fact is particularly problematic for the DD-PT model since many of the test instruments suffer significant reliability deficiencies as noted earlier. Without satisfactory reliabilities, one cannot hope to demonstrate satisfactory validity.

Diagnostic Validity

The diagnostic validity of a test is its ability to discriminate between groups which are known to differ on some other variable, such as reading skill, presence of a learning disability, race, or socio-economic status. Typically, the diagnostic validity of ability tests has been studied by determining whether or not an ability test differentiates between good readers and poor readers. Diagnostic validity is an important consideration because many of the DD-PT tests and subtests are used to classify children as educable mentally retarded, learning disabled, neurologically impaired, educationally handicapped, etc. Such classifications can have rather dramatic effects on a child's life, possibly disqualifying the child from receiving special education services or, alternatively, resulting in placement in a restrictive setting, e.g., a special class.

Psycholinguistic abilities. Three reviews of diagnostic validity concern themselves exclusively with the ITPA (Newcomer & Hammill, 1975; Proger et al., 1973; and Sedlack & Weener, 1973). Newcomer and Hammill (1975) summarized 24 studies that attempted to determine whether ITPA performance distinguished between good and poor readers. Table 6 summarizes the percentage of analyses in which subtests successfully differentiated groups of readers. In the 14 studies which did not

Insert Table 6 about here

control for intelligence, only Grammatic Closure and Sound Blending were successful in differentiating between groups of readers in more than one-half of the studies. In the studies that controlled for intelligence no subtest differentiated between groups more than 33% of the time.

Proger et al. (1973) reviewed two studies which were not reported in the other reviews (Dugger, 1969; Gaskins, 1971). Both studies examined the power of the ITPA to distinguish good and poor readers. Neither study reported significant differences between good and poor readers on any portion of the ITPA. In a recent study Larsen, Rogers, and Sowell (1976) compared learning disabled children who also had reading deficits with non handicapped normal readers on three subtests of the ITPA (Visual and Auditory Sequential Memory and Sound Blending). They could detect no differences between groups.

Visual perceptual abilities. Studies which attempted to assess the diagnostic validity of various measures of visual perception have been identified by Hammill (1972), Hammill and Larsen (1974b), Larsen and Hammill (1975), Sabatino (1973), Silverston and Deichman (1975), and Ysseldyke (1973). In all, 16 studies⁷ were located and consulted as primary sources (see Table 7).

Of the 37 comparisons which attempted to differentiate between good and poor readers, 19 (51%) were successful. This percentage is reduced to 32% in the studies which controlled for intelligence.

Insert Table 7 about here

Analysis by subskill revealed that the majority of significant differences are accounted for by auditory-visual integration measures. Overall, neither visual discrimination, visual memory, nor miscellaneous visual ability measures distinguished between good and poor readers more than 36% of the time.

Auditory perception. Diagnostic validity of auditory ability tests has been reviewed by Hammill (1972), Hammill and Larsen (1974b), and Sabatino (1973). Many of the studies were not described in detail in these reviews; thus, primary sources were consulted whenever possible.

Fifteen studies⁸ examined the diagnostic validity of various auditory abilities for differentiating good from poor readers. Overall, 88% of the auditory perception measures (35/40) showed significant differences between groups of readers. Even with intelligence controlled, six of eight measures (75%) were significant (see Table 7). Larsen et al., (1976) was the single study which attempted to verify the diagnostic validity of an auditory test on a population other than good and poor readers. When they compared learning disabled and normal children on the Auditory Discrimination Test they found no differences.

Summary of diagnostic validity. Neither subtests of the ITPA nor most visual perceptual abilities (except auditory-visual integration) possess satisfactory diagnostic validity for reading. By contrast, auditory measures demonstrate good diagnostic validity for reading. Many instruments frequently used in differential diagnosis have yet to

be examined for diagnostic validity in academic achievement areas other than reading. Except for auditory tests, the diagnostic validity of differential diagnostic instruments is inadequate.

Construct Validity

The construct validity of a test can be viewed as the degree to which the test delineates the dimensions of the theoretical model on which it is based. Proger et al., (1973) describe a construct as an abstract variable, delineated by an individual theoretician, which represents an hypothesis of how a variety of behaviors will correlate with one another. Evidence relating to other types of validity (e.g., concurrent) are relevant to construct validity. However, the present discussion of construct validity is limited to factor analytic studies that describe the psychological dimensions of a test, and to concurrent criterion validity studies which relate other hypothetical constructs, such as intelligence and perceptual motor ability, to the test in question.

ITPA. The ITPA is composed of 12 subtests which attempt to assess language performance at two levels (representational and automatic), two channels (auditory-vocal and visual-motor), and three processes (reception, association, and expression). If the theoretical model underlying the ITPA is valid, then factor analytic studies should distinguish performance along the levels, channels, and processes. Factor analytic studies of the ITPA have been reviewed by Proger et al. (1973), Sedlack and Weener (1973), Meyers (1969), and Waugh (1975).

Conclusions regarding the construct validity of the ITPA differ depending upon how individual investigators conducted their factor analyses. Early studies examined factor groupings of just the ITPA subtests without employing external criteria. Proger (1973) summarized nine such studies as indicating that the factor analytic structure of the ITPA is much simpler than the model on which it is based. Besides a general linguistic factor that is similar to a general intelligence factor, only the channels dimension seemed to be clearly distinct. Sedlack and Weener (1973), who consider 20 factor analytic studies, are even less positive:

The tentative factors that have been identified in factor analytic studies offer scant support for the channel-level-process model on which the ITPA is based....it is difficult to say what kind of factor structure one would predict, based upon the theoretical model of the ITPA...[since] factors which would honor process distinctions would cut across channels and levels, etc. (p. 124)

Waugh (1975) reached an opposite conclusion, that there was indeed empirical support for the ITPA model. Her review included three recent studies which employed not only the ITPA but other reference tests designed to measure the same traits as the ITPA.

Waugh's analysis is supported by Newcomer, Hare, Hammill, and McGettigan (1974), a primary source not included in previous reviews. They factor analyzed the ITPA with twenty criterion tests judged to parallel the functions measured by the ITPA, and found ten factors

accounting for 66% of the total variance. According to these data, the ITPA subtests appear to measure independent abilities with the exception of Visual Sequential Memory, Visual Reception and Auditory Reception. They also found support for the level and process dimensions, but not for visual and auditory channels. If the "visual" tests do not measure a unique function, this could explain why the auditory tests of the ITPA relate higher than do the visual test to measures of academic performance.

A test's construct validity can also be studied by examining the pattern of correlations between it and other "theoretically related" performances. For example, if ITPA subtests correlate highly with criterion measures to which they in theory should correlate (e.g., Auditory Reception with reading comprehension), yet have near zero correlations with other criterion measures from which they should in theory be independent (e.g., Auditory Reception with arithmetic computation), then this would be interpreted as support for the constructs that the subtests claim to measure (Campbell & Fiske, 1959). Sedlack and Weener (1973), reviewing studies that followed this strategy, conclude that "findings from each of these studies were quite mixed, and none showed expected relationships (or non-relationships) between the criterion variables and the ITPA" (p. 123).

Proger et al. (1973) reported data from 24 studies on the relationships of various measures of intelligence and language scores to the ITPA. Twenty-five correlations between intelligence subtests and the ITPA subtests ranged from .14 to .83, with a median of .50. This supports

the general conclusion of other reviewers that intelligence may account for a large portion of the variance in the scores of the ITPA. The consistent relationship between the ITPA and measures of intelligence poses a serious problem for the studies relating the ITPA to academic skills, since in most of these studies intelligence was not partialled out.

DTVP. Since the DTVP is composed of five subtests which are intended to tap distinct aspects of visual perception, this instrument's construct validity would be supported by factor analytic results which indicated five factor groupings. Hammill and Wiederholt (1973) reviewed construct validity studies on the DTVP. They report that nine studies failed to find five separate perceptual factors in the DTVP, "In fact, seven studies found only one factor, while the other two studies found only two" (p. 41). Hammill and Wiederholt also report fourteen studies which correlated IQ measures to the DTVP. These coefficients ranged from .18 to .59, with a median of .39. Correlation between the DTVP and the Bender are reported in three studies, yielding coefficients of .75, .52 and .63. Results, in general, do not show the patterns of high and low correlations that are needed for construct validation, and may be accounted for by a third factor common to each measure, namely intelligence.

Auditory tests. Sabatino (1973) reviews factor analytic studies across a number of auditory ability tests. Although not reporting the specific tests used, he concludes:

Review of the factorial work to date suggests that the aspects of auditory perception defined as important by any given researcher

are simply those he selected to study. The general lack of agreement as to the important dimensions of auditory perception bears out again the complexity of abilities, subskills and skills present in its make-up. (p. 58)

Sabatino also reported that for three studies, correlations of auditory perception with intelligence hovered around .40. Apparently, very little research has addressed the construct validity of auditory perceptual measures.

Summary of construct validity. Reviews on the construct validity of tests used to measure underlying ability provide mixed results. Factor analytic studies indicate that ITPA performance is highly related to IQ, but give considerable support for the level and process dimensions hypothesized by the model. In contrast, there exists no support for the existence of five independent perceptual abilities as suggested by the DTVP. One should bear in mind that even if factor structures were found, it would not mean that the test instrument is educationally useful, or even that the factor structure was properly named. Indeed, Waugh (1975) has suggested that the ITPA is misnamed and is really a measure of cognitive functioning or intelligence rather than of perceptual and psycholinguistic abilities.

Assumption 4A:

Training Weak Abilities

According to the DD-PT model, children who have failed to develop adequate perceptual, motor, or psycholinguistic abilities will encounter

serious problems in acquiring basic academic skills. When weak abilities are identified, they should be trained or strengthened lest they continue to obstruct school progress. Major assumptions inherent in this view are that these skills can be trained and that such training will result in improved academic performance. This section will consider the extent to which systematic training of underlying abilities has been successful in improving those abilities. The effect of ability training on academic achievement measures will be described in the next section.

