

Rowan University

Rowan Digital Works

Theses and Dissertations

4-30-2001

An interrater reliability study of the Psychological Processing Checklist

Sophia Lusinski
Rowan University

Follow this and additional works at: <https://rdw.rowan.edu/etd>



Part of the [Educational Psychology Commons](#)

Recommended Citation

Lusinski, Sophia, "An interrater reliability study of the Psychological Processing Checklist" (2001). *Theses and Dissertations*. 1584.

<https://rdw.rowan.edu/etd/1584>

This Thesis is brought to you for free and open access by Rowan Digital Works. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Rowan Digital Works. For more information, please contact graduateresearch@rowan.edu.

An Interrater Reliability Study of the Psychological Processing Checklist

By
Sophia Lusinski

A Thesis

Submitted in partial fulfillment of the requirements of the
Master of Arts Degree
of
The Graduate School
At
Rowan University
2001

Approved by _____
Professor

Date Approved 4/30/01

ABSTRACT

Sophia Lusinski

An Interrater Reliability Study of the Psychological Processing Checklist
2001

Advisor: Dr. Klanderma

School Psychology

This research was undertaken to conduct an interrater reliability study of the Psychological Processing Checklist, which proposes to assess in children deficits in cognitive processes.

A sample of 30 Hispanic students, male and female, ages ranging from 6 to 10, enrolled in a Bilingual Program in a district of a city in Southern New Jersey was randomly selected to participate.

Two different teachers rated each individual: one bilingual and/or ESL, the other a mainstream regular classroom teacher. Both used the Psychological Processing Checklist to rate each student.

It was hypothesized that there would be a high correlation between raters. This study was analyzed by the use of correlation of raw scores on a rating scale (PPC). The raw scores of items rated were correlated to determine a reliability coefficient.

The results of the study indicated that there was a high correlation between the teachers rating Hispanic bilingual students with the Psychological Processing Checklist.

Mini Abstract

Sophia Lusinski

An Interrater Reliability Study of The Psychological Processing Checklist
2001

Advisor: Dr. Klanderma

School Psychology

The purpose of this study was to conduct an interrater reliability study of the Psychological Processing Checklist (PPC) with Hispanic bilingual students in a school district in South Jersey.

The results of the study indicated that there was a significant correlation between teachers rating thirty Hispanic bilingual students enrolled in a Bilingual Program in a district in South Jersey with the Psychological Processing Checklist.

ACKNOWLEDGEMENTS

I wish to acknowledge my appreciation to Dr. John Klanderman for his assistance with the preparation of this research project.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	ii
CHAPTER I: The Problem.....	1
Purpose of Study.....	1
Hypothesis.....	2
Theory.....	2
Definition of Terms.....	5
Assumptions.....	6
Limitations.....	6
Overview.....	7
CHAPTER II: Review of the Literature.....	8
Reliability.....	8
Assessments used with Hispanic Bilingual Students.....	11
Testing of Language Minorities.....	16
Difficulties in Assessing Learning Disabilities with Hispanic/Bilingual Students.....	18
Summary.....	20
CHAPTER III: Design of the Study.....	22
Sample.....	22
Measures.....	22
Design of Study.....	23
Testable Hypothesis.....	23
Analysis.....	24
Summary.....	24
CHAPTER IV: Analysis of Results.....	25
Interpretation of Results.....	25
Summary.....	26
CHAPTER V: Summary and Conclusions.....	30
Summary.....	30
Discussion.....	31
Conclusions.....	31
Implication for Further Study.....	31
References.....	33
Appendices.....	36

LIST OF TABLES AND GRAPHS

Table 3.1	Sample.....	22
Figure I	Interrater Reliability of Total Score.....	28
Figure II	Interrater Reliability of Auditory Processing.....	29
Figure III	Interrater Reliability in Attention.....	29

CHAPTER I

THE PROBLEM

Need

The Psychological Processing Checklist (PPC), an instrument developed by the Illinois State University is currently undergoing its final stages of development. At this stage, there is a need for interrater reliability to add to the research establishing this checklist as a credible identification instrument. There also tends to be an over classification of Hispanic students and there is a compelling need for developing an instrument that would be valid and reliable for use with Hispanic students including those who are limited English proficient (LEP). Thus the reliability of the Psychological Processing Checklist is well worth studying. In this study, the PPC is going to be administered in order to examine its interrater reliability. This scale provides information that assists in the assessment of cognitive processing deficits and in the development of interventions to facilitate students; learning. It is free of cultural bias and students' attitude is not a factor in determining outcomes. Rating is done by teachers based on observed students' behavior. If it is determined to be reliable, Hispanic students as well as teachers will benefit from it.

Purpose

The purpose of this research is to conduct an interrater reliability study of the Psychological Processing Checklist (PPC) with bilingual students in a school district in Southern New Jersey. If this scale proves to be reliable, upon collection and analysis of statistical data, it may be adopted by public school systems as a measure that will enable educators and practitioners in the field to better assist students with learning disabilities among the Hispanic population.

Hypothesis

Testable Hypothesis

It is hypothesized that there will be a high degree of correlation in the rating of bilingual students in a school district in Southern New Jersey with the Psychological Processing Checklist (PPC) by a bilingual or ESL teacher and a mainstream regular classroom teacher.

Null Hypothesis

It is hypothesized that there will be no correlation in the rating of Hispanic bilingual students in a school district in Southern New Jersey with the Psychological Processing Checklist (PPC) by a bilingual and/or ESL teacher and a mainstream classroom teacher.

Theory

The purpose of this research is to conduct an interrater reliability study of The Psychological Processing Checklist, which proposes to assess in children deficits in cognitive processes such as Auditory Processing, Visual Processing, Visual-Motor, Social Perception, Organization and Attention. Theories pertinent to information processing with learning disabilities were examined.

Kantowitz (1994) defined Information-Processing Theory as one that deals with how people “Attend to, select and internalize information and how they later use it to make decisions and guide their behavior” (p. 245).

The primary domain of information processing psychology is mental processes and intelligent behavior covering memory and thinking. At the general level, theory is directed at properties of mechanisms responsible for the comprehension, storage, retrieval and utilization of information that may initiate in the external environment or in the internal mental states (Kantowitz, 1994).

According to information-processing theories, learning disabled students are deficient in information processing. They have restricted access to information processes activities. Certain inefficient patterns of processing are commonly shared among children who have learning disabilities (Swanson, 1987).

AUDITORY PROCESS. The auditory process is a stage where inner speech occurs. Auditory memory is necessary to remember words, to read and acquire written language as well as to spell words. Many learning disabled students frequently demonstrate a deficit in this area.

The most basic of all auditory processes is auditory discrimination.. It permits sounds to be distinguished from each other. If this area is affected in early life, then all language behavior is affected. A deficit in this area causes confusion of similar words in both speech and reading (Travis, 1971).

