

**TEST-RETEST RELIABILITY OF THE
SLOSSON INTELLIGENCE TEST AND THE
SHIPLEY-HARTFORD INTELLIGENCE TEST**

EDWARD HAROLD LESTER

TEST-RETEST RELIABILITY OF THE
SLOSSON INTELLIGENCE TEST AND THE
SHIPLEY-HARTFORD INTELLIGENCE TEST

An Abstract
Presented to
the Graduate Council of
Austin Peay State University

In Partial Fulfillment
of the Requirements for the Degree
Master of Arts

BY

Edward Harold Lester

August 1971

ABSTRACT

The purposes of this study were to determine the reliability of the Slosson Intelligence Test and the Shipley-Hartford Institute of Living Scale on an adult population and to determine the correlation between the two.

The subjects were 40 undergraduate students currently enrolled in lower division psychology courses at Austin Peay State University, Clarksville, Tennessee. The Slosson Intelligence Test and Shipley-Hartford Institute of Living Scale were administered individually to each subject at the beginning of the Spring Quarter of 1971 and again the week before final examinations in the same quarter.

The Pearson Product-Moment Correlation technique was used to compare SIT IQ scores on the first measure with those of the second measure. The scores on the Shipley-Hartford scale were treated by the same technique in the same manner. The IQ scores on the first measure of the SIT were compared with the scores on the first measures of the Shipley-Hartford Scale using the aforementioned statistical technique. The scores of retests were also compared in this manner. The correlation coefficients obtained in this study ranged from .46 to .83 and were all significant beyond the .01 level.

The results of this study indicate that the SIT and the Shipley-Hartford are reliable instruments for measuring adult intelligence.

TEST-RETEST RELIABILITY OF THE
SLOSSON INTELLIGENCE TEST AND THE
SHIPLEY-HARTFORD INTELLIGENCE TEST

A Thesis
Presented to
the Graduate Council of
Austin Peay State University

In Partial Fulfillment
of the Requirements for the Degree
Master of Arts

BY

Edward Harold Lester


August 1971


To the Graduate Council:

I am submitting herewith a Thesis written by Edward Harold Lester entitled "Test-Retest Reliability of the Slosson Intelligence Test and the Shipley-Hartford Intelligence Test." I recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Arts, with a major in Psychology.

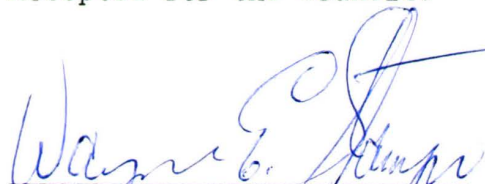

Major Professor

We have read this thesis and
recommend its acceptance:


Minor Professor
or
Second Committee Member


Third Committee Member

Accepted for the Council:


Dean of the Graduate School

ACKNOWLEDGEMENTS

The writer wishes to express sincere appreciation to Dr. John D. Martin, Associate Professor of Psychology, Austin Peay State University, who suggested the problem and who assisted and counseled him during the course of the study; to Dr. Elizabeth Stokes and Dr. Garland E. Blair, Department of Psychology, for their comments and criticisms which have proven invaluable in the completion of this thesis.

The author wishes to thank his fellow graduate students, Gloria Armstrong, Kathy Simpson, and George Serednesky, for their assistance in the administration of the tests used in the research and for their encouragement throughout the study.

Gratitude is also extended to the students who volunteered to serve as subjects in this study.

I wish to express special appreciation to my wife, Brenda, and to my daughter, Tracy, for their support, patience and understanding.

TABLE OF CONTENTS

	Page
LIST OF TABLES	v
Chapter	
I. INTRODUCTION TO THE PROBLEM	1
II. METHOD	8
The sample	8
Description of the Instruments	8
Administration and Scoring	8
III. RESULTS	10
IV. DISCUSSION	11
V. SUMMARY	13
REFERENCES	14

LIST OF TABLES

Table	Page
1. Test-Retest correlations of the SIT and Shipley, and correlations between the SIT and the Shipley.	10

CHAPTER I

INTRODUCTION TO THE PROBLEM

The Slosson Intelligence Test (SIT) was introduced in 1963 by Richard L. Slosson for use as a quick "individual" intelligence test. Slosson's purpose in constructing this test was to provide an abbreviated form of the Stanford-Binet Intelligence Scale, Form L-M, by Terman and Merrill, which could be quickly and easily administered. The test is mainly composed of selected items from the Stanford-Binet and Gesell Development Tests (Rudolph, 1970).