A common research paradigm characterizes ability training investigations. First, children are identified who perform poorly both on an ability assessment and on an academic measure. Second, some of these children are selected to form the experimental group and are given ability training; the others serve as a control and receive the "regular" program of instruction. Finally, after a specified time has passed both groups of children are retested on the original ability test and their performance is compared. The amount of ability improvement is analyzed as a function of the two treatments.

Perceptual Motor Training

Goodman and Hammill (1973), Hallahan and Cruickshank (1973), Hammill (1972), Hammill et al. (1974), Haring and Bateman (1971), Keogh (1974), Kleisius (1972), Krippner (1973), Proger et al. (1973), Robinson (1972), Sabatino (1973), Sedlack and Weener (1973), and Ysseldyke (1973) have all provided reviews of training in different ability areas. These reviews differ in several respects other than the specific ability area

of focus. Some concentrated on specific psychometric ability tests and attempted to determine if various approaches to ability training had been successful in improving performance on these specific tests (Hammill & Larsen, 1974a; Proger et al., 1973; Sedlack & Weener, 1973). Others focused on a particular training program, such as that developed by Frostig and Horne (1964), and recorded that program's success in improving underlying abilities as measured by various instruments (Hallahan & Cruickshank, 1973; Haring & Bateman, 1977; and Robinson, 1972). Other reviewers examined multiple programs and their effects on a variety of tests (Hallahan & Cruickshank, 1973; Kleisius, 1972; Proger et al., 1973; Sabatino, 1973; and Ysseldyke, 1973). Only twice have reviewers attempted to categorize studies on the basis of the population studied (Hallahan & Cruickshank, 1973; Hammill & Larsen, 1974a). Another difference among reviews has been the inclination of some authors to differentiate between well and poorly designed studies. Finally, some reviewers merely reported presence or absence of treatment effects, while others reported in greater detail how particular treatments affected particular measures. These differences among reviews, along with the fact that no review was comprehensive, have contributed to divergent and sometimes conflicting conclusions regarding the degree of success enjoyed by ability training programs.

In an attempt to provide a more complete picture of the training studies, the following strategy was adopted. The most comprehensive reviews are reported first, followed by studies from other reviews which were not included in the more comprehensive reviews. Primary sources

were consulted when reviews presented conflicting information or lacked sufficient detail. Since it was impossible to secure primary sources in every such instance, it was sometimes necessary to report the studies as described in the review. The only studies considered were those in which trained and untrained groups were compared, and in which training focused on a psychological ability considered important for academic achievement.

Table 8 summarizes the success of training for abilities popularized by DD-PT. Studies are classified according to the specific training

Insert Table 8 about here

program employed and as to the adequacy of the research design. The differentiation between well and poorly designed studies was not ours; rather, these were the judgments of past reviewers. For example, to be classified as well designed by Hammill et al. (1974), a study had to have an $N \geq 20$, a control group and training for a minimum of 15 weeks. Hallahan and Cruickshank (1973) classified studies as well or poorly designed depending upon the extent to which a study controlled (e.g. random assignment) for potential sources of bias. It is evident from this table that well designed attempts to train underlying abilities are characterized by failure; only 24% of these studies show success. Ability training more often appeared to succeed in the poorly designed studies (88%). All studies considered, the overwhelming weight of this

research indicates that underlying abilities presumed important for school learning are remarkably resistant to improvement through training with current methods. Significant ability training effects are observed on only 43% of the measures.

The same pattern is evident for individual training programs. Of the studies rated as good, only 12.5% of those employing Frostig training materials report success in improving visual abilities. To include poorly designed studies and those not rated raises the success rate to 55%, but brings into doubt the validity of the claim to success. Well designed studies employing Kephart-Getman procedures show a 24% success rate. Success rate rises only to 34% when all Kephart-Getman studies are included. The success rate with the Delacato approach is 0% for well designed studies and 76% for all studies. Studies employing miscellaneous perceptual-motor training programs show a success rate of only 43%.

Psycholinguistic Training

Training of psycholinguistic abilities as measured by the ITPA has been summarized by Hammill and Larsen (1974a) who considered 39 studies and 280 separate comparisons of trained and untrained groups. They categorized results by training program, (i.e., Peabody Language Development Kit [PDLK] or a selected activities approach), by psycholinguistic ability measured, by population studied, and by degree of individualization (prescriptive vs. non-individualized, where all children were exposed to a set program).

Table 9 reveals that neither training program (PDLK vs. selected activities) produced consistently positive results, with success rates

Insert Table 9 about here

generally below 50%. Similarly, results for the various populations studied indicate that training failed more often than not. Of the specific abilities subjected to training, only verbal expression appeared trainable in more than 50% of the studies. Finally, prescriptive approaches were, in general, no more effective than non-individualized approaches.

Summary of Ability Training

After reviewing over 100 separate studies covering a wide range of auditory, visual and psycholinguistic training programs, one finds little evidence to support the trainability of underlying psychological abilities. Fewer than 50% of training efforts have yielded dividends in ability growth. This is the case whether results are summarized according to specific training programs, to the degree of individualization, or to the populations studied.

Assumption 4B:

Effects on Academic Performance

The crucial test of the DD-PT model is whether training weak abilities leads to increased academic success. The research paradigm

in studies which address this issue is essentially the same as that described in the previous section on ability training. The primary difference, of course, is that training is not evaluated on the basis of improved ability test scores, but instead on improved performance on tests of academic functioning (e.g., reading achievement) or sometimes on measures of intellectual functioning (e.g., the WISC).

While Hallahan and Cruickshank (1973), Hammill and Wiederholt (1973), Keogh (1974), Kleisius (1972), Robinson (1972), and Sedlack and Weener (1973) all provide reviews of this aspect of DD-PT, the most comprehensive reviews are those by Goodman and Hammill (1973), and Hammill et al. (1974).

These reviews cover over 100 separate investigations. In addition, five studies reviewed by Keogh (1974) were not reported in sufficient detail for inclusion in this summary and were not available as primary sources. These studies (Coleman, 1972; Hopper, 1962; Morgan, Note 2; Shearer, Note 3; and Young, Note 4), however, generally reported significant effects on some academic measures.

Table 10 summarizes the studies according to type of training program (e.g., Kephart-Getman, Frostig, etc.), type of outcome measure (e.g., readiness, reading, IQ, etc.), quality of research design as judged by the reviewers, the number of measures and the percent of these measures on which trained groups outperformed untrained controls (i.e., percent successes).

Insert Table 10 about here

Considering all studies together, ability trained groups surpassed control groups on only 38% of the 116 measures reported. This figure reduces to 36% when the studies designated as "poor" are eliminated. Intelligence measures were least often affected (14%) and "miscellaneous other" measures most affected (75%). Many of the latter would not be considered achievement measures. Reading and general achievement combined were positively affected only 35% of the time. When only good studies were considered, this reduced to 33%.

When the effects on reading and general achievement are analyzed by type of training program, the results fail to support any particular approach to ability training. Kephart-Getman (42%) and "other visual perceptual" programs (61%) enjoy more success than do Frostig (17%), Delacato (20%), auditory perceptual (33%) or "other ability" programs (38%).

Assumption 5:

Matching Instruction to Learner Strengths

The final issue of concern is whether knowledge of a child's profile of strengths and weaknesses is useful in planning academic instruction. This approach is based on the supposition that aptitude-treatment interactions exist. Such a strategy may improve learning even though the weak abilities themselves are resistant to training.

To illustrate, for a child who is diagnosed as an "auditory learner," it is recommended that academic instruction be provided through the auditory channel rather than through the weaker visual channel (Johnson & Myklebust, 1967; Wepman, 1971; and Lerner, 1971).

The standard research paradigm employed in these "modality matching" studies involves identifying children's strong and weak modalities through an instrument like the ITPA. Next, some children receive academic instruction through their strong modality while others receive instruction through their weak modality. Finally, achievement is studied to determine if children whose instructional program matched their modality strengths surpassed those whose program matched their modality weakness.

Arter and Jenkins (1977) and Ysseldyke (1973) have reviewed the modality matching research. Arter and Jenkins' review is the more comprehensive, including 15 studies in which children identified as auditory, visual or kinesthetic learners were presented with reading instruction based on auditory, visual or kinesthetic approaches. In 14 of the 15 studies, matching instruction with modality strength failed to produce differential improvement; children learned equally well whether or not instruction was matched to their strong modality. In no study involving elementary aged students was the approach successful.

The consistently negative nature of these results casts considerable doubt on the usefulness of ability assessments in planning academic instruction. However, modality studies to date have been concerned with

reading instruction; other academic areas may be more amenable to modality influence.

SUMMARY AND DISCUSSION

Criticism of the DD-PT model is based upon philosophical, theoretical, and empirical considerations. In this section the validity of the five assumptions which underlie the DD-PT model is discussed in light of existing evidence.

Assumption 1:

Existence and Measurability of Abilities

In some ideal sense educationally relevant abilities may exist and be measurable. However, two major obstacles have thwarted attempts to identify and assess educationally important abilities. One is definitional while the other is measurement-related. The terminology used by DD-PT has posed a significant problem since there is little agreement as to what is meant by many of the ability labels. For example, Hammill (1972) in reviewing 33 studies of "perception" found that some authors considered perception as the entire perceptive process from stimulus reception to cognitive analysis. Other authors made a distinction between "sensation" (receiving stimuli) and "perception" (the remainder of the process). Still other writers distinguished between sensation and cognition, with "perception" subsumed under the rubric of "cognition." Finally, there are those who distinguish among "sensation," "perception" and "cognition." In the latter case, the processes which involve thinking are assigned to cognition, while those dealing with

nonsymbolic, nonabstract properties of stimuli are relegated to perception. These differences in terminology not only make communication difficult, but also make tests and their results ambiguous, especially for those attempting to design remediation materials.

Those attempting to support the first assumption of DD-PT encounter another obstacle, namely the measurement of hypothetical constructs. Test developers attempt to label their instruments to indicate the variable under consideration. When that variable is an abstract concept, there is no guarantee that the measure actually taps that construct. An example comes from the ITPA which purports to test psycholinguistic ability. Waugh (1975), after reviewing the research on reliability and validity of the ITPA, concluded that the test does not measure psycholinguistic functioning, but instead measures cognitive functioning, that is, intelligence.

