

INTERRATER RELIABILITY OF THE BALCONES SENSORY
INTEGRATION SCREENING TEST-REVISED

A THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF MASTER OF ARTS
IN THE GRADUATE SCHOOL OF THE
TEXAS WOMAN'S UNIVERSITY

SCHOOL OF OCCUPATIONAL THERAPY

BY

MARY ANN MONKHOUSE KLEUSER, B.S., OTR

DENTON, TEXAS

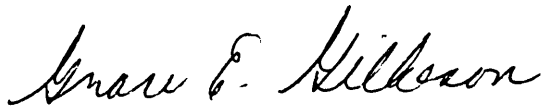
AUGUST 1986

TEXAS WOMAN'S UNIVERSITY
DENTON, TEXAS

May 28, 1986

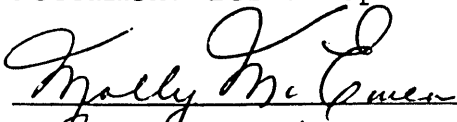
To the Provost of the Graduate School:

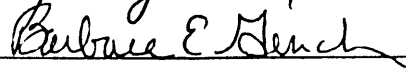
I am submitting herewith a Thesis written by Mary Ann Monkhouse entitled "Interrater Reliability of the Balcones Sensory Integration Screening Test-Revised". I have examined the final copy of this Thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Arts, with a major in Occupational Therapy.



Grace E. Gilkeson, Major Professor

We have read this thesis and
recommend its acceptance:





Accepted



Provost of the Graduate School

ACKNOWLEDGEMENTS

The author is indebted to her committee for their constructive suggestions and encouragement, her colleagues for their willingness to participate in the study, and her family for their support and patience during this seemingly unending endeavor.

INTERRATER RELIABILITY OF THE BALCONES SENSORY
INTEGRATION SCREENING TEST-REVISED

By

Mary Ann Monkhouse Kleuser, B.S., OTR

August 1986

The Balcones Sensory Integration Screening Test-Revised (BSIST-R) was designed to identify children who need additional sensory integrative testing. As with many evaluations utilized by occupational therapists, validity and reliability studies had not been done. This study addressed interrater reliability of the BSIST-R.

Three groups of occupational therapists, with 10 in each group, served as subjects. The groups were delimited by the subjects' methods of learning test administration and by their experience with the test. The therapists completed questionnaires regarding work experience, viewed a videotape of a child being tested with the BSIST-R, and scored the child's performance. These scores comprised the data and were compared both within the groups and among the three groups. The hypotheses stated that there were no significant differences either within or among the three groups.

Generalizability theory (Berk, 1979) was utilized to compute coefficients of reliability within each of the three groups. The generalizability coefficients for 10

observations in each of the three groups were:

Group I--.919; Group II--.981; Group III--.966. These coefficients indicated a high degree of consistency within the three groups. To test for differences among the three groups, an overall multivariate F -test was computed for all groups. It was not significant ($F = 1.57$; $p = .122$). An ANOVA and univariate F -test indicated no difference between the three groups on eight factors identified in the original standardization of the test. Alpha (.05) was readjusted to .006 to protect against a Type I error when interpreting the results of the eight univariate tests. The hypothesis of no significant difference among the scores given by the three groups was supported on both the multivariate F -test ($\alpha = .05$) and the eight univariate F -tests ($\alpha = .006$).

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.	iii
LIST OF TABLES.	vi
Chapter	
I. INTRODUCTION.	1
Statement of the Problem.	1
Statement of the Purpose.	2
Significance of the Study	2
Hypotheses.	4
Assumptions and Limitations	4
Definition of Terms	5
II. REVIEW OF LITERATURE.	7
Evaluation in Occupational Therapy Practice	7
Reliability and Validity.	15
Generalizability Theory	19
Balcones Sensory Integration Screening Test-Revised.	23
III. METHODOLOGY	26
Subjects.	26
Instrument.	27
Procedure	28
Statistical Analysis.	29
IV. PRESENTATION OF FINDINGS.	31
Description of the Subjects	31
Statistical Analysis.	32
Discussion.	39
V. SUMMARY, CONCLUSION AND RECOMMENDATIONS	44
Summary	44
Conclusion.	45
Recommendations	45

REFERENCES.	49
---------------------	----

APPENDICES

A. Subject Questionnaire and Consent Forms.	55
B. Subject Scoring Sheet.	58
C. Visual Motor Forms Test.	60
D. Letter to Participating Therapists	62
E. Balcones Sensory Integration Screening Test-Revised Protocol Manual	65

LIST OF TABLES

Table

1	Experience and SCSIT* Test Training for Subject Groups.	33
2	Work Settings for Subject Groups.	34
3	Frequency of Use of BSIST-R by Subject Groups.	35
4	Analysis of Variance Components for Estimating Interobserver Reliability Across Eight Behaviors.	36
5	Generalizability Coefficients for Subject Groups.	37
6	Univariate F-tests on Each of Eight Factors (Behaviors)	39

CHAPTER I

INTRODUCTION

Adequate and proper assessment for occupational therapy clients provides the "keystone of treatment planning" (Bowker, 1984, p. 25). Bowker (1984) stated that the occupational therapy assessment influences choices of intervention strategies. In addition, in the present atmosphere of fiscal accountability for health care, the occupational therapy assessment may determine the choice of patients who will actually benefit from intervention. For many years occupational therapists have lacked assessment tools specific to their discipline, but now more evaluations designed by therapists are being used (Benson & Clark, 1982; Bowker, 1984; Hopkins & Tiffany, 1983). Therapists both designing and utilizing these tests must use accepted standards of testing science to critically view the quality and usefulness of these assessment tools (Berk & DeGangi, 1979; Hasselkus & Safrit, 1976; Lewko, 1976; Punwar, 1976).

Statement of the Problem

The problem addressed by this study was the absence of established interrater reliability for the Balcones Sensory Integration Screening Test-Revised (Jones & Monkhouse, 1981).

Statement of the Purpose

The purpose of this study was to determine the adequacy of the current Balcones Sensory Integration Screening Test-Revised (BSIST-R) protocol manual through the study of interrater reliability within and among the following three groups: (a) occupational therapists who attended a course in administration of the Balcones Sensory Integration Screening Test-Revised (BSIST-R) and had subsequent clinical experience in its use including at least three administrations within the last year; (b) occupational therapists who had learned administration of the BSIST-R only through independent study of the protocol manual with subsequent clinical experience in its use including at least three administrations within the last year; and (c) occupational therapists who reviewed the BSIST-R protocol manual but had no clinical experience in its use. This comparison was done to determine if the protocol manual alone was sufficient to achieve scoring reliability or if experience and/or a test administration course significantly influenced scoring reliability.

Significance of the Study

The significance of the study was determination of the quality of the Balcones Sensory Integration Screening

Test-Revised (BSIST-R) in terms of its interrater reliability. An acceptable level of interrater reliability improves the test's value as a tool for screening children for sensory integrative dysfunction and for making appropriate clinical judgments regarding the need for additional testing. Also, if scorers tend to give consistent scores on the BSIST-R using only the protocol manual or the manual and clinical experience with the test, then formal training in the test administration would not be necessary. In addition, if interrater reliability is determined to be good, due in part to the adequacy of the protocol manual, then norms for the BSIST-R could be expanded and additional research applications of the test could be considered.

Hypotheses

The following hypotheses were tested at the .05 level of significance.

1. There is no significant difference in the scores given by the therapists within the group who have had both a course and clinical experience in the administration of the Balcones Sensory Integration Screening Test-Revised (BSIST-R).

2. There is no significant difference in the scores given by the therapists within the group who have had only independent study and clinical experience with the BSIST-R.

3. There is no significant difference in the scores given by therapists within the group who have only reviewed the protocol manual but had no experience with the BSIST-R.

4. There is no significant difference among the scores on the BSIST-R given by the three groups of therapists.

Assumptions and Limitations

The following assumptions were identified in completion of this study:

1. Only occupational therapists with a minimum of one year of pediatric clinical experience were used as scorers.

2. Scorers had experience in administering other standardized tests to children.

3. Therapists experienced in administration of the Balcones Sensory Integration Screening Test-Revised (BSIST-R) had administered the test at least 10 times with 3 administrations within the last year.

4. Therapists identified as having taken the training course in administration were enrolled in one of the three courses given in the fall of 1981 sponsored by the Texas Occupational Therapy Association, Inc.

5. Scorers reviewed the protocol manual and gave scores as accurate as possible when observing the child's responses during the test administration.

A limitation of the study was the necessity for mailing the videotape to subjects instead of personally monitoring all viewings of the videotape.

Definition of Terms

The terms used in this study have been defined as follows:

Interrater agreement--the degree to which scores or observations are replicable and consistent across observers (Berk, 1979; Berk & DeGangi, 1979; Greenstein, 1980; Hasselkus & Safrit, 1976; American Psychological Association [APA], 1974). According to Berk (1979), the usual methods for computing interrater reliability are actually measures of interrater agreement based on two-observer studies with categorical data using percentages of observer agreement or interclass correlations.

Interrater reliability--although related to interrater agreement, reliability reflects "the degree of objectivity with which the target behavior can be measured" (Berk, 1979, p. 460). Regardless of who administers the scale to an individual, the observed performance will be the same. "Data should be gathered on the consistency with which two or more

independent therapists observe the set of behaviors measured by the scale" (Berk & DeGangi, 1979, p. 243).

Reliability--"the various types of consistency meaningful to the measurement process" (Berk & DeGangi, 1979, p. 243).

Screening instrument--a standardized test which determines the need for further evaluation in specific areas. Ideally, it separates individuals with the possibility of having a particular trait from a large group of individuals who probably do not have that trait (Hopkins & Tiffany, 1983).

Standardized test--an evaluation method which is administered in a precise and predetermined manner and whose results can be compared to normative data derived from controlled studies of the test (APA, 1974).

CHAPTER II

REVIEW OF LITERATURE

In the review of literature, the topics which are discussed include assessment in the practice of occupational therapy, studies investigating occupational therapists' use and misuse of evaluations in assessment, and the need for reliable and valid evaluation tools and their proper application in the occupational therapy assessment process. Reliability and validity in test design and development are examined, as well as generalizability theory as a method for determining interrater reliability. The Balcones Sensory Integration Screening Test-Revised, an instrument with clinical value which does not have established interrater reliability, is presented as the focus for this study.

Evaluation in Occupational Therapy Practice

The methods and importance of evaluation in occupational therapy has become increasingly sophisticated and important. Gillette (1971) stated, "the evaluation process represents a commitment to a professional responsibility" (p. 79). Evaluation was described as providing the basis from which the treatment objectives and process would evolve and to

which the therapist and patient would return to determine the degree and quality of their success. Gillette felt that evaluation was of critical importance not only to the individual therapist/patient relationship but also to the profession as a whole.

It becomes apparent that one's evaluation procedures measure not only patient performance and ability, but the therapist's own success and even, to some degree, the inherent worth of the occupational therapy process itself...(p. 80)

Additionally, Gillette felt that if therapists want to assume the privileges and rights of being professionals, that they must also assume the responsibility and initiative to develop and utilize effective and appropriate evaluation methods and tools (1971, 1982).

Smith and Tiffany (Hopkins & Tiffany, 1983) differentiated between evaluation and assessment. In their view, assessment was "the sum of the results of the evaluation procedures used" (p. 143). They also stated that it is the process of assessment from which meaningful intervention objectives can be developed. Accurate data collection, observation, and the need to utilize several sources in the assessment process was emphasized by both Gillette (1971) and Hopkins and Tiffany (1983).