VISUAL PROCESS. Visual activity is an important factor in language development because of its effects on reading and writing skills. If visual acuity is poor, eyesight is inadequate causing impediment in the learning of reading and writing.

Visual perception allows organisms to maintain on-going contact with the environment. It has several components that affect the acquisition of written language. Those components are visual recognition, visual memory and visual imagery.

Visual discrimination is one component necessary for writing. Writing is not possible until letters can be discriminated.

Visual recognition is the ability to recognize shape. When it assumes integration memory occurs. Many learning disabled students have difficulties in perceiving and recognizing letters correctly.

Visual memory is needed to record, store, and retrieve information for auditory, visual and motor processes. Acquisition of these processes is essential before written language can be mastered.

Visual imagery is the ability to recall all or parts of experiences auditorily or visually pictured in the mind. If this cannot be internalized, children are unable to use written words (Travis, 1971).

MOTOR PROCESS. Motor development follows a sequential pattern. This pattern starts when a child first holds a pencil and ends when writing becomes a means of communication. Visual perception and motor coordination are essential in handwriting (Kavale, Forness & Bender, 1987).

ORGANIZATION. Organization is a memory strategy. Children who use this strategy mentally group materials to be remembered in meaningful clusters of closely associated items so they have to remember only one part of a cluster to gain access to the rest. This is an area in which nearly all learning disabled students evidence difficulty. Many learning disabled children have trouble developing concepts. The requisite for concept formation is the ability to organize data. Studies have demonstrated that children who do not use organized strategies can be taught to do so even when they have a learning disability (Kavale, Forness & Bender, 1987).

SOCIAL PERCEPTION. Social problems evidenced by learning disabled children may frequently be attributed to difficulty in interpreting figurative speech. There may not only be a problem with the perception of oral language but also with its expression in social settings. Learning disabled children may not appreciate the need to display positive social behavior such as an expression of concern and of verbal affection. This

lack of interpersonal exchanges leads to rejection, isolation and a feeling of being different. Studies on social skills concluded that children can be taught appropriate social skills through modeling and reinforcing of appropriate social behaviors (Kavale, Forness & Bender, 1987).

ATTENTION. A number of experts suggest that attention deficits should be of primary focus in remediation of learning disabled children. A theory was developed concerning the automaticity of information processing in which the primary component is attention that is assumed to be both selective and of limited capacity. For the disabled reader, information at the visual and phonological level is not processed automatically thereby taking a great portion of attention and leaving little to focus at the semantics level and thus reducing understanding. Because of limited processing space, more space is consumed by decoding and less is available for comprehension (Kavale, Forness & Bender, 1987).

Definitions

For the purpose of this study, the following terms are defined.

Hispanic. An individual of Spanish background who may encompass a variety of cultures and races. He/She may or may not use Spanish as the home language.

Reliability. The consistency of scores obtained by the same persons when retested with the identical test or with an equivalent form of the test.

Language Proficiency. Refers to a person's learned, functional capability to use a language system in comprehension, speaking, reading and writing.

Psychological Processing Scale (PPC). Scale developed to assess processing deficits. (currently undergoing validity and reliability studies).

Reliability Coefficient. Is the percentage of score variance attributable to different scores. A reliability coefficient of .85 means that 85% of the variance in test scores depends on true variance in trait measured and 15 % depends on error variance.

Interrater Reliability scorer reliability that can be found by having a sample test independently scored by two examiners. The two scores are correlated and the correlation coefficient is the interrater reliability.

Measurement Error. Is an alternative way of expressing test reliability. It is standard error of a score. It is a measure suited to the interpretation of individual scores.

Limited English Proficient (LE P) Any student who does not have English language skills in listening, speaking, reading and writing at the same level as a native speaker.

Learning Disability. Refers to a retardation, disorder or delayed development in one or more of the processes of speech, language, reading, writing, arithmetic or other school subject resulting from a psychological handicap.

Assumptions

For the purpose of this study, it is assumed that all teachers rating Bilingual students will use the same degree of integrity.

Limitations

This study is limited to 30 Hispanic students enrolled in the Bilingual Program in a school district in Southern New Jersey.

Overview

In Chapter I, the need and purpose of this thesis was stated. In Chapter II, relevant research concerning reliability, interrater reliability, usefulness of rating scales for Hispanic bilingual students, the implications of testing of language minority students and difficulty in identifying learning disabilities with Hispanic bilingual students will be reviewed. In Chapter III, the design of this study will be addressed. This design includes a description of the devices used for measurement, a testable hypothesis and an analysis of the results will be presented in Chapter IV. In Chapter V, a summary and relevant conclusions with discussions will be presented with implications for future research.

CHAPTER 2

REVIEW OF LITERATURE

In this chapter, this researcher will attempt to present a review of literature pertaining to Reliability, Interrater reliability, usefulness of psychological measures used with Hispanic bilingual students, the implications of testing of language minorities and the difficulty in identifying learning disabilities with Hispanic bilingual students.

The importance of reliability to his study is eminent as the purpose of this research is to conduct an interrater reliability study of the Psychological Processing Checklist (PPC). Reliability is one of the essential technical characteristics of an instrument. It is undertaken to provide information, which will be used to evaluate and make decisions about children. Estimating the reliability of an instrument is one way of knowing how much confidence can be placed in the results obtained from a particular instrument. The interrater reliability is the form of reliability used for the purpose of this study since it was hypothesized that there will be a high correlation in the rating of Hispanic bilingual students with the PPC by two different teachers.

Reliability

The reliability of a measure refers to its degree of dependability, stability, consistency, predictability and accuracy (Groth-Marnat, 1997). It is the rating of the precision of a given instrument. The concept of reliability in its various forms is central to the theory and practice of educational and psychological testing. To do its job well, an instrument should yield results that are consistent and pertinent to the specific purpose the administrator has in mind. A measure cannot be of much value if the score it yields

for one student one day is quite different from the score it would have yielded for him/her under similar conditions another day (Stanley, 1972).

Reliability is not a simple basic concept in assessment theory. It serves a particular purpose just as validity. Theoretically, the member that is observed as the reliability estimate of a given test represents the degree to which the constructed test overlaps a perfect measure of the characteristic of interest (Groth-Marnat, 1997).

Psychometrically, reliability is the squared correlation between the observed scores and the true scores for that trait of interest (Goldstein and Hersen, 1990, p. 27). Clinicians should hope for correlations of .90 or higher in tests that will be used to make decisions about individuals (Groth-Marnat, 1997).

There are several computational forms of reliability that are used depending on the characteristics that a test intends to measure, the type of items, and the needs of the test user. All these forms are methods to estimate the reliability of the instrument by systematically controlling for potential sources of error (Goldstein and Hersen, 1990).