An underlying ability is assessed by measuring performance on activities which are thought to require the ability. Unfortunately, no activity can be considered a pure measure of an isolated ability. Any assessment task is susceptible to contamination by irrelevant (with respect to the target ability) features of the task. For example, putting shapes in a sequence depends not only on "visual sequential memory" but also on the motor ability to physically manipulate the shapes and the ability to understand the verbal instructions which detail the task requirements.

Thus, even though Assumption 1 cannot, in principle, be disproven, its acceptance would seem to require either an act of faith or empirical

demonstrations of the efficacy of the DD-PT model. The latter is the subject of the next assumptions.

Assumption 2:

Reliability of the Tests Used in Differential Diagnosis

Measurement authorities suggest that if important decisions are to be made on the basis of a test, then that test should produce retest reliability coefficients of at least .85. The median reliabilities of the most frequently employed tests for the DD-PT do not meet this minimum criterion; the median reported reliability for the ITPA subtests is .71 and for the DTVP Total Score is .79. The median reliabilities of many of the ITPA and DTVP subtests, which are used to prescribe different kinds of instruction, are even lower. Reliabilities of other tests commonly used in DD-PT range from .35 to .90 with a median of .83. Clearly, the evidence on reliability of DD-PT instruments does not justify confidence in their continued use as a basis for making important educational decisions.

Assumption 3:

Validity of the Tests Used in Differential Diagnosis

Instruments employed in DD-PT were examined in connection with four types of test validity. Research indicates that while individual investigations occasionally report satisfactory correlations, overall results have not proven differential ability tests to be concurrently valid with respect to academic achievement. Exceptions include ITPA Grammatic

Closure and Total Score for reading achievement, and the DTVP Total Score for arithmetic achievement. Studies in which IQ is controlled report far fewer differential ability-achievement relationships than do studies which do not partial out IQ.

With regard to the predictive validity of DD-PT instruments, current research is difficult to interpret. To begin with, only a few longitudinal studies have focused upon one instrument, and these studies paint an ambiguous picture. Often one study will suggest that an instrument meets minimum validity standards, but the next study will give a contrary indication. Moreover, nearly all of the longitudinal studies have neglected to control for the contribution of IQ, which itself can account for an apparent relationship between DD-PT measures and later achievement. Given these qualifications, it appears that the strongest case for predictive validity can be made for certain auditory measures (e.g., ITPA Auditory Association).

Studies of the diagnostic validity of DD-PT instruments yield a similar picture. Neither the ITPA nor miscellaneous visual perception tests appear capable of discriminating between good and poor readers. Auditory perceptual tests, in contrast, have an encouraging record. There have been few studies which examine the diagnostic validity of underlying abilities with regard to academic areas other than reading.

The results of studies which consider construct validity have yielded mixed results. There is some support for at least two dimensions in the ITPA. There is no empirical support for the five separate abilities hypothesized by Frostig, nor has there been a consistent series of

dimensions delineated for auditory tests in general. Intelligence appears to be a strong general factor in most of the tests, suggesting to some that "specific" ability tests are, at best, measures of general intellectual ability.

Investigations of reliability and validity do not provide the ultimate test of DD-PT. The major premise of the model is that these abilities are crucial for academic success, that if an ability is weak it can be trained, and that such training will result in improved academic performance. If correlations between underlying abilities and academic achievement were high (which they are not), there would be corroboration for, but not proof of this essential proposition, since correlation does not imply causation. Similarly, if consistent factors were found in factor analytic research, though for the most part they were not, it would again provide corroborating evidence, but not proof of the essential proposition. To prove that the ability training approach is useful, one must identify weaknesses, train them, and observe subsequent improvement on academic performance. Or, alternatively, one must identify strengths and weaknesses, plan instruction in accord with them, and demonstrate that such instruction is differentially effective.

Assumption 4:

Effects of Ability Training

There have been many attempts to train specific abilities. Psycholinguistic, visual perceptual, auditory perceptual and motor abilities have all been the focus of training. The training itself has been based

on various theoretical positions related to the ITPA, Peabody Language Kits, Doman-Delacato Methods, Kephart-Getman programs, the Frostig-Horne program, and a number of miscellaneous perceptual motor programs. Ability training succeeded about 24% of the time in well designed investigations. It is difficult to escape the conclusion that abilities measured in differential diagnosis are highly resistant to training by existing procedures.

Given this, it would certainly be surprising to find that ability training improved academic performance. Indeed, the research shows that more often than not academic performance is not improved. Excluding studies designated as poorly designed, ability trained groups out-performed untrained controls on roughly one-third of the academic measures taken. In the majority of studies, control groups performed as well on both ability and academic measures as did the experimental groups. Thus, Assumption 4 lacks support.

Assumption 5:

Differential Diagnosis Improves Academic Programming

Advocates of DD-PT propose that differential diagnosis helps the teacher to determine how a child should be taught. The particular constellation of psychological abilities identified through differential diagnosis is said to reveal individual learning styles which, in turn, determine special instructional methodologies and materials.

To date there are 14 reported efforts to improve beginning reading by matching instructional materials and procedures to children's modality

strengths. In none of these was reading instruction improved by modality-instructional matching. In one study conducted with secondary aged students the approach appeared to succeed. Assumption 5 appears to lack even minimal empirical support.

Reasons for the Lack of Support for the DD-PT Model

A number of explanations have been posed that could account for this general failure to support the DD-PT model (Arter & Jenkins, 1977; Hammill & Larsen, 1974a; Minskoff, 1975; Newcomer et al., 1975; and Proger et al., 1973).

First, the ability training model may itself be invalid. Underlying abilities may not exist as such, or they may exist but be an unimportant factor in instruction.

Another possible explanation for the failure of DD-PT is that even though underlying abilities may exist and may be functionally related to achievement, they produce only a weak effect that is easily masked by stronger, uncontrolled variables. There are two issues concerning this point. First, Minskoff (1975) and Keogh (1974) argue that ability training has not been successfully demonstrated because of poorly designed studies. They contend that, in general, the research suffers from uncontrolled teacher effects, Hawthorne effects, experimenter effects (no double-blind), and subject selection problems. Responding to this point Newcomer et al., (1975) argue that there would be even less evidence to support DD-PT if the research methodologies were improved, since the uncontrolled sources of bias usually favor the DD-PT groups. In general,

experimental subjects received more one-to-one training than did controls, experiments tended to be conducted by investigators who expected to find treatment effects, novelty effects favored ability trained groups, and regression effects were often not considered. In fact, the studies which employed better research designs less often supported ability training. Quality of research designs is not a plausible explanation for failure to demonstrate DD-PT success.

The second issue related to masking variables is that of classroom usefulness. If the effects of underlying abilities are easily masked by other, more robust variables, then attention should be focused on those stronger variables.

A third explanation for the lack of support for the DD-PT approach is related to the prescriptive/remediation programs. With regard to ability training, the instructional programs themselves may need to be strengthened, or "abilities" identified for training may need to be more carefully selected. Williams (1977) argues that ability training may be useful if the abilities are chosen very carefully so that they are important components of the reading task (e.g., sound blending). However, she cautions against existing auditory ability training programs:

However, we are not proposing that training in auditory perceptual skills, generally speaking, will lead to better achievement in reading. The lessons from the past two or three decades on the relationship of visual-perceptual skills and reading have convinced us otherwise. When we reviewed several recently developed

and popular auditory programs (many of which are listed by Kass, 1972), we were dismayed to find content coverage or methodology which, on the basis of current knowledge about perceptual skills and instructional methodology, was surprisingly poor. For example, there is often excessive emphasis on material unrelated to that of early reading skills (e.g., children are asked to identify the animals who make different barnyard sounds). Sometimes language tasks are presented in a context quite different from that of initial decoding. For example, two voices present two separate messages concurrently and the child must focus his attention on only one of the conflicting messages. Sometimes relevant tasks are presented, but in a way which would tax the child's memory or confuse him. In addition, one program, dealing with the important skill of auditory analysis, develops tasks to a level of difficulty far beyond that required for initial decoding; some programs present what could be classified as practice material but no instruction; and blending as a process is not taught in any of the programs. (pp. 284-285)

With regard to modality matching, instructional programs may not sufficiently emphasize one modality to the exclusion of others. Indeed, this may not even be possible (Vandever and Neville, 1974) since reading seems to require both auditory and visual skills no matter how it is taught.

Fourth, abilities may exist and be useful, but tests have not been developed yet which consistently and accurately reveal ability profiles.

The reliabilities and validities of current DD-PT tests are often unsatisfactory.

A fifth explanation for the lack of positive DD-PT results involves factors relating to individualization of instruction. Minskoff (1975) argued that the DD-PT model is effective, but that it has not been fairly tested. She points out that one premise of the model, and indeed of education in general, is that instruction must be geared to the individual needs of each student. Therefore, studies employing different populations and treatments should not be compared, nor should studies which provide one treatment to a large group of children be expected to show overall effects. She also proposes that most DD-PT techniques could really be expected to work only with severely disabled populations. Newcomer et al. (1975) have countered this argument, citing Hammill and Larsen (1974a) who failed to locate an advantage for individualized over nonindividualized DD-PT programs. Newcomer et al. also point out that most test and program developers recommend their products for use with any children who evidence ability deficits, not just the severely disabled. Moreover, several reviewers (e.g., Hallahan & Cruickshank, 1973; Hammill & Larsen, 1974a) have analyzed results by population, program, and criterion test and still the results remain uniformly negative. While the DD-PT model may, in theory, be responsive to different populations of learners, the burden of proof rests with proponents of the model (Hammill & Larsen, 1974a).

In summary, it is not surprising that DD-PT has not improved academic achievement, since most ability assessment devices have

inadequate reliability and suspect validity. Moreover, abilities themselves have resisted training, and given the low correlations between ability assessments and reading achievement, it is not surprising that modality-instructional matching has failed to improve achievement.