This process of data collection can and should include many sources and methods such as observation, reviewing other tests, checklists, rating scales, inventories and standardized tests (Bowker, 1984; Gillette, 1971; Greenstein, 1980; Hopkins & Tiffany, 1983; Lewko, 1976; Punwar, 1976). Typically, therapists have utilized many of these sources of information as a part of their evaluation and repertoire. Standardized testing has been a part of evaluation, but typically therapists have been creative in their modification of existing tools or in the development of new ones to fill the needs of their client population (Bowker, 1984; Greenstein, 1980; Hopkins & Tiffany, 1983; Lewko, 1976; Punwar, 1976). Lewko (1976) sampled facilities providing services to children. In that sample 256 different published and unpublished tests were being used. His study was across professional categories but occupational therapists represented 27% of the respondents. Lewko identified significant problems with the use of 'standardized' tests. First, it was found that poor information or disregard for the limitations of the tests existed. Second, a large number of respondents used unpublished 'self-made' tests. In responding to Lewko's article, Punwar (1976) concurred with findings as a "reasonably accurate description of current practice in evaluating motor behaviors" (p. 420). Punwar confirmed the use of clinician

devised tests stating that occupational therapists are "prone to the 'adaptation syndrome'" (p. 421) because they do not feel any other evaluation adequately meets the needs of their unique client/patient population. It may be true that existing standardized tests do not meet the needs of therapists for patient evaluation (Banus, 1983), thus the proliferation of 'therapist devised' evaluations. This proliferation may also be a result of therapists' poor understanding of test and measurement theory and requirements for validity and reliability.

A notable exception to the prevalence of poorly standardized tests in the area of occupational therapy evaluation has been the test battery, the Southern California Sensory Integration Tests (SCSIT), developed by A. Jean Ayres, OTR, Ph.D. Dr. Ayres has been involved in the evaluation and treatment of sensory integrative dysfunction for over 30 years (Henderson, Llorens, Gilfoyle, Myers, & Prevel, 1974). Ayres developed the series of tests to detect and to determine the nature of sensory integrative dysfunction in children who have learning and behavior disorders. Originally, parts of the SCSIT were published separately as the Ayre's Space Test, Southern California Motor Accuracy Test, Southern California Figure-Ground Visual Perception Test, Southern California Perceptual Motor Test, and Southern California Kinesthesia and Tactile

Perception Tests. When these tests were combined in 1972 as the SCSIT, the Position in Space Test and Design Copying were added and some other revisions were made. Extensive standardization studies were done "in several different groups of tests, each group administered at a different time to different but similar sample populations of children" (Ayres, 1972, p. 1). Norms were developed for children from age 4 through 8 for most of the subtests but through 10 years on the Design Copying and Motor Accuracy-Revised. The normative sample was 70-125 children (approximately equal numbers of male and female) in each six month interval from 4.0 to 8.11 years. The test manual revised in 1980, includes 83 tables of standardization and normative data. Statistics of reliability and validity are also provided and discussed (Ayres, 1980).

The theory of sensory integration has been fraught with controversy, as has the SCSIT. The test battery has been challenged by many validity and reliability studies (Ayres & Mailloux, 1981; Cermak & Ayres, 1984; Henderson, Llorens, Gilfoyle, Myers & Prevel, 1974; Hsu & Nelson, 1981; Kimball, 1981; Montgomery & Rodell, 1982; Morrison & Sublett, 1983; Nelson, Weidensaul, Shih & Anderson, 1984; Peterson, Goar, Van Deusen, 1985; Peterson & Wikoff, 1983; Punwar, 1982; Royeen, Lesinski, Ciani, & Schneider, 1981; Saeki, Clark & Azen, 1985; Smith, 1983; Ziviani, Poulsen, & O'Brien, 1982).

The Center for the Study of Sensory Integrative Dysfunction (CSSID) was established by Dr. Ayres and others to further the work and study of sensory integration. In 1983, the name of the Center was changed to that of Sensory Integration International (SII). The CSSID and SII have held theory and test administration courses for the purpose of 'certifying' therapists in the administration and interpretation of the SCSIT. The intent of the courses and certifying procedures has been to educate therapists in proper administration procedures, scoring and interpretation of the rather sophisticated and complicated battery of tests. Approximately 1,500 professionals hold SCSIT certification. Most of these individuals are occupational therapists (Center for the Study of Sensory Integrative Dysfunction, 1983). The SCSIT is currently undergoing extensive revision and standardization procedures. The revised instrument will be republished as the Sensory Integration and Praxis Test (Ayres, Mailloux, & Mcatee, 1985).

Unlike the SCSIT, many of the evaluations developed by therapists for particular clinical settings, may have face validity and have been thoughtfully developed, but these assessments have not been subjected to adequate standardization, validity, and reliability studies (Bonder, 1985; Bowker, 1984; Greenstein, 1980; Hasselkus & Safrit,

1976; Punwar, 1976). The procedures to establish validity and reliability are considerable. Articles in the professional journals of occupational therapy in the past few years have explained the process of developing a standardized test and the principles of statistical measurement including the concepts of validity and reliability (Benson & Clark, 1982; Berk & DeGangi, 1979; Greenstein, 1980; Hasselkus & Safrit, 1976; Kielhofner, 1982a; Kielhofner, 1982b). As Hasselkus and Safrit (1976) stated:

Measurement is not simply assigning numbers to behaviors or objects, it is the assigning of numbers according to the rules. It is not enough for a therapist to design a test or rating scale on the basis of common sense and logic. The principles of measurement theory must be followed...(p. 429)

Increasingly, these journals have had articles citing validity and reliability studies for existing and/or developing standardized evaluations (Banus, 1983; Behnke & Fetkovich, 1984; Bledsoe & Shepherd, 1982; Clopton & Martin, 1984; Crowe, Deitz, & Siegner, 1984; Dutton, 1985; Fox & Harlowe, 1984; Gregory-Flock & Yerxa, 1984; Izraelevitz, Fisher, & Bundy, 1985; Kielhofner & Nelson, 1983; Mathiowetz, Weber, Kashman, & Volland, 1985; Ottenbacher, Dauck, Grahn,

Gevelinger, & Hassett, 1985; Royeen & Kannegieter, 1984; Sparling & Rogers, 1985; Wilson, Barber, & Craddock, 1985; Yerxa, Barber, Diaz, Black, & Azen, 1983).

The need for valid and reliable standardized evaluations which are administered to the appropriate populations according to the protocol manuals and instructions continues (Benson & Clark, 1982; Bonder, 1985; Bowker, 1984; Hopkins & Smith, 1983). Bowker (1984) stated that historically, therapists have worked on a one-to-one basis with patients. Rarely have they used systematized evaluations and in fact few therapists have been initially well trained in utilizing standardized tests. This has changed somewhat in recent years. In 1979, a survey done by the special interest sections of the American Occupational Therapy Association showed a "growing trend for therapists to document clinical findings through the use of some structured instruments that are valid and reliable" (Bowker, 1984, p. 29). Research has become part of many undergraduate occupational therapy curricula (American Occupational Therapy Association [AOTA], 1983). In 1984, The Representative Assembly of the American Occupational Therapy Association adopted the "Hierarchy of Competencies Relating to the Use of Standardized Instruments and Evaluation Techniques by Occupational Therapists" as an official document of the association (AOTA, 1984). The

increasing demands on the profession to justify treatment methods and objectively document client change magnify the need for more evaluations properly standardized and utilized in occupational therapy practice. Third party payers of occupational therapy services including government, public agencies such as schools, and private insurance companies may begin to demand that treatment decisions be based on more precise evaluation data as well as clinical experience (Benson & Clark, 1982; Bonder, 1985; Bowker, 1984; Christiansen, 1983; Gillette, 1982).

Reliability and Validity

The validity and reliability of an evaluation tool can be expressed as statistics which are the "quantitative measures of qualitative characteristics" (Berk & DeGangi, 1979, p. 241) of the evaluation in question. The validity of a scale is a measure of how well the evaluation achieves the purpose for which it was constructed. Reliability refers to the consistency of an evaluation in yielding the same data over time and regardless of the rater. "Classical methods of estimating reliability coefficients call for correlating at least two sets of similar measurements" (APA, 1974, p. 48).

Berk and DeGangi (1979) cited types of validity and reliability which they felt should be stated in the

protocol manual of an evaluation instrument. They described discriminant validity, domain validity, and construct validity. Discriminant validity concentrates on how the scale is used and the extent to which the evaluation will discriminate between groups of patients with different diagnoses or dysfunction. Domain validity describes the extent to which the items on the evaluation are representative of all aspects of the 'domain' being evaluated. This aspect of an evaluation is based on judgment rather than a statistical computation. Construct validity is developed over time but begins with the investigator formulating hypotheses about the 'construct' that the evaluation is believed to be measuring. Research that empirically supports the construct of the evaluation is accumulated over time. This research can include predictions about behavior or results of the evaluation that are confirmed or disconfirmed (APA, 1974).

Reliability was divided into two categories by Berk and DeGangi (1979). The dependability of the decisions made based on evaluation results was called decision-making reliability. This type of reliability answers the question, if a patient is put into one treatment program or group on the basis of evaluation scores, would he be put into that same group again on a parallel version of the test.

The second was interrater reliability. They made two points in discussing this type of reliability. First, because motor scales are scored by an observer (rater), it is important that observations are replicable and consistent from one rater to another. "The reliability index reflects the degree to which the different therapists observe the same behaviors" (Berk & DeGangi, 1979; p. 243). Secondly, they distinguished between interrater reliability and interrater agreement stating that indexes of reliability were more "rigorous, precise, and flexible" (p. 243) and hence of more value than those indexes measuring only agreement. Berk (1979) went on to explain that:

The standard psychometric definition of reliability in terms of observed-score and true-score variance components is neither assumed nor considered in the measurement of interobserver [interrater] reliability in the above research domains [APA, 1974]. The term agreement provides a more appropriate label for the statistics that appear in the literature. (p. 460)

Berk (1979) has advocated the use of generalizability theory to determine true rater reliability over more traditional methods of interclass correlation or percentage of rater agreement. It is interrater reliability computed utilizing generalizability theory that is the focus of this study.

According to the manual Standards for Educational and Psychological Tests published by the American Psychological Association (1974), developers of standardized tests have an obligation to determine the validity and reliability of their evaluations or testing instruments. This manual rated certain aspects of a test as 'essential,' 'very desirable' or 'desirable' to be provided. The manual listed predictive and concurrent validity, content validity, and construct validity as essential to be provided. Evidence of reliability including estimates of the standard error of measurement, procedures and sample groups utilized to determine reliability and test-retest reliability were deemed essential by the American Psychological Association (1974). Interrater reliability was only described as "desirable" in the American Psychological Association's manual (p. 50).

Berk and DeGangi (1979) appeared to think that interrater reliability was more important in the development of a motor scale because of the fact that a subject's performance is observed and scored by a therapist. "Ideally, regardless of who administers the scale to an individual, the observed performance should be virtually identical" (Berk & DeGangi, 1979, p. 243). They stated that an interrater reliability could be very high or consistent, it could also be consistently inaccurate. They advocated that

the validity of the observations should be checked as well but this aspect of test standardization was beyond the scope of this study.

Generalizability Theory

The use of the theory of generalizability to determine the interrater reliability in psychological and educational measurement was described in detail by Berk (1979). In this paper, he outlined and critiqued 16 indices of interrater agreement and six methods for estimating coefficients of interrater reliability.