There are many occasions when the reliability of intelligence tests are very crucial to educators. Decisions regarding type of instruction and level of ability of the student are often based upon intelligence test scores (DeCecce, 1968). This is, however, a misuse of intelligence test scores because "IQ scores may be much less reliable predictors of achievement than previous achievement of the student" (DeCecce, 1968, p. 11). Nevertheless, teachers do use IQ scores for predicting success in school and must therefore be provided with as reliable as possible scores. This accentuates the need to provide reliable tests.

Slosson, in researching the reliability of the SIT, used 139 subjects ranging in age from 4 to 50 years. The heterogeneity of his sample probably somewhat inflated his reliability coefficient (r) (Garrett, 1965). A sample from a more homogeneous population was utilized in the present study in order to more adequately determine the reliability of the SIT on a more restricted age range.

A review of the literature indicated the need for a study of the reliability of the SIT by virtue of the almost complete absence of other studies of this nature with the SIT. Also, shorter tests, quick forms or abbreviated forms, characteristically have less reliability than do longer tests (Anastasi, 1961, p. 161).

Rudolph (1970) investigated the validity of the SIT on an adult population when compared with an accepted and proven individual intelligence test. She administered the Wechsler Adult Intelligence Scale (WAIS) and the SIT to each subject in her study. The correlation coefficient obtained from their comparison was .70 for WAIS Full Scale scores, .73 for WAIS Verbal Scale scores, and .49 for WAIS Performance Scale scores. All of these correlation coefficients were significant beyond the .01 level and therefore indicate that the SIT is, in fact, a valid instrument for measuring intelligence with an adult population. Since validity of a test presupposes reliability

of that test, many reviews concerning the validity of the SIT are found in the present study.

Slosson (1963) reports in the test manual that in his standardization of the test he administered the Stanford-Binet, Form L-M, and the SIT to 141 subjects, and obtained a concurrent validity coefficient of .92. Published research on the SIT is very limited; however, as will be noted from the review of the literature, studies indicate that the SIT correlates highly with older and proven measures of mental ability, such as the Binet and the Wechsler scales.

DeLapa's conclusion from his study using the SIT was that it appeared to be measuring approximately the same abilities as the Binet (DeLapa, 1968). The obtained correlation between the SIT and Binet scores for students in special education classes was .60, and for the students from regular class rooms it was .90. The SIT is therefore a valid test when compared to the Binet.

Poissant (1967) tested 36 slow learners (ages 8 to 12 years) with the Stanford-Binet and later the same children were tested with the SIT. An obtained coefficient of .89 led her to conclude that the SIT is a valid tool for use with slow learners when compared to the Binet.

Kaufman and Ivanoff (1969) assessed the validity of the SIT in a correlational study with the Wechsler Adult Intelligence Scale on a rehabilitation population. Impairments of the subjects were of an emotional nature. Intellectual functioning of the subjects ranged from mild retardation to bright normal, based on obtained WAIS scores. An obtained correlation of .93 between the SIT and the WAIS full scale scores indicates that the two tests seem to be measuring the same types of intellectual functioning. An obtained correlation of .70 between the SIT and the performance section of the WAIS, and an obtained coefficient of .96 between the SIT and WAIS verbal led the authors to conclude that the SIT seemed to be a valid measure of intellectual ability.

Hammill (1969) assessed the reliability of the Slosson when administered and scored by teachers who had not been trained in psychological testing. The test-retest reliability coefficient was .97. The validity study obtained coefficients ranging from .70 to .80 when the SIT was compared to other cognitive and intellectual variables. Hammill concluded that the results he obtained relative to interscorer differences, internal consistency, reliability, and concurrent validity confirms the contention by Slosson that the test can be

administered by teachers, principals and other personnel unskilled in psychological testing.