The four primary methods of obtaining reliability are: *Test-retest*, which involves determining the extent to which the test produces consistent results on retesting (reliability time to time). *Alternate Forms*, which is the relative accuracy of a test at a given time (reliability form to form). *Split-Half* which is the internal consistency of the items of a test (reliability item to item). *Inter-scorer* or *Inter-rater*, which is the degree of agreement between two examiners (reliability scorer to scorer). Although there are the main types of reliability, there is a fifth type — the Kuder-Richardson, which is a measurement of the internal consistency of the test items like split-half type (Groth-

Marnat, 1997). For the purpose of this study, only the inter-rater reliability method will be the focus as it is relevant to the hypothesis being studied.

Inter-rater Reliability (or scorer reliability) is a method of estimating test reliability by using multiple examiners to score the same series of responses. In other words, two different individuals score a series of responses from a single client. A variation is to have two different examiners test the same client using the same test and then to determine how close their scores or ratings of the person are. The two sets of scores can then be correlated to determine a reliability coefficient. Any test that requires partial subjectivity in scoring should provide information on interscorer reliability (Goldstein and Hersen, 1990).

Generalizability Theory measurement error is common in all fields of educational and psychological testing. Specialists in the field have devoted a great deal of study in this area. Tests that are relatively free of measurement error are considered to be reliable, and test that contain relatively great measurement error are considered to be unreliable. The Classical test theory was developed in order to deal with the problem of measurement error. A central role in this theory is played by the concept of reliability, that is, the ratio of true to observed score variance. This approach to the assessment of reliability does not take into account multiple sources of error to consider various types of measurement (Kaplan and Saccuzo, 1982). The Generalizability theory, however, provides a flexible structure for examining the dependability of behavioral measurement. This theory assumes that a measurement taken on a person is only a random sample of that person s behavior. The usefulness of the measurement depends on the degree to which that sample allows to generalize accurately the behavior of the same person in a wider set of

situations. The concept of reliability is replaced in the Generalizability theory. Instead of asking how accurately a set of observed scores reflects their corresponding true scores, the generalizability theory asks how accurately a set of observations permits us to generalize about a person's behavior in a universal situation (Gruijter and Kamp, 1990).

The importance of reliability of tests is paramount since it is through a study of reliability that greater accuracy of assessments can be achieved. Without knowledge of reliability, no method of improving the consistency of an assessment can be adopted (Satterly, 1989). Considerations of usability of tests such as convenience, economy and interpretability are important only after questions of reliability have been answered satisfactorily. (Stanley, 1972).

Assessments Used with Hispanic Bilingual Students

Many assessment instruments used with Hispanic bilingual students were received for the purpose of this research/study. Although the reliability of some of the measures is weak, they are considered in light of their usefulness with the bilingual population.

The significantly increased number of students in grades K-12 in our schools whose native language is not English creates a need for educational professionals to make linguistic, intellectual and academic assessments of these students. As part of these assessments, standardized instruments are likely to be utilized. In recent years, legislative action in various states created a demand for instruments that can be used by schools to assess the English proficiency of such students in order to place them in the most appropriate instructional programs. Many assessments were developed in response to this need and their availability and the demands for their use has increased.

The Maculaitis Assessment Program (MAC) is an instrument widely used with language minority students. It is designed to assess the English language competencies of limited English proficient (LEP) students, grades K-12 in the four basic language skills of listening, speaking, reading and writing. According to the publishers, the MAC could be used for selection, placement, diagnosis, proficiency and achievement. This instrument is also available in Spanish. It was published in 1982. National norms are lacking. Consequently, the reliability of the instrument is limited. The test provides only one form and it is currently being revised (Maculaitis, 1985).

Another instrument that was reviewed for this study is the Spanish Assessment of Basic Education second edition (SABE/2). It is a series of norm-referenced tests for grades 1 through 8. This assessment is useful in determining Spanish-speaking students instructional needs and in planning for their transition to English instructional programs. It is also appropriate for evaluating bilingual education programs. This test was normed on a Spanish-dominant sample of over 10,000 students in grades 1 through 12. It was published in 1991-1994. The internal consistency reliability for the subtests at various grade levels fell between .97 and .67. This instrument represents advancement in the field of Bilingual Education.

The Terra Nova Spanish edition called Supera is an achievement test published in 1997. This battery is technically well built with reliability coefficients consistently in the .80s and .90s. It is one of the better batteries of its type. Reviewers report that usability studies were conducted to support test design and content. The test's content and approach match school curriculum framework and is therefore suggested for adoption consideration (CTB/McGraw-Hill, 1997)

The Language Assessment Battery (LAB) has 2 editions: English and Spanish. This battery was published in 1982. The English version of the LAB is used in grades K-12. Its purpose is to assess the language proficiency (reading, writing, speaking and listening) of all students whose English language proficiency is limited. For those students who have a Hispanic background, the administration of the Spanish version may be used to determine language dominance. This test is available in two forms at various grade levels. It is easy to administer. It is used to determine entry and exit to an ESL/Bilingual program. Despite its outdated norms, this instrument was widely used by many school districts for many years to the present time due to its economy, convenience and interpretability {(Abbott (New York City Public Schools Office of Testing Division of Curriculum and Instruction, 1982))}.

The Language Assessment Scales (LAS), like the LAB, also has two editions: Spanish and English. The English LAS is designed to measure the English language skills necessary for functioning in mainstream classes. It is therefore used to produce placement of language minority students into ESL/Bilingual Programs. This instrument was published in 1987-91. It is currently undergoing re-norming in the state of New Jersey. It has two forms and it measures the four major aspects of language: listening, speaking, reading and writing (De Avila, Duncan, 1987). Despite the fact that data reported in the technical report was not adequate to support reliability and validity, this scale was used and continues to be used by many districts as the assessment instrument for placement of ESL students.

The Woodcock Language Proficiency Battery Revised (WLPB-R) was designed in 1991. This battery can be administered to individuals from 2 to 90 years of age. It

measures language functioning. It can be used for evaluating English as a Second Language, diagnosis and program placement. Bateria Woodcock de Proficiencia en el Idioma is the available Spanish version. This battery is quite comprehensive. Inter-rater reliability is reported to be high. Many professionals extensively use this test with Hispanic students, especially in cases of evaluation for referral (Woodcock, 1991).

In the area of intelligence testing, the instrument that is most broadly used with Hispanic students is the Wechsler Intelligence Scale for children — Revised (WISC-R). Its equivalent Spanish version is the Escala de Inteligencia Wechsler Para Ninos — Revisada (EIWN-R) was published in 1982 and developed to most approximately suit all Spanish-speaking groups in the United States. This test was translated from English to Spanish, addressing Spanish-language subgroups by providing alternate expressions which permit its use with children from major Spanish-speaking groups in the United States: The Mexican-American, Puerto-Rican and Cuban-American. The WISC-R is technically well built with high reliability across the entire age range with average coefficients being .94, .90 and .96 respectively. (David Wechsler, 1974).