The repeated failure to support the basic assumptions underlying the DD-PT model casts doubt on the model's validity. We do not intend to suggest that the model is theoretically untenable, or that it may not one day be effectively implemented. Rather, we believe that with the current instructional programs and tests, this model is not useful. A number of authors who have reviewed specific aspects of the DD-PT model have arrived at a similar conclusion (Hammill & Larsen, 1974b; Sedlack & Weener, 1973; Silverston & Deichman, 1975; and Ysseldyke, 1973). For example, with reference to psycholinguistic training, Newcomer et al. (1975) write:

We cannot help but conclude that psycholinguistic training based on the Kirk-Osgood model is not successful because it does not help children to increase their ability to speak or understand language, nor does it aid them in academic skills such as reading, writing or spelling...the wrong skills are being remediated. (p. 147)

Unfortunately, this view does not represent that held by most authorities and practitioners in special education. The DD-PT model is preferred by the vast majority of special education teachers (Arter & Jenkins, 1977). In a state-wide survey of Illinois it was found that 82% of special education teachers believed that they could, and should,

train weak abilities, 99% thought that a child's modality strengths and weaknesses should be a major consideration when devising educational prescriptions, and 93% believed that their students had learned more when they modified instruction to match modality strengths. The same survey provided data to suggest that teacher training programs were, to a large degree, responsible for these views and practices. Unsupported expert opinion and teacher training programs resulting from this opinion appear to have a direct, deleterious effect on teacher behavior and an indirect effect on children's learning. Not only are teachers adhering to an ineffective model, but because they have been persuaded that the model is useful, they are less apt to create variations in instructional procedures which will result in improved learning. We believe that until a substantive research base for the DD-PT model has been developed, it is imperative to call for a moratorium on advocacy of DD-PT, on classification and placement of children according to differential ability tests, on the purchase of instructional materials and programs which claim to improve these abilities and on coursework designed to train DD-PT teachers.

Reference Notes

1. Hammill, D. D., Larsen, S. C., Parker, R., Bagley, M. T., & Sanford, H. G. Perceptual and conceptual correlates of reading. Unpublished manuscript. 1505 Sunny Vale, #217, Austin, Texas, 1974.
2. Morgan, M. W. The effect of visual perceptual training upon subsequent scholastic progress of children with specific visual difficulties. Duncan, Okla.: Optom. Extension Program Foundation, 1966.
3. Shearer, C. W. A study comparing disadvantaged preschool children's test scores in directionality orientation and sensory concepts fundamental to their reading readiness before and after visual-perceptual motor training. Duncan, Okla.: Optom. Extension Program Foundation, 1969.
4. Young, L. M., Motor coordination visual perception study. Duncan, Okla.: Optom. Extension Program Foundation, 1971.

References

- Alshan, L. Reading readiness and reading achievement. In A. Figurel (Ed.), Reading and inquiring. Newark, Del.: International Reading Association, 1965.
- Anatasi, A. Psychological testing (3rd ed.). London: The MacMillan Company, 1968.
- Arter, J. A., & Jenkins, J. R. Examining the benefits and prevalence of modality considerations in special education. Journal of Special Education, 1977, 11, 281-298.
- Bagford, J. Reading readiness scores and success in reading. The Reading Teacher, 1968, 21, 324-328.
- Barrett, T. C. The relationship between measures of prereading visual discrimination and first grade reading achievements: A review of literature. Reading Research Quarterly, 1965, 1, 51-76. (a)
- Barrett, T. C. Visual discrimination tasks as predictors of first grade reading achievement. The Reading Teacher, 1965, 18, 276-282. (b)
- Bateman, B. Three approaches to diagnosis and educational planning for children with learning disabilities. Academic Therapy Quarterly, 1967, 3, 11-16.
- Beery, J. W. Matching of auditory and visual stimuli by average and retarded readers. Child Development, 1967, 38, 827-833.
- Beery, K. & Buktenica, N. Developmental test of visual motor integration. Chicago: Follet Publishing Co., 1967.

- Bender, L. A visual-motor Gestalt test and its clinical use. Research Monographs, No. 3. American Ortho-psychiatric Association, 1938.
- Benton, A. L. Benton visual retention test (Rev. ed.). New York: Psychological Corporation, 1955.
- Bergan, J. R. A study of the relationships between perception and reading. Tucson, Ariz.: University of Arizona, 1966. (ERIC Document Reproduction Service No. ED 017 435)
- Birch, H. G., & Belmont, L. Auditory-visual integration, intelligence, and reading ability in school children. Perceptual and Motor Skills, 1965, 20, 295-305.
- Bonsall, C., & Dornbush, R. L. Visual perception and reading ability. Journal of Educational Psychology, 1969, 60, 294-299.
- Burns, E. Effects of restricted sampling on ITPA scaled scores. American Journal of Mental Deficiency, 1976, 80, 394-400.
- Busby, W. A. & Hurd, D. E. Relationships of auditory and visual reaction times to reading achievement. Perceptual and Motor Skills, 1968, 27, 447-450.
- Chall, J., Roswell, F., & Blumenthal, S. H. Auditory blending ability: A factor in success in beginning reading. The Reading Teacher, 1963, 17, 113-118.
- Cicirelli, V. G., Granger, R., Schemmel, D., Cooper, W., & Holthouse, N. Performance of disadvantaged primary-grade children on the revised Illinois Test of Psycholinguistic Abilities. Psychology in the Schools, 1971, 8, 240-246.

- Coleman, H. M. The West Warwick visual perception study. Journal of the American Opthomologist Association, 1972, 43, Part I, 452-462, Part II, 532-542.
- Connor, J. P. Bender-Gestalt test performance as a predictor of differential reading performance. Journal of School Psychology, 1968-69, 7, 41-44.
- Campbell, D. T., & Fiske, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 1959, 56, 81-105.
- deHirsch, K., Jansky, J. J., & Langford, W. S. Predicting reading failure. New York: Harper, 1966.
- Dennis, R., & Dennis, M. Dennis visual perception scale. Los Angeles: Western Psychological Services, 1969.
- Deputy, E. C. Predicting first-grade reading achievement. Teacher College Contributions to Education, 1930, No. 421.
- Dornbush, R. L., & Basow, S. The relationship between auditory and visual short-term memory and reading achievement. Child Development, 1970, 41, 1033-1044.
- Dugger, J. G. The ITPA as a diagnostic instrument for reading disability problems in mental deficiency. Psychology, 1969, 6, 47-51.
- Egeland, B. M., DiNello, M., & Carr, D. The relationship of intelligence, visual motor, psycholinguistic and reading readiness skills with achievement. Educational and Psychological Measurement, 1970, 30, 451-458.

- Ellenhammer, M. Assessing progress in reading in Denmark. In M. Jenkinson (Ed.), Reading instruction: An international forum. Newark, Del.: International Reading Association, 1966.
- Flynn, P. T., & Byrne, M. C. Relationships between reading and selected auditory abilities of third-grade children. Journal of Speech and Hearing Research, 1970, 13, 731-740.
- Frostig, M. Developmental Test of Visual Perception. California: Consulting Psychologists Press, 1963.
- Frostig, M., & Horne, D. The Frostig Program for the Development of Visual Perception: Teachers guide. Chicago: Follett Publishing Co., 1964.
- Garrett, H. E. Statistics in psychology and education. New York: Longmans Green, 1954.
- Gaskins, A. I. Characteristics that differentiate dyslexic from nondyslexic poor readers (Doctoral dissertation, University of Pennsylvania, 1970). Dissertation Abstracts International, 1970, 31, 3336A. (University Microfilms No. 70-22,862)
- Gates, A. I. The psychology of reading and spelling. Teachers College Contributions to Education, 1922, No. 129.
- Gates, A. I. A study of the role of visual perception, intelligence, and certain associative processes in reading and spelling. Journal of Educational Psychology, 1926, 17, 433-445.
- Gates, A. I. An experimental evaluation of reading readiness tests. Elementary School Journal, 1939, 39, 497-508.

- Gates, A. I., Bond, G. L., & Russell, D. H. Methods of determining reading readiness. New York: Bureau of Publications, Teachers College, Columbia University, 1939.
- Gavel, S. R. June reading achievements of first-grade children. Journal of Education, 1958, 140, 37-43.
- Giebink, J. W., & Birch, R. The Bender-Gestalt Test as an ineffective predictor of reading achievement. Journal of Clinical Psychology, 1970, 26, 484-485.
- Goetzinger, C. P., Dirks, D. D., & Baer, C. J. Auditory discrimination and visual perception in good and poor readers. Annals of Otology, Rhinology, and Laryngology, 1960, 69, 121-136.
- Goldman, R., Fristoe, M., & Woodcock, R. W. Goldman-Fristoe-Woodcock Test of Auditory Discrimination. Circle Pines, Minn.: American Guidance Service, Inc., 1970.
- Goodman, L., & Hammill, D. The effectiveness of the Kephart-Getman activities in developing perceptual-motor cognitive skills. Focus on Exceptional Children, 1973, 9, 1-9.
- Graham, F. K., & Kendall, B. L. Memory-for-Designs Test. Missoula, Mont.: Psychological Test Specialists, 1960.
- Guilford, J. P. Fundamental statistics in psychology and education. New York: McGraw Hill, 1956.
- Hallahan, D. P. & Cruickshank, W. M. Psychoeducational foundations of learning disabilities. Englewood Cliffs, N.J.: Prentice-Hall, 1973.

- Hammill, D. Training visual perceptual processes. Journal of Learning Disabilities, 1972, 5, 39-44.
- Hammill, D., Goodman, L., & Wiederholt, J. L. Visual-motor processes: Can we train them? The Reading Teacher, 1974, 27, 470-479.
- Hammill, D. D., & Larsen, S. C. The effectiveness of psycholinguistic training. Exceptional Children, 1974, 41, 5-14. (a)
- Hammill, D. D., & Larsen, S. C. The relationship of selected auditory perceptual skills and reading ability. Journal of Learning Disabilities, 1974, 7, 40-46. (b)
- Hammill, D. D., & Wiederholt, J. L. Review of the Frostig Visual Perception Test and the related training program. In L. Mann & D. A. Sabatino (Eds.), The first review of special education (Vol. 1). Philadelphia: JSE Press, 1973.
- Haring, N. G., & Bateman, B. Teaching the learning disabled child. Englewood Cliffs, N.J.: Prentice-Hall, 1977.
- Harrington, M. L., & Durrell, D. Mental maturity versus perception abilities in primary reading. Journal of Educational Psychology, 1955, 46, 375-380.
- Haworth, M. R. The primary visual motor test. New York: Grune & Stratton, Inc., 1970.
- Hirshoren, A. A comparison of the predictive validity of the Revised Stanford-Binet Intelligence Scale and Illinois Test of Psycholinguistic Abilities. Exceptional Children, 1969, 35, 517-521.