In discussing interrater agreement, he stated that utilizing the percentage of rater agreement as an indicator of interrater reliability yields "measures that are often spuriously high...due, in part, to their failure to take into account the proportion of agreement due to chance" (Berk, 1979, p. 461). Additionally, the agreement that is reported is usually a study utilizing two raters with categorical data even when the larger study utilizes more than two raters. Berk (1979) stated that this "tends to magnify the distortion of each pairwise statistic" (p. 461). To overcome these problems of agreement, researchers have devised methods for estimating coefficients of interrater reliability. These methods overcome the problem of agreement due to chance but remain limited by the use of

categorical data. The kappa coefficient is considered one of the more "rigorous" (Berk, 1979, p. 462) of these methods and is capable of handling two or more raters but is still not considered appropriate when handling data on mentally/physically handicapped individuals (Hartmann, 1977; Liberty, 1976). Berk's (1979) criticism of the correlations are mainly in the realms of flexibility and computational ease.

The correlations cannot account or control for the numerous sources of extraneous variance that could confound the estimation and interpretation of reliability. Furthermore, the average intercorrelation cannot be calculated efficiently without a computer (p. 462).

Cardinet, Tourneur, and Allal (1976) discuss generalizability from a historical perspective stating that the definition of reliability assumes that repeated use of a test or measurement will be under equivalent conditions to the original standardization. This causes an inherent problem in that there will always be sources of variation in every administration. Generalizability theory according to these authors, "admits...that each observation belongs to a multitude of possible sets of observations" (Cardinet et al., 1976, p. 121).

The interclass correlation-generalizability theory approach expresses the classical theory of measurement error relationship between true and observed variance and is computed from analysis of variance components. Its difference lies in the fact that instead of focusing on the F-ratios, the "major interest is the magnitude of the individual variance components and generalizability coefficients" (p. 463). The coefficients are based on the true-to-observed variance ratio and provide estimates of the extent to which the observed ratings are confounded with error. A sample of observations can be generalized to the universe of observations. In addition, according to Berk (1979), generalizability allows specific identification of variables such as characteristics of the rater, client, and investigator, methods of scoring behavior, and the nature and duration of rater training as sources of error affecting the measurement of target behavior.

Two types of generalizability coefficients are of value to the investigator of interrater reliability. These are: (a) generalizability of a single observations and (b) generalizability of the average of k observations. According to Berk (1979, p. 465-466), single observations are expressed in the following ratio:

$$\rho_1^2 = \frac{\sigma_\rho^2}{\sigma_\rho^2 + \sigma_0^2 + \sigma_e^2}$$

σ_ρ^2 - (numerator) the variance of the true observations of behavior in the universe of which n persons in a study constitute a random sample.

σ_ρ^2 - variance of the true observations

σ_0^2 - variance due to rater bias

σ_e^2 - variance due to the errors of measurement.

These variance component estimates are derived from the mean squares. The sources of variation for determining mean squares include between behaviors, within behaviors, between raters, and residual. Use of these estimates into the formula results in the generalizability coefficient $\Sigma\rho_1^2$ for a single observation.

Substituting the value of $\underline{k}$ in the between-persons variance component (σ_ρ^2) provides different generalizability coefficients for different numbers of raters.

$\underline{k}$ observations are expressed with the following ratio:

$$\rho_2^2 = \frac{\kappa \sigma_\rho^2}{\kappa \sigma_\rho^2 + \sigma_0^2 + \sigma_e^2}$$

$\Sigma\rho_2^2$ - the estimate of ρ_2^2 , which is computed from the same variance components as the coefficient $\Sigma\rho_1^2$ for the single observation.

Balcones Sensory Integration Screening Test-Revised

As stated above, many assessments utilized by occupational therapists do not have adequate validity and reliability studies to determine their value and applicability in the clinical setting. One such assessment is the Balcones Sensory Integration Screening Test-Revised (Jones & Monkhouse, 1981). Published by the Texas Occupational Therapy Association, Inc., the Balcones Sensory Integration Screening Test-Revised (BSIST-R) was developed from a screening protocol and manual which was a result of work done in the Balcones Special Services Cooperative (BSSC) under a Title IV-C, ESEA Grant (1976-1979). The screening protocol was used to screen all students who received special education services from the BSSC. The purpose of developing the screening protocol was to enable therapists and special educators to quickly identify children who might benefit from further sensory integration testing and intervention. Although the Southern California Sensory Integration Tests (SCSIT) are well known and utilized by occupational therapists, the amount of time required to administer and to analyze the results is extensive (Ayres, 1976; 1980). There was no available screening tool that determined which children would need such extensive evaluation. This made the SCSIT not practical in many settings most notably the schools where limited numbers of

therapists are expected to provide services to large numbers of children (Gilfoyle, 1980).

According to the protocol manual of the BSIST-R (Jones & Monkhouse, 1981), the following standardization studies were done. Test items on the BSIST-R were devised from test items in previously published and standardized instruments. Each item was drawn from at least two other standardized tests. In 1979, the test was administered to two classes each of first, second, and third graders. The age range of the children tested was 6 to 9 years. The total sample population was 130 children from two public schools and the racial distribution of the sample was proportional to that of the population of Texas. Factor analysis identified eight distinct factors for which means and standard deviations were computed. The specific formulae and/or methods used to compute these statistics were not provided in the manual.

In 1981, the test was published as the Balcones Sensory Integration Screening Test-Revised by the Texas Occupational Therapy Association. At that time, additional explanatory text and illustrations were added, some titles were changed to reflect test items more accurately, and the format was changed to that of a spiral bound book in order to facilitate administration of the test. The published protocol manual became part of a test kit which contained

all the materials necessary to administer the test with the exception of a hand dynamometer. In addition, three workshops were given in the fall of 1981 to clarify further the test's use to clinicians. Since that time, approximately 300 test kits have been sold through the Texas Occupational Therapy Association. The need for such a screening test seems clear; however, the lack of adequate validity and reliability studies needs to be rectified. It was the lack of interrater reliability that served as the focus of this study.

CHAPTER III

METHODOLOGY

An interrater reliability study was done with scores given to a child tested with the Balcones Sensory Integration Screening Test-Revised (BSIST-R). A videotape of the testing was made and then viewed by 30 occupational therapists. These therapists scored the child's performance on the test. An analysis of variance was computed to determine interrater reliability.

Subjects

The subjects were three groups of occupational therapists. Each therapist had experience of at least one year working with pediatric clients. Ten occupational therapists were in each group. One group had taken a course in administration of the BSIST-R and had subsequent clinical experience in administering the test. The second group had utilized independent study of the protocol manual in learning the test administration and had subsequent clinical experience in administering the BSIST-R. Experience was defined as having had at least 10 administrations of the BSIST-R with three test administrations within the last year.

The third group had a chance to read and review the protocol manual before scoring but did not have any previous clinical experience with the BSIST-R.

The subjects were each asked to sign a consent form and to complete a questionnaire (Appendix A). The questionnaire was developed to determine years of experience in occupational therapy practice and pediatric practice, work setting, subject group, and other experience with sensory integration theory, evaluation, and treatment. The questionnaire and scoring sheet (Appendix B) were assigned subject numbers and were separated from the consent forms. One subject's scoring sheet was not used due to her acknowledgement that she had not read the protocol manual at all prior to viewing the videotape.

Instrument

The investigator tested one child using the BSIST-R. The session was videotaped. The child was 9.5 years of age and had mild to moderate sensory integrative dysfunction as determined by histories taken from parent and therapist. The videotape was viewed by the investigator and a thesis committee member to determine if the child's responses were depicted clearly enough to be adequately scored. Copies of the child's written performance on the Visual Motor Forms

item were duplicated for use by the scoring therapists rather than having them score a page viewed on the videotape (Appendix C).

Procedure

Scorers were to view the videotape at their convenience but under the supervision of the investigator. Although initial viewings were done in this manner, it became logistically impossible to complete the study in this way because of the lack of local therapists who met the subject criteria. A detailed letter was written describing the step by step procedure for viewing the tape and scoring the child's performance (Appendix D). The instructional letter, videotape, protocol manuals (Appendix E), scoring sheets, subject questionnaires and consent forms, and all other materials needed to complete scoring were sent to therapists who met the criteria and who agreed to participate in the study.

Groups and individual therapists viewed the tape but independently scored the child's performance. The importance of not discussing the tape during the viewing was emphasized in the instruction letter. Sufficient time was allowed between each item of the test to allow the scorer to reread the test items and scoring criteria if necessary. The scorer was allowed to see each item several times if

necessary. Instructions on how to do this were included in the letter and on the videotape. The scoring therapist determined when the next item would be shown on the tape.

Statistical Analysis

Statistical analysis of scoring sheets and questionnaires was done utilizing descriptive statistics, correlation coefficients, multivariate analysis of variance, and interclass correlation-generalizability theory. The subjects were divided into the three test groups: (a) Group I--Balcones Course; (b) Group II--Independent Study, and (c) Group III--Manual Review. Descriptive statistics for each group included a frequency distribution, mean, median, mode, range, standard error, variance, and standard deviation for each item on the questionnaire and each of the 29 items scored on the BSIST-R.

A multivariate analysis of variance was computed for each of the eight factors (behaviors) to determine if differences existed among the three groups. These factors were: (a) ocular; (b) proprioception; (c) primitive reflexes; (d) vestibular; (e) tactile; (f) stereognosis; (g) form; and (h) laterality. Univariate F-ratios were computed for each of these eight factors and an overall multivariate F-ratio was computed. The overall multivariate $\underline{F}$ was not significant, so to protect against a Type 1 error, alpha was

adjusted to .006. A Type 1 error is that of "rejecting a 'true' null hypothesis when it should have been supported" (Ottenbacher, 1984, p. 38). The computation for the Type 1 error protection is .05/8 tests which yields the more stringent $\alpha = .006$ (Glass & Hopkins, 1984).

Two generalizability (interclass) coefficients of interrater reliability were computed utilizing two-way analysis of variance components. Sources of variance included between behaviors, within behaviors, between observers and residual. The coefficients were computed based on a single observation for each group and on k (10) observations for each group.

CHAPTER IV

PRESENTATION OF FINDINGS

The findings are presented as description of the subjects and as statistical analysis of the data. The final section is a discussion of the findings.

Description of the Subjects

Descriptive statistics were compared for each of the three groups. Mean years of occupational therapy experience in the groups were: Group I--12 years, Group II--6 years and Group III--8 years. Stated as percentages, 70% of Group I had 6 or more years of occupational therapy experience compared with 30% for Group I and 50% for Group III. Pediatric occupational therapy experience of 6 or more years was held by 80% of the subjects in Group I, 20% in Group II, and 50% in Group III. Group I had more overall experience in the practice of occupational therapy. Although Group I did have more experience than the other groups, all subjects met the experience requirements specified for the study.

The questionnaire included questions as to whether subjects had taken theory or test administration courses or were certified to administer the SCSIT. Group I had 6 or 60%

of the therapists who were certified, Group II had 2 (20%), and Group III had no certified therapists. Percentages for attendance at theory and administration courses for the SCSIT are summarized in Table 1. Data on therapy experience and certification is to be found in Table 1 also.

Table 2 contains information on subjects' work settings. Groups I and III show a majority of therapists working in a public school setting whereas Group III had a majority in outpatient rehabilitation settings. Other work settings included hospitals and inpatient rehabilitation facilities.

Table 3 contains responses to questions regarding use of the Balcones Sensory Integration Screening Test-Revised (BSIST-R). There are very similar responses by Groups I and II. Group III did not respond to any of these items on the questionnaire because of the identified study limitations. This group was to have had no prior experience with the BSIST-R.