The validity of the SIT with two age groups of educable mentally retarded children was investigated by Jongeward (1968). One group of 30 subjects was tested with the Binet and SIT. Obtained correlations of .761 between IQ scores, and .806 between mental ages of the SIT and Binet were both significant beyond the .01 level. Jongeward tested a second group of 30 EMR children with the SIT and the Wechsler Intelligence Scale for Children (WISC). He found a correlation of .537 between SIT IQ scores and WISC full scale scores. The SIT and WISC verbal IQ scores correlated .852. All obtained correlation coefficients were significant beyond the .01 level. Correlating SIT scores and WISC performance scores yielded a coefficient of .204, which did not attain significance. His conclusions were that the SIT should not replace the Binet or WISC as a basic tool for assessing the mental abilities of EMR children, but felt that the SIT does have merit as an instrument for screening children for EMR classes.

As was noted from the aforementioned reviews, most reliability and validity studies of the Slosson Intelligence Test have been done with children or with an atypical population such as educable mental retardates or rehabilitation groups. Slosson indicates that his test is

a valid and reliable instrument for measuring adult intelligence as well as the intelligence of children. One of the purposes of this study was to determine the reliability of the SIT on an adult population.

A second purpose of this study was to determine the reliability of the Shipley-Hartford Intelligence Test, and to determine the correlation, if any, between the SIT and the S-H.

The Shipley-Hartford Intelligence Test, or Shipley Institute of Living Scale, was introduced in 1940 as a test designed to aid in detecting mild degrees of intellectual impairment in individuals of normal original (premorbid) intelligence. It should not be used with very obviously deteriorated cases, intellectual subnormals (feeble-minded and borderline), and individuals with language handicaps.

Standardization of the reliability was with 322 Army recruits with obtained coefficients of .87, .89, and .92 on the vocabulary, abstraction, and combined tests respectively. The sample is, again, rather heterogeneous, in terms of mental ability, and therefore may possibly have inflated the obtained coefficients. The literature is lacking concerning other studies of the reliability of the Shipley-Hartford Intelligence Test.

In the present study, a more homogeneous sample was used to determine the reliability of the S-H. The same subjects were utilized in the reliability study of the SIT.

CHAPTER II

METHOD

The Sample

The sample used in this study was undergraduate students enrolled in lower division psychology courses at Austin Peay State University, Clarksville, Tennessee. All participants volunteered to serve as subjects, and received extra points in the aforementioned classes for their participation in the research. The sample was composed of 40 students, of which 25 were females and 15 males.

Description of the Instruments

The Slosson Intelligence Test is a short form individual intelligence test for use with adults and children. The test is essentially verbal except for lower levels where perceptual-motor items are included. The test yields a mental age and a ratio IQ.

The Shipley-Hartford Institute of Living Scale (Shipley) is a written test which can be administered as a group or individual test. It requires no special materials or detailed instructions and is essentially a self-administering test. It is composed of 40 vocabulary items and 20 abstract items. The raw score is obtained by adding the number of correct responses on the vocabulary test to twice the number of correct responses on the abstraction test.

Administration and Scoring

The SIT and Shipley were administered individually to each subject by the present researcher and fellow graduate students with a 45 day interval between the first test and the retest on both measures. Each test was scored according to the manuals of direction.

CHAPTER III

RESULTS

The Pearson Product-Moment technique was used to compute the correlation coefficients. SIT IQ scores on the initial tests were compared to SIT IQ scores on the retests. Shipley raw scores on the initial tests were compared to Shipley raw scores on the retests. SIT IQ scores on the initial tests were compared to Shipley raw scores on the initial tests. SIT IQ scores on the retests were compared to Shipley raw scores on the retests. Table I summarizes the correlations.

TABLE I

Test-Retest correlations of the SIT and Shipley, and correlations between the SIT and the Shipley.

Item	r*
1. SIT Test-Retest	.83
2. Shipley Test-Retest	.80
3. SIT and Shipley 1st testing	.46
4. SIT and Shipley 2nd testing	.54

*All correlations were significant beyond the .01 level.