- Hopper, J. E. Analysis of Winter Haven project. In A preliminary report. Winter Haven study of perceptual learning. Winter Haven, Fla.: Winter Haven Lions Res. Foundation, 1962.
- Hurd, D. E. A study of the relationship between reading achievement and sense modality shifting. Peoria, Ill.: Bradley University, 1967.
(ERIC Document Reproduction Service No. ED 015 119)
- Johnson, D. Educational principles for children with learning disabilities. Rehabilitation Literature, 1967, 28, 317-322.
- Johnson, D., & Myklebust, H. Learning disabilities: Educational principles and practices. New York: Grune & Stratton, Inc., 1967.
- Katz, P. A., & Deutsch, M. Relation of auditory-visual shifting to reading achievement. Perceptual and Motor Skills, 1963, 17, 327-332.
- Keogh, B. K. The Bender-Gestalt as a predictive and diagnostic test of reading performance. Journal of Consulting Psychology, 1965, 29, 83-84.
- Keogh, B. K. Optometric vision training programs for children with learning disabilities: Review of issues and research. Journal of Learning Disabilities, 1974, 7, 219-231.
- Keogh, B. K., & Smith, C. Visuo-motor abilities for school prediction: A seven-year study. Perceptual and Motor Skills, 1967, 25, 101-110.
- Kephart, N. C. The slow learner in the classroom. Columbus, Ohio: Merrill Inc., 1960.
- Kimmell, G. M., & Wahl, J. Screening test for auditory perception. San Rafael, Cal.: Academic Therapy Publications, 1969.

- Kirk, S. A. Educating exceptional children. Boston: Houghton-Mifflin, 1972.
- Kirk, S., McCarthy, L., & Kirk, W. Illinois Test of Psycholinguistic Abilities (Rev. ed.). Urbana: The University of Illinois Press, 1968.
- Kirk, S. A., & Kirk, W. D. Psycholinguistic learning disabilities: Diagnosis and remediation. Urbana: The University of Illinois Press, 1971.
- Kleisius, S. E. Perceptual-motor development and reading--a closer look. In R. C. Aukerman (Ed.), Some persistent questions on beginning reading. Newark, Del.: International Reading Association, 1972.
- Koppitz, E. M. Bender-Gestalt Test, Visual Aural Digit Span Test and reading achievement. Journal of Learning Disabilities, 1975, 8, 154-157.
- Krippner, S. Research in visual training and reading disability. In B. Bateman (Ed.), Reading performance and how to achieve it. Seattle: Bernie Straub Publishing Co., 1973.
- Lachmann, F. M. Perceptual-motor development in children retarded in reading ability. Journal of Consulting Psychology, 1960, 24, 427-431.
- Larsen, S. C., & Hammill, D. D. The relationship of selected visual-perceptual abilities to school learning. Journal of Special Education, 1975, 9, 281-291.

- Larsen, S. C., Rogers, D., & Sowell, V. The use of selected perceptual tests in differentiating between normal and learning disabled children. Journal of Learning Disabilities, 1976, 9, 85-90.
- Lee, J. M., Clark, W., & Lee, D. M. Measuring reading readiness. Elementary School Journal, 1934, 34, 656-666.
- Leider, A. B. Relationship of visual perception to word discrimination. In H. M. Robinson & H. K. Smith (Eds.), Clinical studies in reading III, Supplemental Educational Monograph, 1967, 97, 104-108.
- Lerner, J. Children with learning disabilities. New York: Houghton-Mifflin Co., 1971.
- Lingren, R. I. Performance of disabled and normal readers on the Bender-Gestalt, auditory discrimination test, and visual-motor matching. Perceptual and Motor Skills, 1969, 29, 152-154.
- Lovell, K., & Gorton, A. A study of some differences between backward and normal readers of average intelligence. British Journal of Educational Psychology, 1968, 38, 240-248.
- Mann, L. Psychometric phrenology and the new faculty psychology: The case against ability assessment and training. Journal of Special Education, 1971, 5, 3-65.
- McCarthy, J. J., & Olson, J. L. Validity studies on the Illinois Test of Psycholinguistic Abilities. Urbana: University of Illinois Press, 1964.
- Meikle, S., & Kilpatrick, D. L. Changes in perceptual, motor, and reading test scores in a remedial reading group. Canadian Psychologist, 1971, 12, 254-269.

- Meyers, C. E. What the ITPA measures: A synthesis of factor studies of the 1961 edition. Educational and Psychological Measurement, 1969, 29, 867-876.
- Minskoff, E. H. Research on psycholinguistic training: Critique and guidelines. Exceptional Children, 1975, 42, 136-144.
- Monroe, M. Children who cannot read: The analysis of reading disabilities and the use of diagnostic tests in the instruction of retarded readers. Chicago: University of Chicago Press, 1932.
- Muehl, S., & Kremenak, S. Ability to match information within and between auditory and visual sense modalities and subsequent reading achievement. Journal of Educational Psychology, 1966, 57, 230-239.
- Mueller, M. W. Prediction of achievement of educably mentally retarded children. American Journal of Mental Deficiency, 1969, 73, 590-596.
- Newcomer, P. L., & Hammill, D. D. ITPA and academic achievement: A survey. The Reading Teacher, 1975, 28, 731-741.
- Newcomer, P., Hare, B., Hammill, D. D., & McGettigan, J. Construct validity of the ITPA subtests. Exceptional Children, 1974, 40, 509-510.
- Newcomer, P., Larsen, S. C., & Hammill, D. D. A response to research on psycholinguistic training. Exceptional Children, 1975, 42, 144-148.
- Nila, M. Foundations of a successful reading program. Education, 1953, 73, 543-555.
- Nunnally, J. C. Psychometric theory. New York: McGraw-Hill, 1967.
- Olson, A. V. Growth in word perception abilities as related to success in beginning reading. Journal of Education, 1958, 140, 25-36.

- Peter, L. J. Prescriptive teaching. New York: McGraw-Hill, 1965.
- Phelan, M. B., Sr. Visual perception in relation to variance in reading and spelling. Washington, D. C.: Catholic Education Press, 1940.
- Plessas, G. P. Reading abilities of high and low auders. Elementary School Journal, 1963, 63, 223-226.
- Poling, D. L. Auditory deficiencies of poor readers. Clinical studies in reading II, Supplementary Educational Monographs, 1953, 77.
- Potter, M. C. Perception of symbol orientation and early reading success. Teacher College Contributions to Education, 1959, No. 939.
- Proger, B. B., Cross, L. H., & Burger, R. M. Construct validation of standardized tests in special education: A framework of reference and application to ITPA research (1967-1971). In L. Mann & D. A. Sabatino (Eds.), The first review of special education (Vol. 1). Philadelphia: JSE Press, 1973.
- Raab, S., Deutsch, M., & Friedman, A. Perceptual shifting and set in normal school children of different reading levels. Perceptual and Motor Skills, 1960, 10, 187-192.
- Resnick, L., Wang, M., & Kaplan, J. Task analysis in curriculum design: A hierarchically sequenced introductory mathematics curriculum. Journal of Applied Behavior Analysis, 1973, 6, 679-710.
- Roach, E. G., & Kephart, N. C. The Purdue Perceptual-Motor Survey. Columbus, Ohio: Charles E. Merrill Publishing Co., 1966.
- Robinson, H. M. Perceptual training--does it result in reading improvements? In R. C. Aukerman (Ed.), Some persistent questions on beginning reading. Newark, Del.: International Reading Association, 1972.

- Robinson, H. M., Letton, M., Mozzi, L., & Rosenbloom, A. An evaluation of the children's visual achievement forms at grade one. American Journal of Optometry and Archives of the American Academy of Optometry, 1958, 35, 515-525.
- Rosenfield, A. G. Psycholinguistic abilities as predictors of reading abilities and disabilities in first grade children (Doctoral dissertation, St. Louis University, 1970). Dissertation Abstracts International, 1970, 32, 798A. (University Microfilms No. 71-03, 274)
- Sabatino, D. A. Auditory perception: Development assessment, and intervention. In L. Mann & D. A. Sabatino (Eds.), The first review of special education (Vol. 1). Philadelphia: JSE Press, 1973.
- Sabatino, D. A., & Ysseldyke, J. E. Effects of extraneous background on visual-perceptual performance of readers and nonreaders. Perceptual and Motor Skills, 1972, 35, 323-328.
- Sedlack, R. A., & Weener, P. Review of research on the Illinois Test of Psycholinguistic Abilities. In L. Mann & D. A. Sabatino (Eds.), The first review of special education (Vol. 1). Philadelphia: JSE Press, 1973.
- Shorr, R., & Svagr, V. Relationship of perceptual and visual skills with reading accuracy and comprehension. Journal of the American Optometric Association, 1966, 37, 671-677.
- Silver, A. A. Diagnostic considerations in children with reading disabilities. Bulletin of the ORTON Society, 1961, 11, 5-11.