Statistical Analysis

Generalizability coefficients were computed to test hypotheses 1, 2, and 3 of interrater reliability, or consistency of scoring, within each of the three groups. The analysis of variance (ANOVA) components for estimation of interrater reliability for Groups I, II, and III are in Table 4. The behaviors from which the sources of variation

Table 1

Experience and SCSIT* Test Training for Subject Groups

	Group I	Group II	Group III
	Balcones Course	Independent Study	Manual Review
	n = 10	n = 10	n = 10
Therapy Experience			
Range (years)	20	18	19
Low to High (years)	5-25	2-20	1-20
Mean	12	6	7.9
6 or more years (%)	70%	30%	50%
Pediatric Experience			
Range (years)	13	13	11
Low to High (years)	2-15	2-15	1-12
Mean	8.6	5.2	5.7
6 or more years (%)	80%	20%	50%
SCSIT* Training			
Theory Course	7	5	7
Test Administration	6	3	5
Certified	6	2	0

*SCSIT--Southern California Sensory Integration Tests

Table 2

Work Settings for Subject Groups

	Group I	Group II	Group III	Total
	Balcones Course	Independent Study	Manual Review	%
	n = 10	n = 10	n = 10	
Public School	5	1	6	40%
Private Practice	2	1	0	10%
Outpatient Rehab Facility	3	6	4	43%
Other Work Settings	0	2	0	7%

Table 3

Frequency of Use of BSIST-R by Subject Groups

	Group I	Group II
	Balcones Course	Independent Study
	n = 10	n = 10
Current Use of BSIST	9	9
BSIST-R as main form of evaluation	2	1
Number of BSIST-R		
Administered /month		
0-2	6	7
3-4	2	2
5-8	0	0
9-10	2	1
/year		
0-3	2	1
4-6	2	4
7-10	2	1
11-20	1	2
21-80	3	2

Table 4

Analysis of Variance Components for Estimating Interrater
Reliability Across Eight Behaviors

Source of Variation	Sum of Squares (SS)	df	Mean Square (MS)	Estimated Variance Component $\hat{\sigma}^2$
<u>Group I--Balcones Course</u>				
Between Behaviors	489.79	7	69.97	6.43
Within Behaviors	379.70	72	5.27	5.27
Between Observers	22.86	9	2.54	-.39
Residual	356.84	63	5.66	5.66
<u>Group II--Independent Study</u>				
Between Behaviors	675.15	7	96.45	9.48
Within Behaviors	130.80	72	1.82	1.82
Between Observers	25.95	9	2.88	.15
Residual	104.85	63	1.66	1.66
<u>Group III--Manual Review</u>				
Between Behaviors	564.79	7	80.68	7.80
Within Behaviors	198.10	72	2.75	2.75
Between Observers	27.26	9	3.03	.04
Residual	170.84	63	2.71	2.71

were computed are the eight factors identified in the original standardization.

The ANOVA components from Table 4 were then used to compute generalizability coefficients. Interrater reliability stated as generalizability coefficients for single observations and the average of 10 observations for each of the three groups are presented in Table 5.

Table 5

Generalizability Coefficients for Subject Groups

	Group I	Group II	Group III
	Balcones Course	Independent Study	Manual Review
Single Observation			
$\Sigma \rho^2_1 =$	.532	.840	.739
k (10) Observations			
$\Sigma \rho^2_2 =$	.919	.981	.966

The lowest generalizability coefficient was a single observation in Group I (Balcones course) at .532 indicating a lack of reliability based on a single observation. When 10 observations were utilized in the computation, the coefficient rose to an acceptable .919 indicating a high degree of agreement in Group I scoring. Group II (Independent Study) had the highest coefficients for both

single and 10 observations at .840 and .981, respectively. These coefficients indicate a high degree of agreement and dependability in a single observation and homogeneity within the group of observations. Group III (Manual Review) achieved coefficients of .739 for individual and .966 for 10 observations.

The test of Hypothesis 4 of no significant difference among the scores given by the three groups was done utilizing a multivariate ANOVA. The overall multivariate F -test computed for all groups was not significant ($F = 1.57$; $p = .122$). This indicates no overall difference in the scores given by all the subjects. The ANOVA and univariate F -tests computed for the three groups on the eight factors indicated no differences among the three groups on the eight factors (behaviors). Since the overall multivariate F -test was not significant ($\alpha = .05$), there was a need to readjust α to protect against a Type 1 error when interpreting the results of the eight univariate tests. Hypothesis 4 of no significant difference among the scores given on the BSIST-R given by the three groups is supported by the lack of significance on both the multivariate F -test ($\alpha = .05$) and the eight univariate F -tests ($\alpha = .006$). Table 6 shows the results of the univariate F -tests.

Table 6

Univariate F-tests on Each of Eight Factors (Behaviors)

Factor	SS	SS error	MS	MS error	F	p
Ocular	6.87	65.0	3.43	2.41	1.43	.258
Proprioception	3.47	111.9	1.73	4.14	0.42	.662
Primitive Reflexes	0.87	85.0	0.43	3.15	0.14	.872
Vestibular	0.20	21.3	0.10	0.79	0.13	.882
Tactile	16.47	64.5	8.23	2.39	3.45	.046
Stereognosis	4.27	31.6	2.13	1.17	1.82	.181
Form	11.27	56.6	5.63	2.10	2.69	.086
Laterality	0.47	13.0	0.23	0.48	0.48	.621

degrees of freedom: 2, 27

Discussion

The results of this study showed no significant differences either within the test groups or among the test groups in the scoring of a child's performance on the BSIST-R. The groups were defined by their experience in the administration of the BSIST-R. The study indicates that there is no difference in the scoring consistency or interrater reliability regardless of whether a scorer took a course in test administration, learned the test through independent study, or read the manual prior to scoring the

child's performance. The protocol manual alone appears to be sufficient to achieve scoring consistency. Formal training is not necessary and perhaps is undesirable.

Group I, by definition, was composed of therapists who had taken a course in administration of the BSIST-R given in the fall of 1981. This group had the lowest correlation coefficients of all the groups for both the single observation (.532) and 10 observations (.919). Although the single observation coefficient is not adequate by testing standards (DeGangi, Berk, & Larsen, 1980; Yerxa, 1982), the 10 observation coefficient is considered adequate.

Group I seemed to be the best prepared in terms of years of experience, amount of formal preparation, and number of SCSIT certifications held. The lower correlation coefficients could be because of the scorers being too sophisticated in their test interpretation skills. Rather than using the protocol manual at 'face value,' they over-interpreted the child's performance and complicated the scoring procedure. This may be a function of the extensive training required to gain certification for the SCSIT. This group had 60% of its subjects who were certified.

Questions about the actual course in administration of the BSIST-R should be raised. The course intended to clarify test procedures and scoring may have only caused

more discrepancies and confusion among course participants. Course information may have been inconsistent with the protocol manual.

Group II was composed of therapists who had learned to administer the BSIST-R through independent study of the protocol manual and experience. This group had the highest correlation coefficients of single observations at .840 and 10 observations at .981. This group did not have the most experience in pediatric therapy nor did they utilize the BSIST-R more. The majority worked in outpatient rehabilitation facilities rather than the schools and only 20% of the group was SCSIT certified. The reason for their high consistency in scoring the test may be because of the therapists not having the sophisticated observation skills which develop with the use of complicated standardized evaluations such as the SCSIT. Therefore, they were more dependent on objectively utilizing the protocol manual.

Group III who only read the manual prior to viewing the videotape had the second highest correlation coefficients of .739 for a single observation and .966 for 10 observations. It is this group's results that most strongly supports the adequacy of using only the protocol manual to achieve scoring consistency or interrater reliability.

There are some factors which may have influenced this overall study. The most critical one is the use of a

videotaped testing session to score the child's performance on the BSIST-R. The logistics and number of subjects precluded carrying out the study in any other way. The subjects' abilities to see the child's reactions on some of the test items may have been less than ideal. In taping the child's performance, camera angles may have accentuated aspects of a child's performance while obscuring others. It was hoped that the viewing of the tape by a committee member helped to minimize these limitations to the study.

Another factor would be the subject's ability to view the child's performance 'several times if necessary.' In an actual testing situation, this would probably not occur. In the six viewings monitored by the investigator, no subject requested to view the tape more than a second time. It is not known if subjects viewing the tape without being monitored took advantage of additional viewings of the same test item. A subject viewing the same test item numerous times and taking an inordinate amount of time to score the test would effect the results. Exactly what this effect might be is difficult to determine. This is a factor for which there was no control in the study.

The original plan for the investigator to monitor the viewing of the videotape by all the subjects had to be altered due to the difficulty in locating subjects who fit the criteria of the groups within the immediate geographical

area. The first six viewings of the tape were done as planned. Using these initial monitored viewings as a guide, a letter was written to accompany the tape as it was mailed to therapists throughout the state (Appendix D). In spite of very specific instructions, discussion of the videotape by subjects during viewing could not be eliminated and may be a factor. However, groups of therapists who viewed the tape together were always representative of at least two of the groups in the study rather than just one.

CHAPTER V

SUMMARY, CONCLUSION AND RECOMMENDATIONS

The summary briefly discusses the purpose and the method of this study. The conclusion summarizes the important findings of the study. The recommendations include suggestions for revisions to the Balcones Sensory Integration Screening Test-Revised and for further validity and reliability studies.

Summary

The purpose of this study was to determine the adequacy of the protocol manual of the Balcones Sensory Integration Test-Revised (BSIST-R) in achieving interrater reliability within and among three groups of ten occupational therapists. The three groups represented: (a) therapists who had taken a course in administration of the BSIST-R and had subsequent clinical experience; (b) therapists who learned BSIST-R administration through independent study and subsequent clinical experience; (c) therapists who had no prior experience with the BSIST-R but had read the protocol manual prior to scoring a child's performance on the test. A videotape of a child being tested with the BSIST-R was

made. The occupational therapists representing the three groups viewed the videotape and scored the child's performance. Interrater reliability using analysis of variance based on a behaviors x scorers (raters) data matrix (Berk, 1979) and a multivariate analysis of variance was computed.

Conclusion

No significant differences were found either within the groups or among the three groups. The data demonstrated that the BSIST-R protocol manual alone is adequate to achieve interrater reliability of .90 or above. A course in test administration and clinical experience had no influence on improving interrater reliability. Actually, the group with the course on test administration had the lowest correlation coefficient although it was still at an acceptable level.

Recommendations

Although it has been demonstrated that the BSIST-R has interrater reliability at an acceptable level of .90 or above, validity and other types of reliability studies are still lacking. Consistency in the scoring may be established but the accuracy of what is being scored must be determined. Discriminant validity studies should be done to

determine how accurate the test is in discriminating between children who need additional evaluation and those who do not. Related to discriminant validity is decision-making reliability. This determines how consistent the test is in discriminating between children. Domain and construct validity have begun to be established in the initial standardization but additional studies should be conducted.

The BSIST-R was originally intended to be used by other professionals as well as occupational therapists. Interrater reliability studies with other groups such as teachers, educational diagnosticians and physical therapists should be done. This would help to determine if the interrater reliability established in this study is a function of the profession (occupational therapy) in the groups studied or the objectivity and clarity of the protocol manual.