As for the norming of EIWN-R, researchers are encouraging the development of norms based on appropriate groups of Spanish-speaking children in the United States (Wechsler 1982). The newest intelligence scale, Edicion de Investigacion, used with the Puerto Rican student population is the WISC-R-PR, which is the Wechsler Intelligence Scale for Children-Revised-Adapted and Standardized in Puerto Rico (EIWN-PR), Spanish version.

Instruments that determine language dominance with Hispanic Students were also examined for this research. The James Language Dominance Test was designed in 1974

to yield a measure of a child's language dominance or bilingualism in both production and comprehension. Test results classify children as Spanish dominant, bilingual with Spanish as the home language, bilingual with both English and Spanish at home, English dominant but bilingual in comprehension, and English dominant in both comprehension and production. The English production scores were compared with monolingual English speakers. The technical manual of the tests does not provide information on reliability and validity information is limited (James, 1974).

The Pictorial Test of Bilingual and Language Dominance was developed in 1975-76 for bilingual screening and language assessment for school placement of Mexican-American children. This test was designed for use with children in grades Kindergarten and One. The test manual presents description of validity and reliability. The split-half reliability is very high: .85 for English and .89 for Spanish, with excellent reliability coefficients. This test was well constructed based on modern concepts of linguistic theory and measurements (Nelson, Fellner & Norrel, 1975). This instrument was proven very useful in Texas, California and other states.

Another instrument, also used to determine language dominance, is the Woodcock Language Battery Revised (WLPB-R), previously mentioned in this research. It is also available in Spanish. Current professionals extensively use this battery.

The Ber-Sil Spanish Test developed in 1972-77 assesses Spanish-speaking school-age children in their dominant language. This is a culture and language sensitive test patterned in the Peabody Picture Vocabulary Test, but with pictures that resemble the cultural environment of Spanish-Speaking children of diverse nationalities. The reliability data are reported difficult to interpret (Beringer, 1972).

Lastly, the Test of Nonverbal Intelligence, Second Edition (TONI-2), published in 1982-90 was examined, as it is an alternative method of assessment used with Hispanic or limited English proficient students. This test is designed to provide a language-free method of assessing general intelligence. It uses picture problem-solving items. Its culture-reduced measure is optimal for use with individuals with different linguistic backgrounds. Reliabilities exceed .80 for all ages (Brown, Sherbenou & Johnsen, 1982).

Testing of Language Minorities

The testing of students with limited or non-English knowledge has become an increasing concern among educators and professionals in assessment practices. Research in this area has been slow and has not been as successful as in other areas of assessment. The adaptation and implementation of valid testing practices with language minorities is a complicated matter. It is linked to the population validity of tests and to the non-biased test development and test administration practices (Duran and Linn ed. 1989).

Baca, Yawkey, Gonzalez and Vega (1997) remark that an important point in the psychometric characteristics of standardized tests is that any given instrument is normed only for the particular sample that participated in the norming process. In another linguistic cultural context, this process is meaningless and therefore any information derived from these tests are biased due to lack of validity and reliability. Suzuki, Meller, and Ponterotto (1996) state that when verbal or language-based tests may disadvantage the examinee, the sampling process is restricted (p. 249). Jones (1988) adds that there is a crisis in the psycho educational assessment of minority and language minority students. The public Law 94-142 requires that children be evaluated using instruments that are not racially or culturally biased.

Concerning non-biased test development, which is linked to valid testing practices, Jones (1988) commented that bias could exist at the content level where decisions are first made about what items to include in a test. The perspectives and experiences of minority group children are often excluded.

Regarding test administration, practices with minority students, Jones (1988) stresses that one of the criteria that evaluation of cultural minorities must meet is that persons who are familiar with the patterns of language, behavior and customs of the person being examined must administer tests in the child's native language.

An assessment of Hispanic limited English proficient students is difficult because of differences in language, culture and thought pattern. These factors, which influence test performance, will be examined. Studies addressing language proficiency and its effects on assessment implicate that linguistic barriers inhibit test performance. Although there are theoretical disagreements on how to best define language proficiency, research in this area suggests that limited familiarity with a language used in testing can affect test performance and the ability to make valid inferences about the meaning of this performance. Consequently, low proficiency in language can make test performance statistically unreliable (Duran and Linn 1989).

One of the areas that pose difficulties in testing students of non-English backgrounds is their cultural influence on thinking. Recent cross-cultural research suggests, There are intimate connections among the ways people perceive the nature of problem solving situations, problem solving tasks, the language surrounding tasks, and social cultural experiences (Duran 1989, p. 574).

Familiarity with the language of assessment might now be significant to ensure that individuals understand tasks. To perform tasks adequately, people have to first understand the social and cultural context of assessment situations the way language is used and the mode of thinking expected (Duran, 1989).

The ways of thinking associated with English-speaking students reflect social and learning experiences, which emphasized those modes of thought. Language minority students reflect a different social and cultural background from that of the mainstream American English speakers. The unrecognized differences in backgrounds of students might create inaccurate assumptions about their testing performance. Therefore, to understand how cultural and linguistic actors influence assessment is a challenge that specialists and educators must face (Kaiser, 1998).

Difficulties in Assessing Learning Disabilities with Hispanic/Bilingual Students

Tests play a central role in identifying learning disabilities with Hispanic bilingual students. Concern over the use of valid testing for these purposes among language minority children has been evident over the past twenty years. This concern was stimulated in the 1970 s as a result of the court case in the state of California (Diana v. State Board of Education) which claimed that too many Hispanic students were in EMR classes and because of misuse of intelligence test. (Duran, 1989).

Gonzalez (1997) states that the situation of assessing Hispanic leaning disabled students changed in the last twenty to thirty years and particularly in the areas of bilingual/multicultural education, and bilingual special education has made tremendous progress even though the challenge for proper diagnoses still exists.

An aid in identifying learning disabilities with Hispanic students is the translation of assessment measures from the English to the Spanish language. The WISC-R, the WISC III, updated and normed in the United States with Hispanic students, the WLPB-R, and the EIWN-R are instruments widely used to assess Hispanic students. Despite these translation efforts, there is still the challenge of having an assessment administered by skilled individuals who are fluent in the examinee's language and who can separate problems related to language and those associated with deficits in learning. (Suzuki et al. 1996).

Although translated tests serve of great benefit in the assessment process in identifying learning disabilities with Hispanic students, the need for more measures available in Spanish still exists. In reference to the discrepancy formula used to identify learning-disabled children whereby combinations of ability and achievement test that yield difference score reliabilities higher than .80, should be used when classifying children, (Shuttle and Borich, 1984) often times psychologist use the WISC-R in Spanish to determine ability levels and the achievement portion is determined by the WIAT which is available only in English. This discrepancy between ability level obtained in one language and achievement in another poses questions of validity of obtained score.