- Silverston, R. A., & Deichmann, J. W. Sense modality research and the acquisition of reading skills. Review of Educational Research, 1975, 45, 149-172.
- Smith, N. B. Matching ability as a factor in first-grade reading. Journal of Educational Psychology, 1928, 19, 560-571.
- Steinbach, M. N. An experimental study of progress in first grade reading. Catholic University of America Educational Research Monographs, 1940, 12, 1-118.
- Sterritt, G., & Rudnick, M. Auditory and visual rhythm perception in relation to reading ability in fourth grade boys. Perceptual and Motor Skills, 1966, 22, 857-864.
- Stuart, I. Perceptual style and reading ability. Perceptual and Motor Skills, 1967, 24, 135-138.
- Vandever, T. R., & Neville, D. D. Modality aptitude and word recognition. Journal of Reading Behavior, 1974, 6, 195-201.
- Waugh, R. P. ITPA: Ballast or bonanza for the school psychologists. Journal of School Psychology, 1975, 13, 201-208.
- Weiner, M., & Feldmann, S. Validation studies of a reading prognosis test for children of lower and middle socio-economic status. Education and Psychological Measurement, 1963, 23, 804-814.
- Weiner, P. S., Wepman, J., & Morency, A. A test of visual discrimination. Elementary School Journal, 1965, 65, 330-337.
- Wepman, J. M. Auditory discrimination, speech and reading. Elementary School Journal, 1960, 60, 325-333.

- Wepman, J. M. Auditory discrimination test (Rev. ed.). Chicago: Language Research Associates, 1973.
- Wepman, J. M. The perceptual basis for learning. In E. C. Freerson & B. Barbe (Eds.), Educating children with learning disabilities: Selected readings. New York: Appleton-Century-Crafts, 1967.
- Westinghouse Learning Corporation/Ohio University. The impact of Head Start on children's cognitive and affective development (Vols. 1 & 2). Springfield, Va.: U. S. Department of Commerce Clearinghouse, 1969.
- Whisler, N. G. Visual memory training in first grade: Effects on visual discrimination and reading ability. Elementary School Journal, 1975, 75, 50-54.
- Williams, J. Building conceptual and cognitive strategies into a reading curriculum. In A. S. Reber & D. L. Scarborough (Eds.), Toward a psychology of reading. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1977.
- Wilson, F. T. Early achievement in reading. Elementary School Journal, 1942, 42, 609-615.
- Wilson, F. T., & Burke, A. Reading readiness in a progressive school. Teachers College Record, 1937, 38, 565-580.
- Ysseldyke, J. E. Diagnostic-prescriptive teaching: The search for aptitude-treatment interactions. In L. Mann & D. A. Sabatino (Eds.), The first review of special education (Vol. 1). Philadelphia: JSE Press, 1973.
- Ysseldyke, J. E., & Salvia, J. Diagnostic-prescriptive teaching: Two models. Exceptional Children, 1974, 41, 17-32.

Zigmond, N. Learning patterns in children with learning disabilities.

Seminars in Psychiatry, 1969, 1, 344-353.

Zurif, E. B., & Carson, G. Dyslexia in relation to cerebral dominance and temporal analysis. Neuropsychologia, 1970, 8, 351-361.

Footnotes

¹Bugsby and Hurd, 1968; Gates, 1926; Harrington and Durrell, 1955; Leider, 1968; Robinson, Letton, and Mozzi, 1958; Rudnick and Sterritt, 1966; and Stuart, 1967.

²Bergan, 1967; Gates, 1922; Hurd, 1967; Nila, 1953; Phelan, 1940; and, Schorr and Svagr, 1966.

³Egeland, DiNello, and Carr, 1970; Hirshoren, 1969; Mueller, 1969; and Westinghouse, 1969.

⁴deHirsch, Jansky, and Langford, 1966; Egeland, DiNello, and Carr, 1970; Giebink and Birch, 1970; Keogh, 1965; and, Keogh and Smith, 1967.

⁵Bagford, 1968; Barrett, 1965b; deHirsch, Jansky, and Langford, 1966; Deputy, 1930; Egeland, DiNello, and Carr, 1970; Gates, 1939; Gates, 1940; Gates, Bond, and Russell, 1939; Gavel, 1958; Lee, Clark, and Lee, 1934; Muehl and Kremenak, 1966; Olson, 1958; Potter, 1959; Smith, 1928; Steinback, 1940; Weiner and Feldman, 1963; Wilson, 1942; and, Wilson and Burke, 1937.

⁶Bagford, 1969; Chall, Rosewell, and Blumenthal, 1963; deHirsch, Jansky, and Langford, 1966; and, Egeland, DiNello, and Carr, 1970.

⁷Beery, 1967; Birch and Belmont, 1965; Bonsall and Dornbush, 1969; Busby and Hurd, 1968; Connor, 1969; Goetzinger, Dirks, and Baer, 1960; Hammill, Larsen, Parker, Bagley, and Sanford, Note 1; Harrington and Durrell, 1955; Katz and Deutsch, 1963; Koppitz, 1975; Lachman, 1960; Larsen, Rogers, and Sowell, 1976; Lovell and Gorton, 1968; Raab, Deutsch, and Friedman, 1960; Sabatino and Ysseldyke, 1972; Weiner, Wepman, and Morency, 1965; Zigmond, 1969; and, Zurif and Carson, 1970.

⁸Dornbush and Basow, 1970; Eilenhammer, 1966; Flynn and Byrne, 1970; Goetzinger, Dirks, and Baer, 1960; Harrington and Durrell, 1955; Lingren, 1969; Lovell and Gorton, 1968; Monroe, 1932; Plessas, 1963; Poling, 1953; Silver, 1961; Weiner, Wepman, and Morency, 1965; Wepman, 1960; Zigmond, 1969; and, Zurif and Carson, 1970.

Table 1
Median Correlation Coefficients Between ITPA Subtests
and Measures of Academic Performance

<u>ITPA</u> subtests	Academic skills		
	Reading ^a	Spelling ^b	Arithmetic ^c
Auditory reception	.24	NS	.31
Visual reception	.25	NS	.24
Auditory association	.39	NS	.40
Visual association	.27	NS	.31
Verbal expression	.21	NS	.25
Motor expression	NS	NS	.22
Grammatical closure	.42	.41	.40
Auditory sequential memory	.31	NS	.27
Visual sequential memory	.24	NS	.26
Auditory closure	.29	NS	NS
Visual closure	NS	NS	NS
Sound blending	.38	.31	NS
Total score	.42	.30	.51

Note. Adapted from Newcomer and Hammill (1975).

^aBased on 820 coefficients.

^bBased on 178 coefficients.

^cBased on 154 coefficients.

Table 2

Median Coefficients Depicting the Concurrent Criterion Validity
of Selected Tests of Visual Perception
as Predictors of Academic Achievement

Visual-perceptual measures	Academic abilities	
	Reading	Arithmetic
<u>DTVP</u>		
Eye hand	.32	.32
Figure ground	.27	.40
Form constancy	NS	.40
Position in space	.29	.43
Spatial relations	.32	.42
Total	.34	.51
<u>WISC</u>		
Block design	NS	---
Coding	NS	.27
<u>Bender-Gestalt</u>	NS	.25
<u>Birch-Belmont-like</u>	NS	---
<u>Metropolitan Readiness Test</u>		
Matching	.21	---

Note. Adapted from Larsen and Hammill (1975). The number of coefficients contributing to each median was not specified.

Table 3

Median Coefficients Depicting the Relationship Between
Visual-Perceptual Abilities and Academic Achievement

Visual-perceptual ability	Academic achievement			
	Reading	Arithmetic	Spelling	Total
Memory	NS	.26	NS	NS
Discrimination	.26	.30	.20	.26
Spatial relations	.29	.26	.24	.28
Audio-visual Integration	NS	---	---	NS
Total	.24	.29	.23	.25

Note. Adapted from Larsen and Hammill (1975). The number of coefficients contributing to each median was not specified.

Table 4

Median Coefficients Associated With Auditory Abilities
and Reading Performance

Auditory Ability	Reading Skill		
	Word Recognition	Reading Comprehension	Composite Reading
Auditory-visual Integration	NS	.37	NS
Sound Blending	.24	NS	.31
Memory	.22	NS	NS
Auditory Discrimination- phonemic	.26	.26	.17
Auditory Discrimination- nonphonemic	NS	NS	NS

Note. Adapted from Hammill and Larsen (1974b).

Table 5
Median Correlations Between Specific Visual-Motor Abilities
and Later Achievement

Visual-Motor Ability	Achievement					Total
	Comprehension	Reading Vocabulary	Writing	Spelling	Arithmetic	
Visual Discrimination	.35 (59)	.24 (20)	.18 (2)	.25 (2)	.18 (2)	.33 (85)
Auditory-Visual Integration	.46 (2)	.39 (1)	---	---	---	.43 (3)
Gross and Fine Motor	.15 (6)	---	.23 (6)	.17 (6)	---	.16 (18)
Laterality and Body Image	.15 (2)	---	.15 (2)	.18 (2)	---	.16 (6)
Total	.35 (69)	.24 (21)	.18 (10)	.17 (10)	.18 (2)	.27 (112)

Note. Numbers in parentheses indicate the number of correlation coefficients contributing to median value.

Table 6
 Percentage of Significant Discriminations Between Reading Groups
 by ITPA Subtests

<u>ITPA</u> subtest	I.Q. controlled	I.Q. not controlled
Auditory reception	25	31
Visual reception	0	8
Auditory association	22	46
Visual association	11	31
Verbal expression	0	15
Motor expression	13	15
Grammatical closure	25	52
Auditory sequential memory	20	43
Visual sequential memory	13	36
Auditory closure	17	17
Visual closure	17	17
Sound blending	33	57

Note. Adapted from Newcomer and Hammill (1975).

Table 7

Percentage of Successful Discriminations Between Reading Groups
by Various Measures of Visual Abilities and Auditory Abilities

Type of ability measure	Percent of successful discriminations					
	IQ not controlled		IQ controlled		Total	
Visual abilities ^a						
Visual discrimination	50	(2)	25	(4)	33	(6)
Visual memory	38	(8)	33	(3)	36	(11)
Auditory-Visual Integration	100	(9)	100	(1)	100	(10)
Miscellaneous	43	(7)	27	(11)	30	(10)
Total	62	(26)	32	(19)	51	(37)
Auditory abilities ^b						
Auditory discrimination	88	(8)	100	(5)	92	(13)
Auditory memory	100	(3)	0	(1)	75	(4)
Sound blending	100	(3)	---		100	(3)
Auditory reception	100	(9)	---		100	(9)
Miscellaneous	77	(9)	50	(2)	73	(11)
Total	91	(32)	75	(8)	88	(40)

Note. Numbers in parentheses indicate the number of measures upon which percentage is based.