The final recommendation for further studies on the BSIST-R involve the format of the protocol manual. Initial standardization included a factor analysis of the test items. Although this information is of value in evaluating the construct validity of the test, its inclusion in the protocol manual can be misleading. The eight factors are listed on a profile sheet included in the back of the protocol manual (see Appendix E). On this profile sheet, the child's scores are totaled and a standard score is tabulated for each factor. These standard scores are

inadequate because in several of the factors, a score lower or higher than -2.0 or +2.0 standard deviations from the mean cannot be achieved. In addition, the profile developed as a result of these standard scores implies that the test is more than a screening test. As a screening test, it should lead to additional testing but in some situations, it is the main form of evaluation utilized. On the questionnaire, subjects were asked if they used the BSIST-R as a main form of evaluation. The responses indicate that 10% of the subjects do use the test in this manner (see Table 3). It is recommended that the scoring procedure should be changed from this profile format to one in which a single score is achieved. This single score should then be compared to a table or range of scores. This comparison would determine whether or not the child's score is at a level indicating possible dysfunction and warranting further evaluation. It may be possible to achieve this redesign of the score tabulation procedure utilizing the original standardization data. However, another standardization study which incorporates recommended studies for validity and reliability would be a more desirable approach.

According to Benson and Clark (1982), the determination of interrater reliability is an important and appropriate step in the development of the BSIST-R. Additional studies as outlined are the next step which should be taken if the

BSIST-R is to be regarded as an adequately developed evaluation to be used by occupational therapists.

REFERENCES

- American Occupational Therapy Association, Inc. (1983). Essentials and Guidelines of an Accredited Educational Program for the Occupational Therapist, Rockville, MD.
- American Occupational Therapy Association, Inc. (1984). Hierarchy of competencies relating to the use of standardized instruments and evaluation techniques by occupational therapists. American Journal of Occupational Therapy, 38, 803-804.
- American Psychological Association (1974). Standards for Educational and Psychological Tests. Washington, DC: Author.
- Ayres, A. J. (1972). Southern California Sensory Integration Test-Manual. Los Angeles: Western Psychological Services.
- Ayres, A. J. (1976). Interpreting the Southern California Sensory Integration Tests. Los Angeles: Western Psychological Services.
- Ayres, A. J. (1980). Southern California Sensory Integration Test-Manual Revised (1980). Los Angeles: Western Psychological Services.
- Ayres, A. J., & Mailloux, Z. (1981). Influence of sensory integration procedures on language development. American Journal of Occupational Therapy, 35, 383-390.
- Ayres, A. J., Mailloux, Z., & Mcatee, S. (1985, September). An update of the sensory integration and praxis tests. Sensory Integration Special Interest Section Newsletter, p. 1.
- Banus, B. (1983). The Miller Assessment for Preschoolers (MAP): An introduction and review. American Journal of Occupational Therapy, 37, 33-340.
- Bartko, J., & Carpenter, W. (1976). On the methods and theory of reliability. The Journal of Nervous and Mental Diseases, 163, 307-317.

- Behnke, C., & Fetkovich, M. (1984). Examining the reliability and validity of the play history. American Journal of Occupational Therapy, 38, 94-100.
- Benson, J., & Clark, F. (1982). A guide for instrument development and validation. American Journal of Occupational Therapy, 36, 789-800.
- Berk, R. (1979). Generalizability of behavioral observations: A clarification of interobserver agreement and interobserver reliability. American Journal of Mental Deficiency, 83, 460-472.
- Berk, R., & DeGangi, G. (1979). Technical considerations in the evaluation of pediatric motor scales. American Journal of Occupational Therapy, 33, 240-244.
- Bledsoe, N., & Shepherd, J. (1982). A study of reliability and validity of a preschool play scale. American Journal of Occupational Therapy, 36, 783-788.
- Bonder, B. (1985). American Occupational Therapy Foundation-Standardized Assessments: Ethical principles for use. American Journal of Occupational Therapy, 39, 473-475.
- Bowker, A. (1984). Assessment: The keystone of treatment planning. Occupational Therapy in Health Care, 1, 25-32.
- Cardinet, J., Tourneur, Y., & Allal, L. (1976). The symmetry of generalizability theory: Applications to educational measurement. Journal of Educational Measurement, 13, 119-135.
- Center for the Study of Sensory Integrative Dysfunction (1983). Directory of CSSID Membership Certified in Administration and Interpretation of the SCSIT. Los Angeles, CA: Author.
- Cermak, J., & Ayres, A. J. (1984). Crossing the body midline in learning-disabled and normal children. American Journal of Occupational Therapy, 38, 35-39.
- Clopton, N., & Martin, A. (1984). A criticism of interrater reliability procedures for the Brazelton Neonatal Behavioral Assessment Scale. Physical and Occupational Therapy in Pediatrics, 4, 55-65.

- Crowe, T., Deitz, J., & Siegner, C. (1984). Postrotary nystagmus response of normal four-year-old children. Physical and Occupational Therapy in Pediatrics, 4, 19-28.
- DeGangi, G., Berk, R., & Larsen, L. (1980). The measurement of vestibular based functions in preschool children. American Journal of Occupational Therapy, 34, 452-459.
- Dutton, R. (1985). Reliability and clinical significance of the Southern California Postrotary Nystagmus Test. Physical and Occupational Therapy in Pediatrics, 5, 57-64.
- Fox, J., & Harlowe, D. (1984). Construct validation of occupational therapy measures used in CVA evaluation: A beginning. American Journal of Occupational Therapy, 38, 101-106.
- Gilfoyle, E. (Ed.). (1980). Training Occupational Therapists Educational Management in Schools. Module 3. Rockville, MD: American Occupational Therapy Association, Inc.
- Gillette, N. (1971). Occupational therapy and mental health. In H. S. Willard & C. S. Spackman (Eds.), Occupational therapy (4th ed.) (pp. 51-131). Philadelphia: Lippincott.
- Gillette, N. (1982). A data base for occupational therapy: Documentation through research. American Journal of Occupational Therapy, 36, 499-501.
- Glass, G., & Hopkins, K. (1984). Statistical methods in education and psychology (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.
- Greenstein, L. (1980). Teaching research: An introduction to statistical concepts and research terminology. American Journal of Occupational Therapy, 34, 320-327.
- Gregory-Flock, J., & Yerxa, E. (1984). Standardization of the prone extension postural test on children ages 4 through 8. American Journal of Occupational Therapy, 38, 187-194.
- Hartmann, D. (1977). Considerations in the choice of interobserver reliability estimates. Journal of Applied Behavior Analysis, 10, 103-116.

- Hasselkus, B., & Safrit, M. (1976). Measurement in occupational therapy. American Journal of Occupational Therapy, 30, 429-436.
- Henderson, A., Llorens, L., Gilfoyle, E., Myers, C., & Prevel, S. (Eds.). (1974). The development of sensory integrative theory and practice--A collection of the works of A. Jean Ayres. Dubuque: Kendall/Hunt.
- Hopkins, H., & Tiffany, E. (1983). Occupational therapy--A problem solving process. In H. L. Hopkins & H. D. Smith (Eds.), Willard and Spackman's occupational therapy (6th ed.) (pp. 89-99). Philadelphia: Lippincott.
- Hsu, Y., & Nelson, D. (1981). Adult performance on the Southern California Kinesthesia and Tactile Perception Tests. American Journal of Occupational Therapy, 35, 788-791.
- Izraelevitz, T., Fisher, A., & Bundy, A. (1985). Equilibrium reactions in preschoolers. Occupational Therapy Journal of Research, 5, 154-169.
- Jones, C., & Monkhouse, M. (1981). Balcones Sensory Integration Screening Test-Revised. Austin, TX: Texas Occupational Therapy Association, Inc.
- Kielhofner, G. (1982). Qualitative research: Part one--Paradigmatic grounds and issues of reliability and validity. Occupational Therapy Journal of Research, 2, 67-79.
- Kielhofner, G., & Nelson, C. (1983). A study of patient motivation and cooperation/participation in occupational therapy. Occupational Therapy Journal of Research, 3, 35-46.
- Kimball, J. (1981). Normative comparison of the Southern California Postrotary Nystagmus Test: Los Angeles vs. Syracuse data. American Journal of Occupational Therapy, 35, 21-25.
- Lewko, J. (1976). Current practices in evaluating motor behavior of disabled children. American Journal of Occupational Therapy, 30, 413-419.
- Mathiowetz, V., Volland, G., Kashman, N., & Weber, K. (1985). Adult norms for the box and block test of manual dexterity. American Journal of Occupational Therapy, 39, 386-391.

- Mathiowetz, V., Weber, K., Kashman, N., & Volland, G. (1985). Adult norms for the nine-hole peg test of finger dexterity. Occupational Therapy Journal of Research, 5, 24-38.
- Montgomery, P., & Rodell, D. (1982). Effect of state on nystagmus duration on the Southern California Postrotary Nystagmus Test. American Journal of Occupational Therapy, 36, 177-182.
- Morrison, D., & Sublett, J. (1983). Reliability of the Southern California Postrotary Nystagmus Test with learning disabled children. American Journal of Occupational Therapy, 37, 694-698.
- Nelson, D., Weidensaul, N., Shih, L., & Anderson, V. (1984). The Southern California Postrotary Nystagmus Test and electronystagmography under different conditions of visual input. American Journal of Occupational Therapy, 38, 535-540.
- Ottenbacher, K. (1984). The significance of power and the power of significance: Recommendations for occupational therapy research. Occupational Therapy Journal of Research, 4, 38-50.
- Ottenbacher, K., Dauck, B., Grahn, V., Gevelinger, M., & Hassett, C. (1985). Reliability of the behavioral assessment of oral functions in feeding. American Journal of Occupational Therapy, 39, 436-440.
- Petersen, P., Goar, D., & Van Deusen, J. (1985). Performance of female adults on the Southern California Visual Figure-Ground Perception Test. American Journal of Occupational Therapy, 39, 525-530.
- Petersen, P., & Wikoff, R. (1983). The performance of adult males on the Southern California Figure-Ground Visual Perception Test. American Journal of Occupational Therapy, 37, 554-560.
- Price, A., Gilfoyle, E., & Myers, C. (Eds.). (1976). Research in sensory-integrative development. Rockville, MD: American Occupational Therapy Association, Inc.
- Punwar, A. (1976). A commentary on Dr. Lewko's article. American Journal of Occupational Therapy, 30, 420-421.

- Punwar, A. (1982). Expanded normative data: Southern California Postrotary Nystagmus Test. American Journal of Occupational Therapy, 36, 183-187.
- Royeen, C., & Kannegieter, R. (1984). Factors affecting textural discrimination in normal children. Occupational Therapy Journal of Research, 4, 261-270.
- Royeen, C., Lesinski, G., Ciani, S., & Schneider, D. (1981). Relationship of the Southern California Sensory Integration Tests, the Southern California Postrotary Nystagmus Test, and clinical observations accompanying them to evaluations in otolaryngology, ophthalmology, and audiology: Two descriptive case studies. American Journal of Occupational Therapy, 35, 443-450.
- Saeki, K., Clark, F., & Azen, S. (1985). Performance of Japanese and Japanese-American children on the Motor Accuracy-Revised and Design Copying tests of the Southern California Sensory Integration Tests. American Journal of Occupational Therapy, 39, 103-113.
- Smith, S. (1983). Performance difference between hands in children on the Motor Accuracy Test-Revised. American Journal of Occupational Therapy, 37, 96-101.
- Sparling, J., & Rogers, J. (1985). Feeding assessment: Development of a biopsychological instrument. Occupational Therapy Journal of Research, 5, 3-23.
- Wilson, D., Baker, L., & Craddock, J. (1985). Functional test for the hemiparetic upper extremity. American Journal of Occupational Therapy, 38, 159-164.
- Yerxa, E. (1982). The issue--A response to testing and measurement in occupational therapy: A review of current practice with special emphasis on the Southern California Sensory Integration Test. American Journal of Occupational Therapy, 36, 399-404.
- Yerxa, E., Barber, L., Diaz, O., Black, W., & Azen, S. (1983). Development of a hand sensitivity test for the hypersensitive hand. American Journal of Occupational Therapy, 37, 176-181.
- Ziviani, J., Poulsen, A., & O'Brien, A. (1982). Correlation of the Bruinks-Oseretsky Test of Motor Proficiency with the Southern California Sensory Integration Tests. American Journal of Occupational Therapy, 36, 519-523.