An important factor in assessing Hispanic students is for the practitioner to be aware of different expressions used by individuals who speak the same language but originate from different geographic regions. Failure to account for differences can inappropriately penalize individuals for using terms acceptable within their culture. (Suzuki et al., 1996).

An asset in assessing learning disabilities with limited English proficient students is the availability of alternative methods of assessment such as nonverbal intelligent tests,

which provide a language-free method of assessing general intelligence. These tests are useful when language makes it difficult to reliably administer standard test.

(Gonzalez, Vega, Yawkey, & Baca, 1997).

Another benefit that is brought into the testing process for identification of learning disabilities with Hispanic students is the administration of language dominance tests, which is mandated by the state of New Jersey. Since the challenge of validity of standardized test use with diverse ethnic groups, the state of New Jersey mandates that children who are limited English proficient (LEP) be tested for language dominance before they are administered a traditional intelligence test. (Jones, 1988). This mandate is fair and should be practiced by practitioners in the field as many by pass it making the assumption that students are dominant in the language they are casually conversing.

As Gonzalez (1997) stated, there has been great positive changes in assessment practices with LEP students due to legislation mandates in the past twenty years, however there is still a need for more reliable assessment instruments available in Spanish and normed with the population in question.

Summary

This review of literature examined the definition and importance of reliability, forms of reliability, the inter-rater type of reliability, which is relevant to the hypothesis of this research, and the generalizability theory. Reliability of an assessment is a must for accuracy of assessments to be achieved. It is an important aspect of professional accountability.

Reliability of scales and assessment instruments used with Hispanic limited English proficient students is weak and sometimes lacking. Most test examined were outdated

and in need of re-norming. They are still useful for assessment of language dominance, proficiency, intelligence and achievement because of their availability

The continued growth in the number of language minority students, including a large percentage of Hispanic students requiring linguistic and educational assessment has challenged test developers. The EIWN-R was translated into Spanish in 1982 and the WISC III was normed in the United States with Hispanic students. Precaution is needed in the administration and interpretation of standardized measures. These measures are required by state statutes with the intent to assist local schools in meeting the needs of diverse limited English proficient students. The development of norms based on appropriate groups of Spanish speaking children is encouraged. The need for reliable instruments useful with the Hispanic population is great as few highly reliable and valid tests are available for grade k-12. Tests normed with the population intended and those that are culturally unbiased would lessen the difficulties in identifying learning disabilities with Hispanic students.

There has been progress in the areas of fair testing practices and placement of students, yet there is much work in the future that still awaits to be accomplished in order to best serve the students who will become the future majority in our schools.

CHAPTER 3
DESIGN OF THE STUDY

Sample

The sample included 30 Hispanic male and female students enrolled in the Bilingual Program in a district of a city in Southern New Jersey. This city is one of the three urban areas in rural Cumberland County in Southern New Jersey. The total school population in this district consists of 5,848.00 students of which 3,618.00 approximately 62% are Caucasian; 1,398.00 approximately 24% are African American; and 749 approximately 13% of the total school population are Hispanic (O Donnel, 2000). The Hispanic population is predominantly Puerto Rican. There are 100 students enrolled in the Bilingual Program. The sampled population ranged from ages 6 to 10. These students were randomly selected. Their class grades ranged from first grade to fifth grade. Among the male students, 7 were first graders, 1 was a second grader, 1 was a third grader, 2 were fourth graders, and 3 were fifth graders.

N=30 SAMPLE

	Male	Female
Grade 1	7	5
Grade 2	1	2
Grade 3	1	4
Grade 4	2	2
Grade 5	3	3
Total	14	16

Table 3.1

Measures

The measure that was used with this study is the Psychological Processing Checklist (PPC), which is currently undergoing validity and reliability studies. It was developed by the Department of Psychology of the Illinois State University. Upon

completion of these studies, this scale is intended to be used to assess in students the cognitive processes necessary to differentiate learning disabilities from underachievement and other disabling conditions. The Psychological processing Checklist (PPC) is a 35-item paper and pencil checklist, which is to be completed by a student's teachers. The students are rated on each of the 35 items using the rating scale of (0) Never; (1) Seldom, (2) Sometimes, (3) Often. The estimated time for the completion of the checklist is 20 minutes. Scoring is based on adding the frequency ratings for the items grouped by processing areas. The items in the checklist are grouped into six areas including Auditory, Visual Processing, Visual-Motor, Social Perception, Organization and Attention. The items included in the six areas are scattered through the checklist. This checklist includes behavioral characteristics that are observable and measurable. Examples of items included in the rating scale are: Has difficulty seeing similarities in pictures, letters numbers, words, and groups of objects. Has difficulty immediately recalling information presented orally.

Design of Study

Two different teachers rated each individual student: one bilingual and/or ESL, the other a mainstream regular classroom teacher. Both used the Psychological Processing Checklist to rate each student.

Testable hypothesis

Null: There will be low correlation in the rating of bilingual students in a city in Southern New Jersey with the Psychological Processing Checklist (PPC) by a bilingual and/or ESL teacher and mainstream regular classroom teacher.

Alternate Hypothesis

There will be a high degree of correlation in the rating of bilingual students in a district in Southern New Jersey with the Psychological Processing Checklist (PPC) by a bilingual and/or ESL teacher and mainstream regular classroom teacher.

Analysis

The correlational study was used in this research. This study was analyzed by the used of raw scores on a rating scale (PPC). The raw scores of items rated will be correlated to determine a reliability coefficient.

Summary

A sample of thirty Hispanic students enrolled in the Bilingual Program in a district in Southern New Jersey were rated by two different teachers who used the Psychological Processing Checklist as the rating measure. The resulting correlations will be found in Chapter Four and the implications for use in determining students processing deficits and in developing interventions to facilitate students learning is discussed in Chapter Five.

CHAPTER IV

ANALYSIS OF RESULTS

Interpretation of Results

The purpose of this research was to conduct an interrater reliability study of the Psychological Processing Checklist (PPC) with Hispanic bilingual students in a school district in South Jersey.

The sample consisted of 14 boys and 16 girls (See Table 3.1) enrolled in a Bilingual Program. The population ranged in ages 6 to 10. These students were randomly selected. Their class grades ranged from first grade to fifth grade. Among the male students, 7 were first graders, 1 was a second grader, 1 was a third grader, 2 were fourth graders, and 3 were fifth graders. Among the female students, 5 were first graders, 2 were second graders, 4 were third graders, 2 were fourth graders, and 3 were fifth graders. (See Table 3.1).

The hypothesis of this study states that there will be a high degree of correlation in the rating of Hispanic bilingual students with the Psychological Processing Checklist(PPC) by a bilingual and/or ESL teacher and a mainstream regular classroom teacher.