^aBased upon 16 studies.

^bBased upon 15 studies.

Table 8

Success of Training Differential Abilities
by Various Training Programs

Reviewer	Reviewer's judgement of quality of research ^a	Number of measures/ (studies)	Percent success ^b
Frostig			
Hallahan & Cruickshank, 1973	P	3/(3)	100
Hammill et al., 1974	G	8/(8)	12.5
	P	8/(8)	75
Ysseldyke, 1973	N	1/(1)	100
Subtotal		20/(20)	55
Kephart-Getman			
Goodman & Hammill, 1973	G	25/(11)	16
Hallahan & Cruickshank, 1973	G	6/(3)	17
	P	6/(2)	100
Keogh, 1974	N	1/(1)	0
Kleisius, 1972	G	3/(1)	100
Subtotal		41/(18)	34
Delacato			
Hallahan & Cruickshank, 1973	G	3/(2)	0
	P	14/(13)	93
Subtotal		17/(15)	76
ITPA based			
Hammill & Larsen 1974	N	85/(13)	30 ^c
Sedlack & Weener, 1973	N	13/(1)	8
Subtotal		13/(1)	8

Table 8 Continued

Reviewer	Reviewer's judgement of quality of research	Number of measures/ (studies)	Percent success
Miscellaneous perceptual-motor			
Hallahan & Cruickshank, 1973	P	7/(7)	86
Hammill, 1972	G	3/(3)	33
	P	3/(3)	67
Keogh, 1974	N	1/(1)	100
Kleisius, 1972	G	3/(2)	67
Meikle & Kilpatrick, 1971	N	14/(1)	28
Proger et al., 1973	N	17/(2)	35
Robinson, 1972	N	2/(2)	100
Sabatino, 1973	N	14/(4)	21
Whisler, 1975	N	1/(1)	100
Ysseldyke, 1973	N	7/(2)	43
Subtotal		72/(28)	43
Grand totals			
Poor only		36/(41)	88
Good only		12/(51)	24

^aReviewers categorized research design as G = Good; P = Poor; or, N = No Judgement.

^bPercentage of measures from the studies which show significant differences between trained and untrained groups.

^cThis is the median of the percentages of success across subtests. This figure is not included in the totals, since the true success rate was not known.

Table 9

The Percentage of Analyses, by Subgroup, Which Found Psycholinguistic Training to be Successful

Subgroups	ITPA Subtests							Total			
	Reception Aud. Vis.	Association Aud. Vis.	Expression Verb. Motor	Closure Grammatic Vis.	Sequential Aud.	Memory Vis.					
Subjects											
Retarded	33	25	13	33	40	44	22	--	22	20	50
Disadvantaged	27	29	59	50	50	29	27	--	33	21	40
Preschool	27	27	54	45	42	27	27	--	33	18	31
Elementary	31	25	43	46	57	54	23	--	31	29	56
Approach											
Prescriptive	17	17	33	57	57	57	17	--	29	14	50
Nonindividualized	32	28	52	39	50	33	25	20	37	26	46
Activities											
Selected	29	29	38	29	44	50	29	--	33	25	42
PLDK	27	30	42	40	55	30	17	--	30	18	47
Total	28	24	48	44	52	40	23	20	35	23	47

89

Note. Adapted from Hammil and Larsen (1974a).

Table 10

Effects of DD-PT Training Programs on Achievement and Intelligence

Reviewer	Quality of Research Design	Reading and school readiness		Reading and general achievement		Intelligence		Other ^b		Total percent successes	
		Number of measures/ (studies)	Percent success	Number of measures/ (studies)	Percent success	Number of measures/ (studies)	Percent success	Number of measures/ (studies)	Percent success	All studies	Excluding poor studies
Kephart-Getman											
Goodman & Hammill, 1973	G	10/(9)	40	7/(7)	57	6/(4)	0	1/(1)	100		
Hallahan & Cruickshank, 1973	G	-	-	-	-	-	-	3/(1)	67		
	P	-	-	-	-	1/(1)	0	1/(1)	100		
Keogh, 1974	N	-	-	2/(1)	100	-	-	1/(1)	100		
Klesius, 1972	G	1/(1)	0	2/(1)	0	-	-	1/(1)	0		
Subtotals		11/(10)	36	11/(9)	55	7/(5)	0	7/(5)	71	42	41
Frostig											
Hallahan & Cruickshank, 1973	P	-	-	1/(1)	0	-	-	-	-		
Hammill et al., 1973	N	7/(7)	43	14/(14)	7	-	-	-	-		
Robinson, 1972	N	1/(1)	0	-	-	-	-	-	-		
Ysseldyke, 1973	N	-	-	-	-	1/(1)	0	-	-		
Subtotals		8/(8)	38	15/(15)	7	1/(1)	0	-	-	17	17

Table 10 Continued

Reviewer	Quality of Research Design	Reading and school readiness		Reading and general achievement		Intelligence		Other		Total percent successes	
		Number of measures/ (studies)	Percent success	Number of measures/ (studies)	Percent success	Number of measures/ (studies)	Percent success	Number of measures/ (studies)	Percent success	All studies	Excluding poor studies
Delacato											
Hallahan & Cruickshank, 1973	G	-	-	2/(2)	0	-	-	-	-		
	P	-	-	1/(1)	0	-	-	2/(2)	50		
Subtotals		-	-	3/(3)	0	-	-	2/(2)	50	20	0
Other visual-perceptual											
Hallahan & Cruickshank, 1973	P	-	-	7/(7)	71	3/(3)	67	5/(5)	80		
Hammill, 1972	N	-	-	5/(5)	40	-	-	-	-		
Robinson, 1972	N	1/(1)	0	1/(1)	0	-	-	1/(1)	100		
Subtotals		1/(1)	0	13/(13)	54	3/(3)	67	6/(6)	83	61	38
Auditory perceptual											
Robinson, 1972	N	-	-	5/(4)	40	-	-	-	-		
Sabatino, 1973	N	4/(3)	50	4/(2)	0	-	-	1/(1)	100		
Ysseldyke, 1973	P	1/(1)	0	-	-	-	-	-	-		
Subtotals		5/(4)	40	9/(6)	22	-	-	1/(1)	100	33	36

Table 10 Continued

Reviewer	Quality of Research Design	Reading and school readiness		Reading and general achievement		Intelligence		Other		Total percent successes	
		Number of measures/ (studies)	Percent success	Number of measures/ (studies)	Percent success	Number of measures/ (studies)	Percent success	Number of measures/ (studies)	Percent success	All studies	Excluding poor studies
		Other ability									
Klesius, 1972	G	3/(3)	67	6/(5)	50	3/(3)	0	-	-	-	-
Sedlack & Weener, 1973	N	-	-	1/(1)	0	-	-	-	-	-	-
Subtotals		3/(3)	67	7/(6)	43	3/(3)	0	-	-	38	38
Totals											
All studies	G,N,P	28/(26)	39	58/(52)	33	14/(12)	14	16/(14)	75	38	38
Excluding Poor Studies	G,N Only	27/(25)	41	49/(45)	29	11/(9)	0	8/(6)	75	36	36

^aNote. Percent success is the percentage of measures from the studies which show significant differences between trained and untrained groups.

^aReviewers categorized research design as G = Good; P = Poor; or, N = No Judgement.

^bOther measures included language, social skills, adaptive behavior, handwriting, naming digits, following directions, etc.

CENTER FOR THE STUDY OF READING

READING EDUCATION REPORTS

- No. 1: Durkin, D. Comprehension Instruction--Where Are You?, October 1977.
- No. 2: Asher, S. R. Sex Differences in Reading Achievement, October 1977.
- No. 3: Adams, M., Anderson, R. C., & Durkin, D. Beginning Reading: Theory and Practice, October 1977.
- No. 4: Jenkins, J. R., & Pany, D. Teaching Reading Comprehension in the Middle Grades, January 1978.

CENTER FOR THE STUDY OF READING
TECHNICAL REPORTS

* Available only through ERIC

- *No. 1: Halff, H. M. Graphical Evaluation of Hierarchical Clustering Schemes, October 1975. (ERIC Document Reproduction Service No. ED 134 926, 11p., HC-\$1.67, MF-\$.83)
- *No. 2: Spiro, R. J. Inferential Reconstruction in Memory for Connected Discourse, October 1975. (ERIC Document Reproduction Service No. ED 136 187, 81p., HC-\$4.67, MF-\$.83)
- *No. 3: Goetz, E. T. Sentences in Lists and in Connected Discourse, November 1975. (ERIC Document Reproduction Service No. ED 134 927, 75p., HC-\$3.50, MF-\$.83)
- *No. 4: Alessi, S. M., Anderson, T. H., & Biddle, W. B. Hardware and Software Considerations in Computer Based Course Management, November 1975. (ERIC Document Reproduction Service No. ED 134 928, 21p., HC-\$1.67, MF-\$.83)
- *No. 5: Schallert, D. L. Improving Memory for Prose: The Relationship Between Depth of Processing and Context, November 1975. (ERIC Document Reproduction Service No. ED 134 929, 37p., HC-\$2.06, MF-\$.83)
- *No. 6: Anderson, R. C., Goetz, E. T., Pichert, J. W., & Halff, H. M. Two Faces of the Conceptual Peg Hypothesis, January 1976. (ERIC Document Reproduction Service No. ED 134 930, 29p., HC-\$2.06, MF-\$.83)
- *No. 7: Ortony, A. Names, Descriptions, and Pragmatics, February 1976. (ERIC Document Reproduction Service No. ED 134 931, 25p., HC-\$1.67, MF-\$.83)
- *No. 8: Mason, J. M. Questioning the Notion of Independent Processing Stages in Reading, February 1976. (Journal of Educational Psychology, 1977, 69, 288-297)
- *No. 9: Siegel, M. A. Teacher Behaviors and Curriculum Packages: Implications for Research and Teacher Education, April 1976. (ERIC Document Reproduction Service No. ED 134 932, 42p., HC-\$2.06, MF-\$.83)
- *No. 10: Anderson, R. C., Pichert, J. W., Goetz, E. T., Schallert, D. L., Stevens, K. V., & Trollip, S. R. Instantiation of General Terms, March 1976. (ERIC Document Reproduction Service No. ED 134 933, 30p., HC-\$2.06, MF-\$.83)
- *No. 11: Armbruster, B. B. Learning Principles from Prose: A Cognitive Approach Based on Schema Theory, July 1976. (ERIC Document Reproduction Service No. ED 134 934, 48p., HC-\$2.06, MF-\$.83)
- *No. 12: Anderson, R. C., Reynolds, R. E., Schallert, D. L., & Goetz, E. T. Frameworks for Comprehending Discourse, July 1976. (ERIC Document Reproduction Service No. ED 134 935, 33p., HC-\$2.06, MF-\$.83)