APPENDIX A

Subject Questionnaire and Consent Forms

Subject Number: _____

SUBJECT QUESTIONNAIRE AND CONSENT FORM

This questionnaire is designed to provide information about the subject group of which you are a part. This information is needed for the research data. Your signature on the second page of this form indicates your consent to be a part of this project. Your name will not be used nor will personal identification be possible. Your questionnaire and scoring form will have subject numbers only. Please provide information as accurately as possible. Thank you for your time and willingness to participate.

Years of occupational therapy experience _____ Years of pediatric experience _____

Setting in which you are currently employed _____

Experience with the Balcones Sensory Integration Screening Test-Revised (BSIST-R)

Please check statements that apply:

- _____ 1. Took course in test administration given in the fall of 1981.
- _____ 2. Have learned to use the BSIST-R through independent study of the protocol manual and experience.
- _____ 3. Have never used the BSIST-R but have read the manual prior to watching the videotape.
- _____ 4. Currently use the BSIST-R.
- _____ 5. Have used the BSIST-R but not currently (within the last year).
- _____ 6. Use the BSIST-R as a main form of evaluation.
- _____ 7. Use the BSIST-R in conjunction with other forms of evaluation. Please list other evaluations used: _____

- _____ 8. Please estimate the number of BSIST-R assessments administered in an average month _____.
- _____ 9. Please estimate the total number of times you have administered the BSIST-R within the last year _____.

Experience with the Southern California Sensory Integration Test (SCSIT)

Please check statements that apply:

- _____ 10. Have taken the SCSIT theory course.
- _____ 11. Have taken the SCSIT test administration course.
- _____ 12. Am certified to administer the SCSIT.
- _____ 13. Use sensory integrative evaluation and treatment methods in the past but do not currently use them.
- _____ 14. Use sensory integrative evaluation and treatment methods in my current practice.
- _____ 15. Have never used sensory integrative evaluation and treatment methods in my pediatric practice of occupational therapy.

page 2

I hereby release the Texas Women's University and the undersigned part acting under the authority of the Texas Women's University from any and all claims arising out of such information gathering, recording, reproducing or publishing as is authorized by the Texas Women's University.

Signature of Participant _____ Date _____

The above consent form and questionnaire was read, discussed, and signed in my presence. In my opinion, the person signing said consent form did so freely and with full knowledge and understanding of its content.

Representative of Texas
Women's University _____ Date _____

APPENDIX B

Subject Scoring Sheet

Subject Number: _____

Balcones Sensory Integration Screening
Protocol

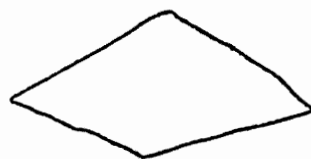
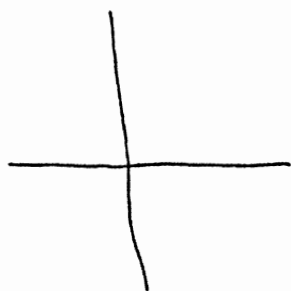
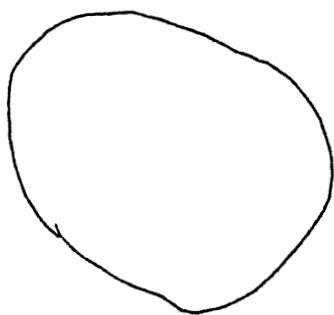
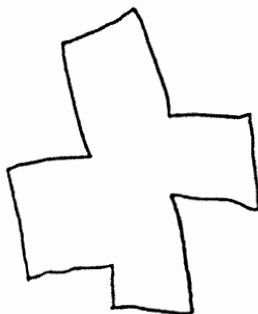
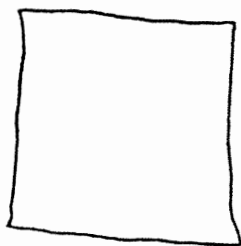
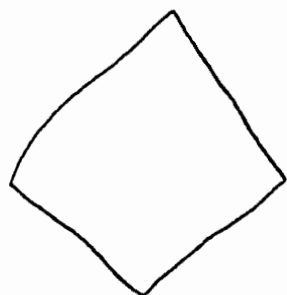
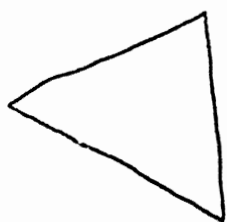
STUDENT: _____ SEX: ♀ ♂ DOB: _____ AGE: _____
RESOURCE _____
TEACHER: _____ SCHOOL: _____ GRADE: _____
PREFERRED _____
HAND: _____ DATE OF SCREENING: _____ EXAMINER: _____

TASK	1	2	3	4
FINGER TO NOSE				
EYES OPEN: RIGHT				
LEFT				
EYES CLOSED: RIGHT				
LEFT				
HEEL-TOE WALKING				
STANDING BALANCE				
EYES OPEN: RIGHT				
LEFT				
EYES CLOSED: RIGHT				
LEFT				
DYNAMOMETER (2 ATTEMPTS)				
RIGHT: _____				
LEFT: _____				
TLR PRONE: _____				
SUPINE: _____				
ATNR (WEIGHTBEARING)				
RIGHT: _____				
LEFT: _____				
OCULAR MOTOR				
RIGHT: _____				
LEFT: _____				
BOTH: _____				
CONVERGENCE: _____				
VISUAL MOTOR FORMS				
FORM: _____				
ORGANIZATION: _____				
ARM POSTURES				
STEREOGNOSIS				
RIGHT: _____				
LEFT: _____				
TACTILE GRAPHICS				
RIGHT: _____				
LEFT: _____				
SCORING TOTALS				

COMMENTS: _____

APPENDIX C

Visual Motor Forms Test



Steven Robbins

APPENDIX D

Letter to Participating Therapists

March 12, 1985

Dear

I want to thank you for your willingness to participate in my study of interrater reliability of the Balcones Sensory Integration Screening Test. In the large mailing envelope you will find:

- 1) the videotape (this tape is VHS format not Beta or 1" videotape)
- 2) copies of the protocol manual ()
- 3) numbered protocols with questionnaires and consent forms attached (#17 → 30)
- 4) the visual motor test cards and the stereognosis test card
- 5) copies of the child's performance of the Visual Motor Forms item
- 6) letters to other therapists who are helping with this study- you will be mailing to one of these therapists when done
- 7) an additional mailing envelope in case this one becomes too tattered.

What you need to do for me is this:

- 1) Find a VHS tape deck and television on which to play the tape.
 - 2) Identify those therapists who can participate in the study according to the following three groups- if you are the only therapist viewing the tape, skip this step-
 - Group 1- pediatric therapists with ^{at least} one year experience who took the Balcones course and have given the test at least 10 times with 3 test administrations in the last year.
 - Group 2- pediatric therapists with ^{at least} one year experience who learned to use the Balcones through reading the protocol manual and have given the test 10 times in the last year with 3 administrations in the last year.
 - Group 3- pediatric therapists with at least one year experience who have never given the Balcones but who have read the protocol manual at least once before viewing the tape. (this last part is important for this group)
 - 3) Have each person take a protocol form and fill out the attached questionnaire and consent form. Explain to the participants that they will not be identifiable from their protocols- only the subject numbers will be used.
 - 4) View the tape. Instructions are included on the tape but to review them: You may view each item several times if necessary. By pushing the 'reverse search' button, the beginning of the item can more easily be located. After viewing the item, 'stop' the tape player and score the item on the protocol sheet either by checking the correct column or by writing the number score in the first column. The protocol manual can be read and used for scoring. Items with 'eyes open' and 'eyes closed' can be viewed and scored individually just as you would do in an actual testing situation.
- PLEASE DO NOT DISCUSS YOUR SCORING WITH OTHERS VIEWING THE TAPE. This will bias the results. After you've completed the tape, feel free to discuss it but DO NOT CHANGE YOUR PROTOCOL SHEETS.

page 2

Other notes for viewing the tape-

Dynamometer- there will be only one score computed in spite of the fact that there are two spaces next to the item on the protocol.

Ocular Motor- this item can be viewed and scored in parts ie right eye, then left eye, then both.

Visual Motor Forms- this is scored by looking at copies of the child's drawings which are included in the envelope. The test cards are included as well with the numbered sequence of presentation on the back.

Stereognosis- the test card is included and be sure to point out the position of the circle and oval as they appear distorted on the video and are hard to distinguish.

Once again, PLEASE DON'T DISCUSS YOUR SCORING WITH OTHERS VIEWING THE TAPE UNTIL AFTER YOU ARE FINISHED AND DO NOT CHANGE YOUR PROTOCOL SHEETS ONCE COMPLETED.

- 5) When done- it should take at most 30 to 45 minutes- put the protocols and questionnaires on the bottom of the stack, put the videotape, protocol manuals, test cards etc. back in the mailing envelope and mail to the person on the enclosed mailing label using the label and stamps provided. Please staple the envelope so it can be used again. If too torn up, use the new one provided.
- 6) Any questions, call me-(817) 731-9919.

I can't thank you all enough and I hope that this has been spelled out clearly enough to not be confusing or frustrating to do for me.

Sincerely,

Mary Ann

APPENDIX E

Balcones Sensory Integration Screening Test-Revised Protocol Manual

BALCONES SENSORY INTEGRATION SCREENING

Protocol Manual

***Revised Edition
1981***



**Cynthia Jones, MS, OTR
Mary Ann Monkhouse, OTR**

Balcones Sensory Integration Screening Protocol

Contents

I. Rationale..... 2

II. Historic Development..... 3

III. Introduction to Revised Edition..... 4

IV. Administration Instruction..... 5

V. Administration and Scoring Criteria..... 6

VI. Cited References.....17

VII. Balcones Sensory Integration Screening Protocol.....18

VIII. Balcones Sensory Integration Screening Profile.....19

Rationale

With the increasing knowledge of brain behaviors, it was the intent of this instrument to provide a screening base for the most basic brain stem and associated area responses.

An instrument was needed which could be used alone as a sensory integrative screening instrument or as a component of a more detailed neuropsychological battery. This composite style instrument was designed to aid in problem identification.

C.G.J.

Historic Development

The Sensory Integrative Screening Protocol was developed under a Title IV-C, ESEA Grant (1976-1977; 1977-1978; 1978-1979) to be used as the primary method of overall screening of all special education students K-12 who were serviced by the Balcones Special Services Cooperative.

The intent of the screening was to identify the students who might benefit formal sensory-motor/sensory integrative testing and programming. Using the information gained from the screening can be helpful in identifying special neuro-behavioral; behavioral; and/or classroom performance. The findings can also aid the classroom teacher's expectation of the child's performance.

Due to the time constraints on the original project, development items selected for the screening instrument have a prenormative base. That is, the items were developed from previously accepted and/or original batteries. Tests utilized are listed in the attached references.

After two and one half years of use, minimal changes were made. The resulting Screening Instrument changes included the addition of tactile based screening items as well as "clarifying" the language in the criteria.

Special acknowledgement must be given to the progressive, creative staff who believed strongly in the "whole" child and neuropsychological functioning as it relates to learning in the public school setting. The entire staff of the Balcones Special Services Cooperative from early in 1976 was involved in the IV-C Grant. Special thanks should be given to Vikki Vernon who labored with the early instrument design and made Motor Labs work and Diane Friou for her direction and assistance in the development and norming of the Profile Grid. And, of course, the driving force behind all of us during those years, Ruth A. Haak, PhD., Director of the Balcones Coop.