The correlational study was used in this research. This study was analyzed by the raw scores on the PPC rating scale. The raw scores of items rated were correlated to determine a reliability coefficient.

The results of this study indicated that the null hypothesis was rejected and the alternate hypothesis was accepted as there is a high correlation in the rating of Hispanic students with the Psychological Processing Checklist by a bilingual and/or ESL teacher

and/or ESL teacher and a mainstream regular classroom teacher as shown on Interrater Reliability of Total Score Figure 4.1.

The Psychological Processing Checklist, developed by the Department of Psychology of the Illinois State University is currently undergoing reliability study. This 35-item checklist is intended to help with the assessment of learning disabilities among children. The PPC, which is completed by the child's teachers is designed for use in kindergarten through fifth grade. It assesses a range of basic psychological processes, including auditory processing, visual processing, visual-motor skills, social perception, organization and attention.

According to information-processing theories, learning disabled students are deficient in basic psychological processing. The auditory process is a stage where inner speech occurs. This process is essential for normal academic learning since it deals with language. With linguistically disadvantaged students, this process is vital in assessing distinction of deficit from language interference. Likewise, theory states that attention is crucial to learning, especially in acquiring a new language. It involves storing and retrieving information processed.

The areas of auditory processing and attention were correlated as shown on Figure 4.1 and Figure 4.2. The interrater reliability in both attention and auditory processing demonstrates a very high degree of correlation.

Summary

Based on results of this study, the PPC is proven to be a useful instrument with the Hispanic bilingual population as well as with mainstream regular classroom students since the interrater reliability coefficient is very high. Reliability is an essential

characteristic of an instrument. It is undertaken to provide information, which will be used to evaluate and make decisions about children. Estimating the reliability of an instrument is one way of knowing how much confidence can be placed in the results obtained from a particular instrument. The Psychological Processing Checklist would appear to be very reliable instrument to be used with Hispanic bilingual students.

Figure 4.1

Interrater Reliability of Total Score

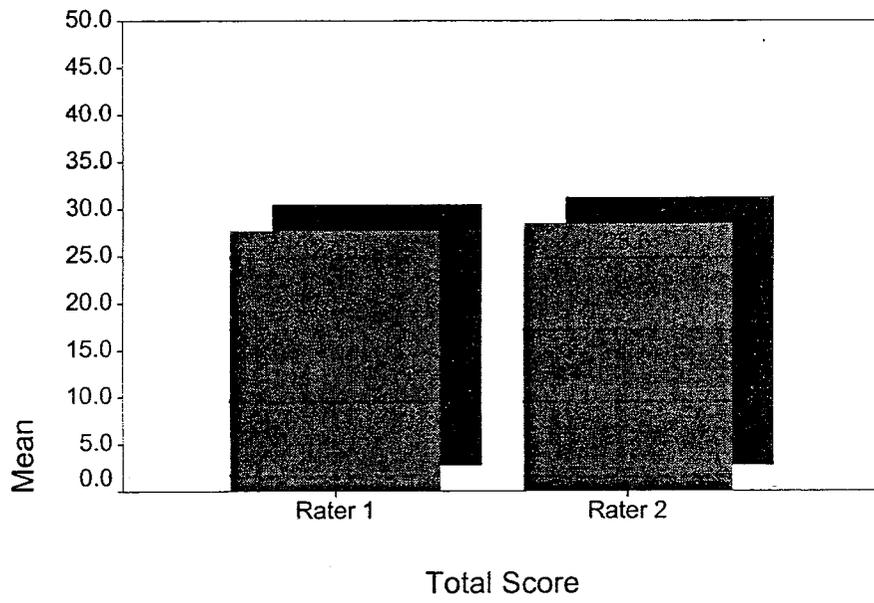


Figure 4.2

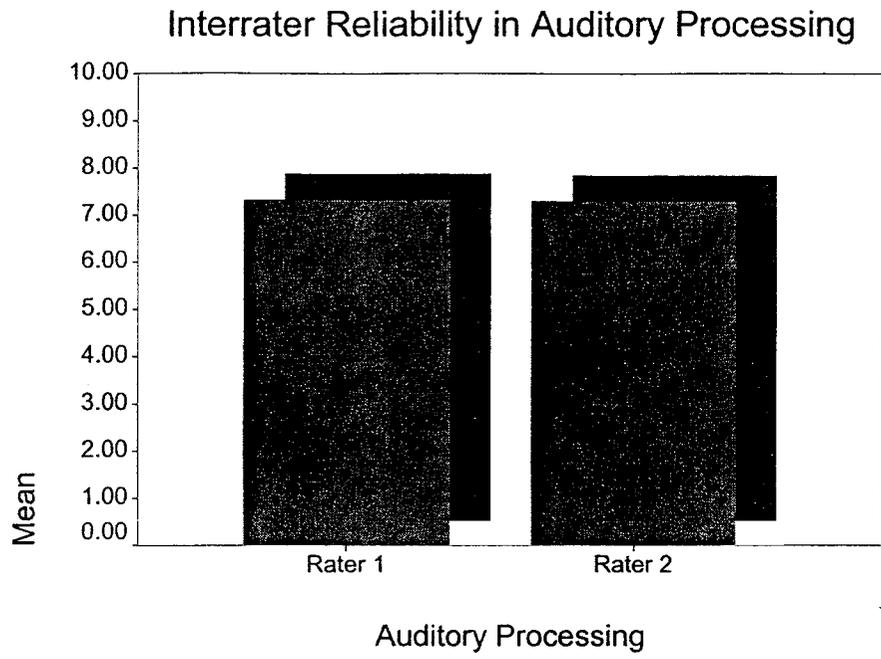
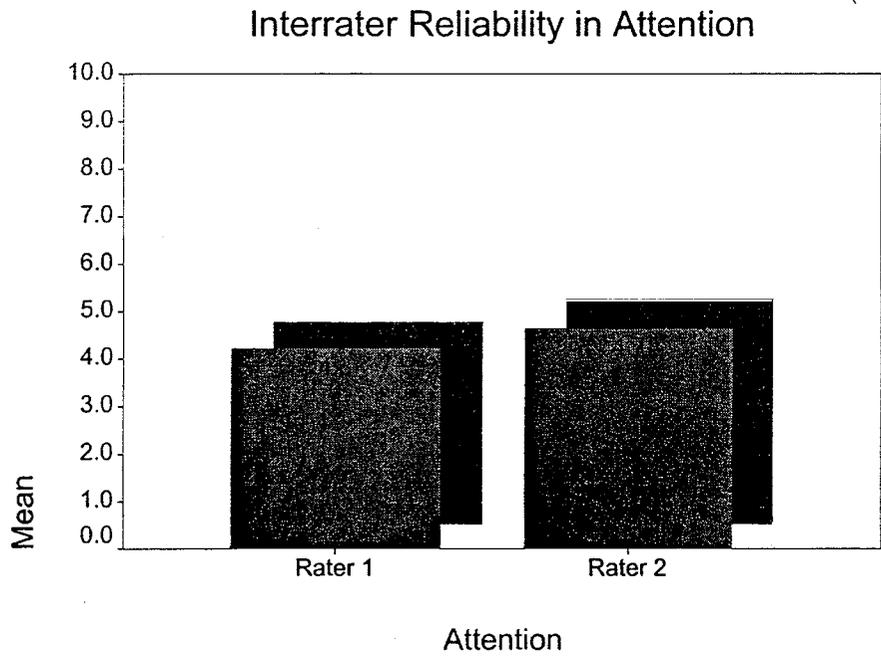