- No. 13: Rubin, A. D., Bruce, B. C., & Brown, J. S. A Process-oriented Language for Describing Aspects of Reading Comprehension, November 1976. (ERIC Document Reproduction Service No. ED 136 188, 41p., HC-\$2.06, MF-\$.83)
- No. 14: Pichert, J. W., & Anderson, R. C. Taking Different Perspectives on a Story, November 1976. (ERIC Document Reproduction Service No. ED 134 936, 30p., HC-\$2.06, MF-\$.83)
- No. 15: Schwartz, R. M. Strategic Processes in Beginning Reading, November 1976. (ERIC Document Reproduction Service No. ED 134 937, 19p., HC-\$1.67, MF-\$.83)
- No. 16: Jenkins, J. R., & Pany, D. Curriculum Biases in Reading Achievement Tests, November 1976. (ERIC Document Reproduction Service No. ED 134 938, 24p., HC-\$1.67, MF-\$.83)
- No. 17: Asher, S. R., Hymel, S., & Wigfield, A. Children's Comprehension of High- and Low-Interest Material and a Comparison of Two Cloze Scoring Methods, November 1976. (ERIC Document Reproduction Service No. ED 134 939, 32p., HC-\$2.06, MF-\$.83)
- No. 18: Brown, A. L., Smiley, S. S., Day, J. D., Townsend, M. A. R., & Lawton, S. C. Intrusion of a Thematic Idea in Children's Comprehension and Retention of Stories, December 1976. (ERIC Document Reproduction Service No. ED 136 189, 39p., HC-\$2.06, MF-\$.83)
- No. 19: Kleiman, G. M. The Prelinguistic Cognitive Basis of Children's Communicative Intentions, February 1977. (ERIC Document Reproduction Service No. ED 134 940, 51p., HC-\$3.50, MF-\$.83)
- No. 20: Kleiman, G. M. The Effect of Previous Context on Reading Individual Words, February 1977. (ERIC Document Reproduction Service No. ED 134 941, 76p., HC-\$4.67, MF-\$.83)
- No. 21: Kane, J. H., & Anderson, R. C. Depth of Processing and Interference Effects in the Learning and Remembering of Sentences, February 1977. (ERIC Document Reproduction Service No. ED 134 942, 29p., HC-\$2.06, MF-\$.83)
- No. 22: Brown, A. L., & Campione, J. C. Memory Strategies in Learning: Training Children to Study Strategically, March 1977. (ERIC Document Reproduction Service No. ED 136 234, 54p., HC-\$3.50, MF-\$.83)
- No. 23: Smiley, S. S., Oakley, D. D., Worthen, D., Campione, J. C., & Brown, A. L. Recall of Thematically Relevant Material by Adolescent Good and Poor Readers as a Function of Written Versus Oral Presentation, March 1977. (ERIC Document Reproduction Service No. ED 136 235, 23p., HC-\$1.67, MF-\$.83)
- No. 24: Anderson, R. C., Spiro, R. J., & Anderson, M. C. Schemata as Scaffolding for the Representation of Information in Connected Discourse, March 1977. (ERIC Document Reproduction Service No. ED 136 236, 18p., HC-\$1.67, MF-\$.83)

- No. 25: Pany, D., & Jenkins, J. R. Learning Word Meanings: A Comparison of Instructional Procedures and Effects on Measures of Reading Comprehension with Learning Disabled Students, March 1977. (ERIC Document Reproduction Service No. ED 136 237, 34 p., HC-\$2.06, MF-\$.83)
- No. 26: Armbruster, B. B., Stevens, R. J., & Rosenshine, B. Analyzing Content Coverage and Emphasis: A Study of Three Curricula and Two Tests, March 1977. (ERIC Document Reproduction Service No. ED 136 238, 22 p., HC-\$1.67, MF-\$.83)
- No. 27: Ortony, A., Reynolds, R. E., & Arter, J. A. Metaphor: Theoretical and Empirical Research, March 1977. (ERIC Document Reproduction Service No. ED 137 752, 63 p., HC-\$3.50, MF-\$.83)
- No. 28: Ortony, A. Remembering and Understanding Jabberwocky and Small-Talk, March 1977. (ERIC Document Reproduction Service No. ED 137 753, 36 p., HC-\$.206, MF-\$.83)
- No. 29: Schallert, D. L., Kleiman, G. M., & Rubin, A. D. Analysis of Differences Between Oral and Written Language, April 1977.
- No. 31: Nash-Webber, B. Anaphora: A Cross-Disciplinary Survey, April 1977.
- No. 32: Adams, M. J., & Collins, A. A Schema-Theoretic View of Reading Comprehension, April 1977. (ERIC Document Reproduction Service No. ED 142 971, 49 p., HC-\$2.06, MF-\$.83)
- No. 33: Huggins, A. W. F. Syntactic Aspects of Reading Comprehension, April 1977. (ERIC Document Reproduction Service No. ED 142 972, 68 p., HC-\$3.50, MF-\$.83)
- No. 34: Bruce, B. C. Plans and Social Actions, April 1977.
- No. 35: Rubin, A. D. A Theoretical Taxonomy of the Differences Between Oral and Written Language, January 1978.
- No. 36: Nash-Webber, B., & Reiter, R. Anaphora and Logical Form: On Formal Meaning Representations for Natural Language, April 1977. (ERIC Document Reproduction Service No. ED 142 973, 42 p., HC-\$2.06, MF-\$.83)
- No. 37: Adams, M. J. Failures to Comprehend and Levels of Processing in Reading, April 1977.
- No. 38: Woods, W. A. Multiple Theory Formation in High-Level Perception, April 1977.
- No. 40: Collins, A., Brown, J. S., & Larkin, K. M. Inference in Text Understanding, December 1977.
- No. 41: Anderson, R. C., & Pichert, J. W. Recall of Previously Unrecallable Information Following a Shift in Perspective, April 1977. (ERIC Document Reproduction Service No. ED 142 974, 37p., HC-\$2.06, MF-\$.83)

- No. 42: Mason, J. M., Osborn, J. H., & Rosenshine, B. V. A Consideration of Skill Hierarchy Approaches to the Teaching of Reading, December 1977.
- No. 43: Collins, A., Brown, A. L., Morgan, J. L., & Brewer, W. F. The Analysis of Reading Tasks and Texts, April 1977.
- No. 44: McClure, E. Aspects of Code-Switching in the Discourse of Bilingual Mexican-American Children, April 1977. (ERIC Document Reproduction Service No. ED 142 975, 38 p., HC-\$2.06, MF-\$.83)
- No. 45: Schwartz, R. M. Relation of Context Utilization and Orthographic Automaticity in Word Identification, May 1977.
- No. 46: Anderson, R. C., Stevens, K. C., Shiffrin, Z., & Osborn, J. Instantiation of Word Meanings in Children, May 1977. (ERIC Document Reproduction Service No. ED 142 976, 22 p., HC-\$1.67, MF-\$.83)
- No. 47: Brown, A. L. Knowing When, Where, and How to Remember: A Problem of Metacognition, June 1977.
- No. 48: Brown, A. L., & DeLoache, J. S. Skills, Plans, and Self-Regulation. July 1977.
- No. 50: Anderson, R. C. Schema-Directed Processes in Language Comprehension, July 1977. (ERIC Document Reproduction Service No. ED 142 977, 33 p., HC-\$2.06, MF-\$.83)
- No. 51: Brown, A. L. Theories of Memory and the Problems of Development: Activity, Growth, and Knowledge, July 1977.
- No. 52: Morgan, J. L. Two Types of Convention in Indirect Speech Acts, July 1977.
- No. 53: Brown, A. L., Smiley, S. S., & Lawton, S. C. The Effects of Experience on the Selection of Suitable Retrieval Cues for Studying from Prose Passages, July 1977.
- No. 54: Fleisher, L. S., & Jenkins, J. R. Effects of Contextualized and De-contextualized Practice Conditions on Word Recognition, July 1977.
- No. 56: Anderson, T. H., Standiford, S. N., & Alessi, S. M. Computer Assisted Problem Solving in an Introductory Statistics Course, August 1977.
- No. 57: Barnitz, J. G. Interrelationship of Orthography and Phonological Structure in Learning to Read, January 1978.
- No. 58: Mason, J. M. The Role of Strategy in Reading in the Mentally Retarded, September 1977.
- No. 59: Mason, J. M. Reading Readiness: A Definition and Skills Hierarchy from Preschoolers' Developing Conceptions of Print, September 1977.
- No. 60: Spiro, R. J., & Esposito, J. Superficial Processing of Explicit Inferences in Text, December 1977.
- No. 65: Brewer, W. F. Memory for the Pragmatic Implications of Sentences, October 1977.

- No. 66: Brown, A. L., & Smiley, S. S. The Development of Strategies for Studying Prose Passages, October 1977.
- No. 68: Stein, N. L., & Nezworski, T. The Effects of Organization and Instructional Set on Story Memory, January 1978.
- No. 77: Nash-Webber, B. L. Inference in an Approach to Discourse Anaphora, January 1978.
- No. 78: Gentner, D. On Relational Meaning: The Acquisition of Verb Meaning, December 1977.
- No. 79: Royer, J. M. Theories of Learning Transfer, January 1978.
- No. 80: Arter, J. A., & Jenkins, J. R. Differential Diagnosis-Prescriptive Teaching: A Critical Appraisal, January 1978.