Introduction to Revised Edition

Increased and expanded use of the original Balcones Sensory Integration Screening has necessitated further revisions or clarification. The Balcones Sensory Integration Screener, Revised Edition, has been published by Texas Occupational Therapy Association, Inc. to meet this expanded use. It is presented in a more comprehensive format with test materials provided (except dynamometer and stopwatch). It is important that the standardized method developed is followed by examiners. This is a screening instrument and is designed so that the examiner is to encourage the subject's best performance.

Participants in testing workshops presented in August, 1981, will contribute to additional data needed for expanded standardization of the instrument.

Use of the Balcones Sensory Integration Screening can assist the occupational therapist in providing quality services to the increasingly diverse populations our profession is serving.

CGJ/MAM

COPIES AVAILABLE FROM:
Texas Occupational Therapy Assn., Inc.
P. O. Box 2042
Austin, Texas 78768
512/479-8792

Administration Instructions

- I. Complete Identifying Information - (must be completed)
 - ...following subject's name; circle sex of the subject
 - ...date of birth (DOB) and age to the nearest month
 - ...resource teacher or primary contact in the nearest school
 - ...school's name
 - ...subject's current grade
 - ...date of the screening
 - ...identification of examiner
 - ...indicate hand dominance (dominance determined by hand used to draw Visual Motor Forms - if both hands used, please note)
- II. Complete Scoring Section according to directions for each subtest
 - ...procedure for administration is stated with each subtest
 - ...note letter of scoring criteria next to check in appropriate score column
- III. Comments
 - Be sure to note:
 - ...testing behaviors
 - ...comments of the student
 - ...unusual or extreme reactions, i.e., overflow; right-left confusion; random behaviors
 - ...recommendations for further intervention
- IV. Materials needed
 - ...protocol manual
 - ...stop watch
 - ...pen light
 - ...dynamometer
 - ...10 feet of tape (to be placed on the floor)
 - ...2 pencils
 - ...blank paper (8 1/2 x 11)
 - ...visual motor forms
 - ... 8 "shape-o-ball" forms
 - ...pointing response card
 - ...screening protocol and profile forms
- V. Totaling the Score
 - A numerical total can be reached by adding the number of (✓) per column, then multiplying each by its receptive column number. Cut-off points can be established as to population being served.
- VI. S-I Screening Implications
 - See Sensory Integration Profile

Administration and Scoring Criteria

I. Finger-Nose-Finger (B, G)

Procedure

- Subject sits in a supporting chair with feet flat on the floor. Examiner demonstrates. After demonstration, the subject extends LEFT arm with LEFT Index finger up (eyes open). Take RIGHT INDEX finger and touch tip of the LEFT INDEX finger and then the nose, index finger then nose, etc. The moving arm is scored. Have subject repeat the pattern for 10 seconds.
- REPEAT with RIGHT ARM extended.....eyes open
- REPEAT with LEFT ARM extended.....eyes closed
- REPEAT with RIGHT ARM extended.....eyes closed

Scoring

- 4.....smooth direct movements of arm
 extended arm steady with no elbow flexion
 lightly touches nose and extended finger tip
- 3.....a. irregular or weaving movements of arm in motion
 b. minimal tremor or swaying of extended arm
 c. missed contact of either finger or nose (2 or less times)
 (score for 7 years or more)
 d. minimal bending of extended arm (less than 30°)
 e. minimal pushing of the tip of the nose or finger (2-3
 times)
 f. slow deliberate movements
 (2 or more errors in 3 score 2 points)
- 2.....a. confused jerky movements
 b. marked tremor or swaying of either arm
 c. missed contact with both nose and finger (2 or more times)
 d. missed contact with either nose or finger more than 20% of
 the time (1 out of 5 times)
 e. marked bending of extended arm (more than 30°)
 (2 or more errors in 2 score 1 point)
- 1.....a. unable to sustain pattern
 b. 2 or more errors in #2

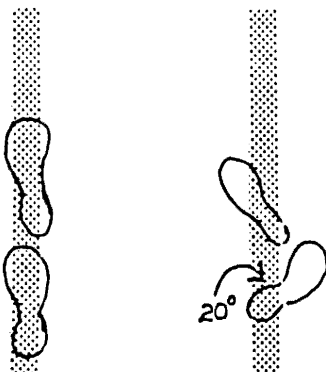
II. Heel Toe Walking (A, B, G)

Procedure - examiner may demonstrate item.

- subject walks 10 feet on a tape or chalk line without shoes
- hands are to be on hips
- foot placement is on the line
- heel is to be within 1 inch of toe on each placement

Scoring

- 4....hands remain on hips
feet are placed on the line
heel toe distance remains 1" or less
smooth forward walk
- 3....a. removes hand from hip (one hand)
b. stepped off once
c. heel toe distance off 1 or 2 times
d. steps correctly and rotates less than 20°
(2 or more errors in 3 score 2)
- 2....a. removes both hands from hips
b. stepped off 2 times
c. heel toe distance off 3 times
d. steps rotate more than 20°
(2 or more errors in 2 score 1)
- 1....a. removes hands from hips 2 or more times
b. stepped off 3 or more times
c. heel toe distance off 3 or more times
d. cannot maintain consistent placement of feet
e. 2 or more errors on #2



III. Standing Balance (A, C, G)

Procedure

- subject stands with no external support, eyes open, shoes remain off
- check to see if the subject knows RIGHT and LEFT foot
- arms are folded across chest
- instruct the subject to "stand on this foot" with eyes open
(Ex. touches S RIGHT leg just below the knee - standing leg is scored)

REPEAT - LEFT leg with eyes open

REPEAT - RIGHT leg with eyes closed

REPEAT - LEFT leg with eyes closed

Scoring

- 4.....gentle swaying
maintains position for 25 seconds
- 3.....a. marked swaying but arms remain folded across chest
b. maintains position 20+ seconds
- 2.....a. jerky swaying or hopping
b. opposite foot makes brief contact with the floor 2 or more times
c. maintains position less than 20 seconds
(2 or more errors in 2 score 1)
- 1.....a. unable to maintain position without wrapping one leg around the other
b. able to maintain position on one foot 5 seconds or less
c. arms unfolded, balance maintained

IV. Dynamometer (B, D, G)

Procedure (kilogram scales are preferred because of the smaller increments of measure)

- subject is given 2 readings with each hand
- subject stands with arms straight and hands at side
- subject grasps dynamometer with RIGHT hand and holds with arm about 6" from the body
- squeeze dynamometer
- rest and repeat with the RIGHT hand
- repeat the sequence for the LEFT hand
- write dynamometer readings on screening form by corresponding hand

Scoring

- scoring completed after entire Screening is administered
 - "dominance" is determined by hand choice for pencil activities on Visual Motor Forms Items (#8)
 - average the 2 scores for each hand (i.e., RIGHT 39/41 = 40)
(LEFT 38/38 = 38)
- 4.....dominant hand 10% stronger than non-dominant hand
 - 3.....dominant hand is less than 10% stronger than non-dominant hand (but not equal)
 - 2.....a. grip strength is equal
 - b. dominant hand is greater than 10% stronger than non-dominant hand
 - c. non-dominant hand stronger but less than 10% of the dominant hand
 - 1.....non-dominant hand 10% greater than dominant hand

V. Tonic Labyrinthine Reflex (TLR) (A, E, F)

Procedure

- each position will be assumed and maintained 20 seconds
- Ex demonstration is allowed including placing subject into position
- subject must assume position from resting position and maintain

PRONE----subject lays on his stomach, arms are flexed at side (as shown below)----subject is instructed to "bow-up" as if to rock----knees should not be flexed



SUPINE----subject lays on his back, arms flexed and resting on chest----subject is instructed to "curl-up" and to hold that position----flexion at all joints should be noted (ankles can be crossed)



Scoring

- 4.....can assume and maintain 20 seconds
- 3.....can assume and maintain 20 seconds with obvious strain
- 2.....can maintain with difficulty less than 20 seconds
 - a. decreased respiration
 - b. extreme muscle tension noted
 - c. knee flexion in prone
- 1.....unable to assume or maintain

or

 maintains 5 seconds or less

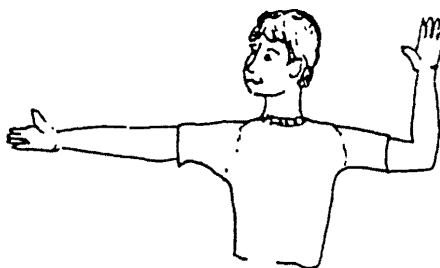
VI. Assymetrical Tonic Neck Reflex (ATNR) (E, F)

(QUADRIPEDE - WEIGHTBEARING)

Procedure

- subject will assume an all-fours position (hands and knees) and slightly flex both elbows (elbows need to be "unlocked")
- subject will then be told to relax his head and allow the examiner to turn it

NOTE: When head is turned toward the RIGHT shoulder and LEFT elbow flexes, this is a positive ATNR-to the RIGHT (pictured below)



Scoring

- 4.....Relaxes head easily
Allows examiner to rotate head
No elbow flexion or shoulder depression noted with rotation
- 3.....a. Allows examiner to rotate head but minimal shoulder depression can be noted
b. Difficulty in relaxing enough to allow examiner to turn head
c. Anticipated direction of movement and leads
(2 or more errors in #3 score 2)
- 2.....Shoulder depression and minimal elbow flexion
- 1.....a. Head held rigidly - very difficult to rotate
b. Extreme elbow flexion or falling to the floor

VII. Ocular Control (A, G)

Procedure

- subject should be seated in a good supporting chair with feet on the floor
- Instruct the subject to "follow the pen light with your eyes only and do not move your head" (With young subjects slight assistance may be needed to stabilize head on binocular items.)
- Cover LEFT eye (allow subject to cover own eye with LEFT hand)
- Move the pen light through the visual field in the following patterns:



REPEAT with RIGHT eye covered (same patterns)

REPEAT with BOTH eyes (binocular usage) ((same patterns))

CONVERGENCE: Hold pen light 6 to 8" from the bridge of the nose.
Slowly move light towards the bridge (to within 1" of bridge)



Scoring

RIGHT-LEFT-BOTH

- 4.....Smooth, even movements in all planes
- 3.....Basic smooth movements with slight wandering or uneven movement at extremes only
- 2.....a. Uneven or jerky movements
 - b. Midline avoidance (2 or more eye blinks at midline)
 - c. Loses target 1 or 2 times
(2 errors on #2 score 1)
- 1.....a. Eyes do not work together (binocular item only)
 - b. Obvious midline jump or darting
 - c. Loses target 3 or more times (unable to maintain contact)

CONVERGENCE:

- 4.....Smooth, even convergence at 3" or less
- 3.....Slight delay or hesitation, but still converge
- 2.....Jerky, slow, inaccurate
- 1.....Eyes break apart - did not converge

VIII. Visual Motor Forms (A, D)

Procedure

- Subject is seated at the table
- Pencil is centered on the table in front of the subject
- The examiner hands subject a sheet of paper and allows him to place it on the table
- The forms are placed in front of the subject (centered above the paper)
- Tell the subject, "There will be 7 forms approximately this size. You need to draw them this size so that they will all fit on this page."
- One form is exposed at a time - if subject appears to be confused or grossly in error, additional instructions can be given.