Figure 4.3



CHAPTER V

SUMMARY AND CONCLUSIONS

Summary

The purpose of this research is to conduct an interrater reliability study of the Psychological Processing Checklist(PPC) with Hispanic bilingual students in a district in Southern New Jersey. There is a compelling need for developing an instrument that would be valid and reliable for use with Hispanic students including those who are limited English proficient.

It was hypothesized that there would be a high degree of correlation in the rating of Hispanic bilingual students in a district of Southern New Jersey with the Psychological Processing Checklist(PPC) by a bilingual and/or ESL teacher and a mainstream classroom teacher.

Discussion

The design of the project was a correlational study to determine the interrater reliability of The PPC. The population consisted of thirty bilingual students, male and female ages ranging from 6 to 10 in grades 1 to 5. Two teachers rated each student: a bilingual and/or ESL and a mainstream classroom teacher. Analysis of the results supports the hypothesis for there is a high correlation ($r = .988 < .000$) between raters.

As stated in Chapter II, review of literature pertinent to the topic of this research found that many assessment instruments used with language disadvantaged students produce evidence of weak reliability. With the increased number of students in grade k-12 in our schools whose native language is not English creates a need for educational professionals to make linguistic, intellectual and academic assessments of these students. As part of

these assessments, standardized instruments are likely to be used. Reliability of an instrument is a must for accuracy of assessment to be achieved. It is an important aspect of professional accountability. Estimating the reliability of an instrument is one way of knowing how much confidence can be placed in the results obtained from a particular instrument. This is especially important when evaluations and decisions are made about children.

It was impressive to note that both a bilingual and/or ESL teacher and a mainstream classroom teacher rated students with the same degree of congruency as demonstrated by the Total Score of Interrater Reliability figure. Also, it was noted, that although a rater rated a few students low on some processing areas whereas another rated the same student differently, the composite score of processing areas shows a high degree of correlation.

Conclusion

Results of this research indicated that the Psychological Processing Checklist (PPC) would appear to be a useful and reliable instrument to be used in the assessment of cognitive deficits with Hispanic bilingual students. This checklist is free of cultural bias and students attitude is not a factor in determining outcomes.

Implications for further study

Although the Psychological Processing Checklist (PPC) appears to be a reliable instrument which can be used with bilingual students, there is still a need for the further development of psychological instruments to be used with Hispanic bilingual students including those who are limited English proficient, especially with the significant increase of number of students in our schools whose native language is not English.

Also, a similar study could be conducted with students in middle school and high school.

REFERENCES

- Abbott, Muriel. (1982). Language Assessment Battery. New York: Office of Testing.
- Buros, O. (1985). Mental Measures Year Book (9th). Pictorial Test f Bilingualism El Circo.
- Buros, O. (1991-1994). Mental Measures Year Book (13th). Language Assessment Scales.
- Buros, O. (1997). Mental Measures (13th). Spanish Assessment Battery (SABE).
- Buros, O. (1997). Mental Measures (13th). Supera. Spanish ed.: Terra Nova.
- Buros, O. (1972). Mental Measures (9th). The Bel-Sil Spanish Test.
- Buros, O. (1982-1990). Mental Measures (12th). Test of Non-Verbal Intelligence. 2nd ed.: TONI —2.
- Buros, O. (1991). Mental Measures (12th). Woodcock-Johnson Language Proficiency Battery.
- Campione, J., & Brown, A. (1986). Psychological Theories and the Study of Learning Disabilities. American Psychologist, 14 (10), 1059-1068. Handbook.
- Duncan S. & De Avila, E. LAS (Language Assessment Scales). (1990). Monterey: CTB.
- Goldstein, G. & Hersen, M. (Eds.). (1990). Test Construction. Handbook of Psychological Assessment. (2nd ed., pp.29-32). New York: Pergamon Press.
- Gonzalez, V., Brusca-Vega, R., Yawkey, T. and Baca, L. (1997). Assessment and Instruction of Culturally and Linguistically Diverse Students with or At-Risk of Learning Problems. Boston: Allyn and Bacon.
- James, P. (1974). James Language Dominance Test. Texas: Learning Concepts.
- Jones, R. (1988). Psychoeducational Assessment of Minority Group Children: A Casebook. California: Cobb & Henry Publishers.
- Kantowitz, B. H. (1994) . Information-Processing Theory. (Vol. Ed.), Encyclopedia of Psychology: Vol.2. (2nd ed., pp. 245-248). New York: Wiley.

Kavale, K. A., Forness, S. R., & Bender, M. (1987). Social Problems, Comprehension Problems, and Reading Disabilities. Handbook of Learning Disabilities: Vol. 1. Dimensions and Diagnosis (pp.158-223). Boston: College-Hill Press.

Linn, R. & Duran, R. (Eds.). (1989). Testing of Linguistic Minorities. Educational Measurement (3rd ed., pp. 573-583). New York: McMillan Publishing Company.

O Donnel, Catherine (2000). Director of Public Relations. Millville Public Schools. New Jersey.

Groth-Marnat, G. (1999). Evaluating Psychological Tests. Handbook of Psychological Assessment. (2nd ed., pp. 10-14). New York: Wiley.

Groth-Marnat, G. (1997). Patterns of Test Usage in Clinical Assessment. Handbook of Psychological Assessment. (3rd ed., pp.12-16). New York: Wiley.

Gruijter, D. & VanderKamp, L (1990). Generalizability Theory. Advances in Educational and Psychological Testing: Theory and Applications. Boston: Kluwer Academic Publishers.

Hoyt, W. & Melby, J. (1999). Dependability of Measurement in Counseling Psychology: An Introduction to Generalizability Theory. The Counseling Psychologist, 27 (3), 325-352.

Kaplan, R. & Saccuzzo, D. (1982). Psychological Testing: Principles, Applications, and Issues. California: Brooks/Cole Publishing Company.

Kayser, Hortensia (1998). Assessment and Intervention Resource for Hispanic Children. California: Singular Publishing Group, Inc.