CHAPTER IV

DISCUSSION

The reliability coefficient of .83 on the SIT obtained in the present study is slightly lower than the reliability coefficient reported in the SIT manual (1963), but is still significant beyond the .01 level. The reliability coefficient of .80 on the Shipley is likewise slightly lower than the reliability coefficient reported in the Shipley manual (1940, 1967), but it is also significant beyond the .01 level. Correlations between these two tests are also significant beyond the .01 level. Although the latter coefficients are significant beyond .01 level, they are somewhat lower than is desirable for reliability coefficients, which usually fall in the .80's or 90's (Anastasi, 1968, p. 78). They are therefore not comparable to a degree which would merit their use as alternates of each other. Each of the tests has been proven by virtue of this study to be within the range of acceptably reliable tests. Lowered reliability coefficients may be explained in terms of lack of sophistication of the examiners and lack of ideal testing conditions. Variability of the test scores, especially the abstraction subtests of the Shipley, may also be explained in terms of the effects of experience with these measures on the second tests. It was also observed that many of the examinees were appreciably less anxious on the second test than they were on the initial

test. The effects of anxiety perhaps contributed to a lessening of the scores on the initial tests. After having become accustomed to the examiner and these tests, the subjects were less anxious and consequently performed better on the second tests than on the initial tests. One such example was one subject who showed a gain of 23 IQ points on the SIT and 11 raw score points gain on the Shipley. This, however, is not indicative of the gains showed by other subjects. The average gain on the SIT for the total population, including the aforementioned score, was 4.3 IQ points. The average gain on the Shipley, including the aforementioned score, was 1.475 raw score points. In view of these slight gains, the effects of practice and reminiscence appear to be inconsequential.

CHAPTER V

SUMMARY

The purposes of this study were to determine the reliability of the SIT and the Shipley and to determine the correlation between the SIT and the Shipley.

The sample used in this study was 40 undergraduate college students. All of the subjects used were volunteers.

The correlation coefficients obtained in this study ranged from .46 to .83 and were all significant beyond the .01 level.

In the light of this study, it is concluded that the SIT is a reliable instrument for measuring adult intelligence. The Shipley has also demonstrated its reliability in this study. The correlations between these two measures were significant beyond the .01 level, and it is therefore concluded that the SIT and the Shipley appear to be measuring the same type of intellectual functioning. Since the SIT has been found to be a valid instrument for measuring adult intelligence by virtue of its high correlations in comparison to establish proven measures of intelligence, it seems logical to conclude that the Shipley, by virtue of its significant correlations with the SIT in this study, also is a valid instrument for measuring adult intelligence.

REFERENCES

- Anastasi, A. Psychological Testing. New York: The Macmillan Company, 1969
- DeCecco, J. P. The Psychology of Learning and Instruction: Educational Psychology. Englewood Cliffs, N. J.: Prentice Hall, 1968
- DeLapa, G. The Slosson Intelligence Test: a screening and retesting technique for slow learners. Journal of School Psychology, 1968, 4 (3), 224-225.
- Garrett, Henry E. Slosson intelligence test (SIT) for children and adults. The Catholic School Journal, April 1965, 65(4).
- Hammill, D. The Slosson Intelligence Test as a quick estimate of mental ability. Journal of School Psychology, 1968-69, 7 (4), 33-37.
- Jongeward, P. A validity study of the Slosson Intelligence Test for use with educable mentally retarded students. Journal of School Psychology, 1968-69, 7 (4), 59-63.
- Kaufman, H., & Ivanoff, J. The Slosson Intelligence Test as a screening instrument with a rehabilitation population. Exceptional Children, 1969, 35 (9). (Reprint: New York: Slosson Educational Publications, 1969)
- Mc Rae, J. A comparison of the Wechsler Preschool and Primary Scale of Intelligence with the Slosson Intelligence Test. Unpublished Master's thesis, Eastern Washington State College, 1968.
- Poissant, P. A product moment correlation between the Stanford-Binet and the Slosson Intelligence Test with Publications, 1969. (Abstract)
- Rudolph, Linda B. Correlates of the Wechsler Adult Intelligence Scale, the Slosson Intelligence Test, ACT Scores and Grade Point Averages. Unpublished Master's Thesis, Austin Peay State University, 1970.

REFERENCES

Slosson, R. Slosson Intelligence Test (SIT) for children and adults. New York: Slosson Educational Publications, 1963.

Soft, L. Yellowstone Boys Ranch School, Route 1, Box 212, Billings, Montana. Personal communication to Donald Hammill. c.a. Hammill, D. The Slosson Intelligence Test as a quick estimate of mental ability. Journal of School Psychology, 1968-69, 7 (4), 33-37.