Chronological Age Difference

- 5.0 to 5.11 use forms 1-4
- 6.0 to 6.11 use forms 1-5
- 7.0 and above use all forms

Scoring - Form

- 4.....Adequate performance throughout
- 3.....a. Mirror distortions (particularly in diamonds)
 - b. Sketchy lines (light lines or repetitive light lines)
 - c. Shakey drawing (fairly firm lines but very unsure)
 (2 or more errors in #3 score 2)
- 2.....Segmented drawings but recognizable
- 1.....a. Dog ears
 - b. Gross segmenting
 - c. Any of drawn forms unrecognizable

Organization

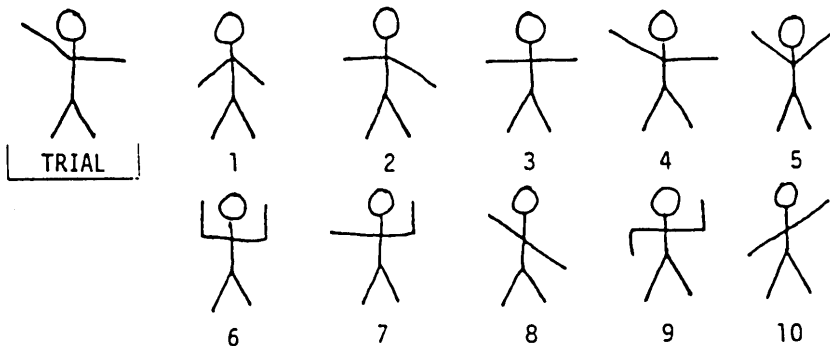
- 4.....Adequate LEFT to RIGHT progression or top to bottom; size adequate
- 3.....a. Other organization form
 - b. Five or more forms organized on the page (if 7 attempted - 4 out of 5)
- 2.....a. Markedly too large or small
 - b. Less than 5 drawings are organized on the page (if 7 attempted)
 - c. Overlapping of drawings
- 1.....No organization apparent (random placement)

* NOTE PREFERRED HAND.*

IX. Arm Postures (A, F, G)

Procedure

- Subject and examiner stand face to face
- Subject is told to "Do exactly as I do."
- Allow one trial and correct if subject mirrors (only once) then continue allowing mirroring (with younger subjects, mirroring is to be expected)
- Examiner should mirror diagrams



Scoring

- 4.....a. Postures are exact with examiner (RIGHT-RIGHT and LEFT-LEFT) OR
 - b. Consistently and correctly mirrors
- 3.....Self correct after first posture and imitates or mirrors correctly the remaining postures
- 2.....a. Hesitation (3 seconds or more)
 - b. Shoulders flexed forward or arms internally rotated
 - c. One uncorrected error
 (2 or more errors in #2 score 1)
- 1.....More than 1 error

REMINDER: This is a test of motor planning through proprioception not visual motor. Visual correction of postures should count against subject.

X. Stereognosis (B, C, D)

Procedure

- Shield forms with response card. The subject's manipulating hand is behind the response card out of sight.
- Present forms 1-4 in order shown below to RIGHT hand
- Repeat with forms 5-8 on LEFT hand (do not alternate hands)
- Subject must point to correct response on card with non-manipulating hand

ORDER OF PRESENTATION:

<u>RIGHT HAND</u>		<u>LEFT HAND</u>	
1.	 (circle)	5.	 (triangle)
2.	 (trapezoid)	6.	 (pentagon)
3.	 (oval)	7.	 (rocker bottom triangle)
4.	 (square)	8.	 (cross)

Scoring

4.....4 forms correct
 3.....3 forms correct
 2.....2 forms correct
 1.....1 form correct

XI. Tactile Graphics (B, C, D)

Procedure

- With vision occluded examiner draws on back of subject's hand with rounded eraser
- Subject retraces design with vision

STIMULUS FORMS (use "I" as teaching form) (→ denotes direction of drawing)

RIGHT HAND	LEFT HAND
1. ↓	5. — →
2. ○ ↻	6. X ↘
3. ∞ ↻	7. 3 ↻
4. < ↗	8. ○ ↻

Scoring (Right and Left hand are scored separately)

- 4.....a. Draws each shape correctly as drawn
 b. Draws each shape correctly but rotates location
- 3.....3 shapes correct
- 2.....2 shapes correct
- 1.....1 - 0 shape correct

NOTE tactile behaviors while testing in comments section

Cited Test References*

- A THE PURDUE PERCEPTUAL MOTOR SURVEY
Eugene G. Roach, Ph. D.
Newell C. Kephart, Ph. D.
1966
Charles E. Merrill Publishing Company
Columbus, Ohio
- B McCARRON ASSESSMENT OF NEUROMUSCULAR DEVELOPMENT
Lawrence McCarron, Ph. D.
McCarron-Dial Systems
P. O. Box 45628
Dallas, Texas 75245
- C SOUTHERN CALIFORNIA SENSORY INTEGRATION TEST
A. Jean Ayres, OTR, Ph. D.
Western Psychological Services
Los Angeles, California 90025
- D CLINICAL NEUROPSYCHOLOGY: CURRENT STATUS AND APPLICATIONS
Ralph M. Reitan, Editor
Halstead Press Book
John Wiley & Sons
New York, 1974
- E REFLEX TESTING METHODS FOR EVALUATING CNS DEVELOPMENT
Mary Fiorentino, OTR
Charles C. Thomas, Publishers
Springfield, Illinois
- F OBSERVATION OF CLINICAL SIGNS
Ayres, A. Jean adapted by BSSCO
1978
- G ESSENTIALS OF A NEUROLOGICAL EXAMINATION
Francis A. Vazuka, M. D.
Smith, Kline & French Laboratories
1968

* Following each subtest title are one or more capital letters.
These letters reflect which of the above major standardized
tests were used as a basis for the item.

Balcones Sensory Integration Screening
Protocol

STUDENT: _____ SEX: ♀ ♂ DOB: _____ AGE: _____
RESOURCE _____
TEACHER: _____ SCHOOL: _____ GRADE: _____
PREFERRED _____
HAND: _____ DATE OF SCREENING: _____ EXAMINER: _____

TASK	1	2	3	4
FINGER TO NOSE				
EYES OPEN: RIGHT				
LEFT				
EYES CLOSED: RIGHT				
LEFT				
HEEL-TOE WALKING				
STANDING BALANCE				
EYES OPEN: RIGHT				
LEFT				
EYES CLOSED: RIGHT				
LEFT				
DYNAMOMETER (2 ATTEMPTS)				
RIGHT:				
LEFT:				
TLR PRONE:				
SUPINE:				
ATNR (WEIGHTBEARING)				
RIGHT:				
LEFT:				
OCULAR MOTOR				
RIGHT:				
LEFT:				
BOTH:				
CONVERGENCE:				
VISUAL MOTOR FORMS				
FORM:				
ORGANIZATION:				
ARM POSTURES				
STEREOGNOSIS				
RIGHT:				
LEFT:				
TACTILE GRAPHICS				
RIGHT:				
LEFT:				
SCORING TOTALS				

COMMENTS: _____

Balcones Sensory Integration Screening Profile

The Balcones Sensory Integration Screening was administered in the Spring of 1979 to two first grade classes, two second grade classes and two third grade classes in two public schools. The total population tested was 130. The age of the subjects ranged from six to nine.

The racial distribution of the sample was that of the population at large in Texas.

The items measured by the Sensory Integration Screening were factor analyzed and eight major components were defined. These factors were used to make up the Sensory Integration Screening Profile.

To use the Sensory Integration Screening Profile, first enter the raw scores of each item from Protocol into the column titled "Item Raw Score."(*) Add the item raw scores and enter the total in the column labeled "Total Raw Score." Each total raw score can then be plotted on the scale line to the right. A profile or graph can be constructed by joining the plotted points.

The vertical lines below the "0" in the section "Raw Score in Standard Score Units" represent mean score on each scale of the normative sample. From -1.0 standard deviation to +1.0 standard deviation represents the range of scores of the middle 68% of the students in the normative sample.

(* The abbreviations for the items can be found on the Profile Factor Sheet - next page.)

Balcones Sensory Integration Screening Profile Factor Sheet

Factor 1	<u>OCULAR MOTOR CONTROL</u>	
	Ocular Control - Right.....OCR	→ M = 7.68
	Ocular Control - Left.....OCL	SD = 3.14
	Ocular Control - Both.....OCB	
Factor 2	<u>PROPRIOCEPTION</u>	
	Finger-Nose-Finger (eyes open) Right.....FNFOR	→ M = 14.60
	Finger-Nose-Finger (eyes open) Left.....FNFOL	SD = 3.06
	Finger-Nose-Finger (eyes closed) Right.....FNFCR	
	Finger-Nose-Finger (eyes closed) Left.....FNFCL	
	Arm Postures.....AP	
Factor 3	<u>Primitive Reflex-Brain Stem</u>	
	<u>Midline Interaction</u>	
	Tonic Labyrinthine - Prone.....TLR-P	→ M = 12.05
	Tonic Labyrinthine - Supine.....TLR-S	SD = 2.60
	Asymmetrical Tonic Neck - Right.....ATNR-R	
	Asymmetrical Tonic Neck - Left.....ATNR-L	
Factor 4	<u>VESTIBULAR</u>	
	Standing Balance (eyes open) Right.....SBO-R	→ M = 9.76
	Standing Balance (eyes open) Left.....SBO-L	SD = 2.55
	Standing Balance (eyes closed) Right.....SBC-R	
	Standing Balance (eyes closed) Left.....SBC-L	
Factor 5	<u>TACTILE FEEDBACK (perception)</u>	
	Tonic Labyrinthine - Prone.....TLR-P	→ M = 12.74
	Tonic Labyrinthine - Supine.....TLR-S	SD = 2.36
	Tactile Graphics - Right.....TG-R	
	Tactile Graphics - Left.....TG-L	
Factor 6	<u>STEREOGNOSIS</u>	
	Stereognosis - Right.....ST-R	→ M = 6.22
	Stereognosis - Left.....ST-L	SD = 1.28
Factor 7	<u>FORM PERCEPTION (visual/motor)</u>	
	Visual Motor Forms.....VMF	→ M = 4.78
	Visual Motor Organization.....VMO	SD = 1.61
Factor 8	<u>LATERALIZATION</u>	
	Heel Toe Walking.....HTW	→ M = 7.15
	Preferred Hand.....PH	SD = 1.73
	Dynamometer.....DYNO	

Balcones Sensory Integration Screening
Profile Sheet

STUDENT NAME: _____ DATE SCREEN: _____

FACTOR	ITEM RAW SCORE	TOTAL RAW SCORE	RAW SCORE IN STANDARD SCORE UNITS						
			-3	-2	-1	0	+1	+2	+3
1. OCULAR MOTOR CONTROL	OCR ____ OCL ____ OCB ____								
				2	5	8	11	14	
2. PROPRIOCEPTION	FNFOR ____ FNFDL ____ FNFCR ____ FNFDL ____ AP ____		6	9	12	15	18	21	
3. PRIMITIVE REFLEX- BRAIN STEM MIDLINE INTERACTION	TLR-P ____ TLR-S ____ ATNR-R ____ ATNR-L ____		3	6	9	12	15	18	
4. VESTIBULAR	SBO-R ____ SBO-L ____ SBC-R ____ SBC-L ____		1	4	7	10	13	16	
5. TACTILE FEEDBACK (Perception)	TLR-P ____ TLR-S ____ TG-R ____ TG-L ____		7	9	11	13	15	17	
6. STEREOGNOSIS	ST-R ____ ST-L ____		3	4	5	6	7	8	
7. FORM PERCEPTION (Visual/Motor)	VMF ____ VMO ____			2	3	5	7	9	
8. LATERALIZATION	HTW ____ PH ____ DYN ____								
				3	5	7	9	11	