Linn, R. (Ed.) (1989). Testing of Linguistic Minorities. Educational Measurement: (3rd ed., pp. 573-583). New York: McMillan Publishing Company.

Maculaitis, J. (1982). Technical Manual for Maculaitis Assessment Program: A Multi-Purpose Test for Non-Native Speakers of English in Kindergarten Through Twelfth Grade. New York: Alemany Press.

Schulte, A., & Borich G. (1984). Considerations in the Use of Difference Scores to Identify Learning-Disabled Children. Journal of Psychology, 22, 381-390.

Stanley, J. (1972). Psychological and Cultural Factors That Influence Performance on Cognitive Tests. Educational and Psychological Measurement and Evaluation. New Jersey: Prentice-Hall Inc.

Stanley, J.C. (1972). Test Reliability. Educational and Psychological Measurement and Evaluation. (pp.114-128). New Jersey: Prentice-Hall.

Suzuki, L., Meller, P. & Ponterrotto, J. (Eds.). (1996). Handbook of Multicultural Assessment. California: Jossey-Bass Publishers.

Swanson, L. (1987). Information-Processing Theory and Learning Disabilities: A commentary and Future Perspective. Journal of Learning Disabilities, 20, 155-163. New York: Wiley.

Travis, E. (1971) Handbook of Speech Pathology and Audiology. New York: Appleton-Century Crofts.

Wechsler, D. (1974). Wechsler Intelligent Scale for Children-Revised. Texas: Harcourt Brace Jonanovich, Inc.

APPENDIX A

Psychological Processing Checklist (Grades K-5)



PPC (GRADES K-5)

by Mark E. Swerdlik, Peggy Swerdlik, Tim Thomas, & Jeffrey H. Kahn

Instructions: Please rate the student on each item using the following scale:

- 0 = Never; Does not engage in the behavior**
- 1 = Seldom; Exhibits behavior one to several times per month**
- 2 = Sometimes; Exhibits behavior one to several times per week**
- 3 = Often; Exhibits behavior one to several times per day**

	0	1	2	3
1. Requires more than 5 seconds to respond to known questions presented orally	⊙	⊙	⊙	⊙
2. Has difficulty seeing similarities in pictures, letters, numbers, words, and groups of objects	⊙	⊙	⊙	⊙
3. Paper-and-pencil tasks are messy (e.g., scratches out rather than erases)	⊙	⊙	⊙	⊙
4. Disturbs others with inappropriate physical contacts (e.g., pokes peers)	⊙	⊙	⊙	⊙
5. Has difficulty managing time (e.g., late for school)	⊙	⊙	⊙	⊙
6. Has difficulty paying attention during group instruction	⊙	⊙	⊙	⊙
7. Has difficulty seeing differences in pictures, letters, numbers, words, and groups of objects	⊙	⊙	⊙	⊙
8. Has difficulty with hand/eye coordination tasks (e.g., cutting with scissors, catching a ball)	⊙	⊙	⊙	⊙
9. Has difficulty understanding nonverbal feedback (e.g., a scowl)	⊙	⊙	⊙	⊙
10. Has difficulty locating appropriate materials (e.g., loses complicated assignments)	⊙	⊙	⊙	⊙
11. Has difficulty completing independent seatwork or homework	⊙	⊙	⊙	⊙
12. Has difficulty immediately recalling information presented orally	⊙	⊙	⊙	⊙
13. Has difficulty following multi-step directions presented visually	⊙	⊙	⊙	⊙
14. Has difficulty forming letters when printing or writing	⊙	⊙	⊙	⊙
15. Disturbs others with inappropriate noises (e.g., mouth sounds)	⊙	⊙	⊙	⊙
16. Has difficulty following classroom routines (e.g., lunchroom lineup)	⊙	⊙	⊙	⊙
17. Has difficulty changing from one task/subject to the next task/subject	⊙	⊙	⊙	⊙
18. Has difficulty with spacing between letters, words, sentences, or numbers	⊙	⊙	⊙	⊙
19. Has difficulty telling or writing a logical story	⊙	⊙	⊙	⊙
20. Has difficulty ignoring environmental distractions	⊙	⊙	⊙	⊙
21. Has difficulty remembering past information presented orally	⊙	⊙	⊙	⊙
22. Has difficulty following directions presented visually (e.g., demonstrations)	⊙	⊙	⊙	⊙
23. Has difficulty with coloring and/or painting (e.g., staying within the lines when coloring)	⊙	⊙	⊙	⊙
24. Has difficulty following multi-step oral directions	⊙	⊙	⊙	⊙
25. Confuses left from right when presented with visual materials (e.g., reading, math computation)	⊙	⊙	⊙	⊙
26. Has difficulty copying from the board or a book	⊙	⊙	⊙	⊙
27. Repeats oral directions, but does not follow them	⊙	⊙	⊙	⊙
28. Disturbs others with repetitive motor movements (e.g., rocking in chair)	⊙	⊙	⊙	⊙
29. Has difficulty planning and placing a written product on a page (e.g., forgets heading)	⊙	⊙	⊙	⊙
30. Has difficulty following a conversation	⊙	⊙	⊙	⊙
31. Has difficulty applying conversational rules (e.g., interrupts conversations, does not wait for his or her turn)	⊙	⊙	⊙	⊙
32. Has difficulty listening to stories, unless they are accompanied by pictures	⊙	⊙	⊙	⊙
33. Has difficulty recognizing the same word when repeated in a sentence or passage	⊙	⊙	⊙	⊙
34. Asks for oral questions and directions to be repeated	⊙	⊙	⊙	⊙
35. Has difficulty noticing visual changes in the environment (e.g., a new bulletin board)	⊙	⊙	⊙	⊙



APPENDIX B

**Psychological Processing Checklist (Grades K-5)
Data Summary Section**

Psychological Processing Checklist K-5 (4)

DATA SUMMARY SECTION

- 1 -

SUBSCALE

RAW SCORE

SECTION I

AUDITORY
PROCESSING

1. _____
12. _____
21. _____
24. _____
27. _____
32. _____
34. _____

SECTION II

VISUAL
PROCESSING

2. _____
7. _____
13. _____
22. _____
25. _____
33. _____
35. _____

SECTION III

VISUAL MOTOR

3. _____
8. _____
14. _____
18. _____
23. _____
26. _____

Psychological Processing Checklist K-5 (5)

DATA SUMMARY SECTION

- 2 -

SUBSCALE		RAW SCORE
<hr/>		
SECTION IV		
SOCIAL PERCEPTION	4.	_____
	9.	_____
	15.	_____
	28.	_____
	31.	_____
<hr/>		
SECTION V		
ORGANIZATION	5.	_____
	10.	_____
	16.	_____
	19.	_____
	29.	_____
<hr/>		
SECTION VI		
ATTENTION	6.	_____
	11.	_____
	17.	_____
	20.	_____
	30.	_____
<hr/>		