International Journal of Advanced Educational Research
ISSN: 2455-6157
Impact Factor: RJIF 5.12
www.educationjournal.org
Volume 3; Issue 1; January 2018; Page No. 173-175

Inter-scorer Reliability Estimation of SIS-II percepts

[1] Umed Singh, [2] Meenakshi
[1] Professor (Rtd), Department of Psychology K.U.K, Haryana, India
[2] Ph.D. Scholar, Department of Psychology, K.U.K, Haryana, India

Abstract
Somatic Inkblot Series-II (SIS-II) was developed as an improvised and comprehensive Inkblot Test over its earlier version as well as over other Inkblot Tests. It was developed in an attempt to make the Inkblot testing free from test anxiety, effect of gender of the examiner, for use in group setting with self administration mode, and to increase its research and clinical utility. The present study is an attempt to establish inter- scorer reliability coefficient for the 14 scores obtained from eleven categories of scales. For the realization of research purpose, 100 SIS-II protocols were randomly selected from the data collected by the investigator for her PhD research work. These protocols were independently scored by two scorers, one the investigator herself and the other who had the experience of administration, scoring and interpretation of one thousand SIS-II Protocols. Obtained data were analyzed with descriptive statistics to ascertain the normalcy of data, and Pearson's correlations to obtain the interscorer reliability coefficients. Obtained reliability coefficients have ranged from.408 to.883, almost comparable to those reported by authors of tests and some other earlier studies.

Keywords: somatic inkblot series, inter-scorer reliability, correlation

Introduction
Psychometric status of projective techniques has been the most controversial issue in the fields of personality assessment (Lilienfeld *et al*. 2000) [5]. Literature on projective techniques in variably exposes a reader to a striking paradox. On the one hand, during the past five decades, a good number of personality assessment researchers (e.g. Anastasi, 1982, Klein, 1986, Dawes, 1994) [1] have decried the reliability and validity of most of the projective techniques (Lilienfeld, 1999, Lilienfeld *et al*, 2000) [5, 10].

Jenson's (1965) [11] famous quotation, although 50 years old, still attracts the attention of contemporary researchers toward the inkblot tests and other projective techniques: "The rate of scientific progress in clinical psychology might well be measured by the speed and thoroughness with which it gets over's the Rorschach". On the other hand, clinicians across the globe have continued to use projective techniques regularly, and even many of them have asserted that without using inkblot and other projective techniques, it is not possible to practice clinical psychology (Watkins *et al*, 1995; Lilienfeld *et al*, 2000; carter *et al*, 2012; Piotrowski, 2017) [5, 6, 12]. The crux of this paradox has been summed up by Anastasi (1982) [1] observing that "Projective techniques present a discrepancy between research and practice. Evaluated as psychometric instruments, most of the projective tests make a poor show, yet in clinical practice their popularity is unabated". Despite such withering criticisms, projective techniques, as per numerous survey, findings, still enjoys wide spread popularity in clinical practice (Dawes, 1994, Piotrowski, 2017).

But Jensen (1965) [11] has opined in this regard with the assertion that if projective techniques are deficient in psychometric properties, fault lies with the clinical researchers not with the techniques themselves. He has further stated that as a result of the historically separate development of projective techniques outside the traditions of psychological measurement, this important of type testing has been practically ignored in a central concept of psychological test theory. Psychometric status of projective techniques has been seldom mentioned in the projective literature implies that the problem of psychometric qualification of projective techniques has never been taken seriously. Even the developers of projective techniques made fallacious argument that concepts of reliability and validity do not apply to projective tests (Jensen, 1965, Cronbach, 1949) [11, 13]. It is clear from the literature that early developers of projective techniques, particular inkblot techniques, viz. Frank, Hertz, Klopfer, Beck did not bother for the estimation of reliability and validity estimates of the Rorschach. To this issue first attention was paid by Holtzman, who attempted to overcome these psychometric deficiencies of the Rorschach, and develop his own inkblot technique popularly known as Holtzman Inkblot Techniques (HIT) in which he established reliability and validity estimates along with the development of objective scoring of inkblot percepts (Holtzman *et al*, 1961), Holtzman made genuine attempts while developing HIT to meet the technical psychometric requirements of a standardized test (Holtzman,1968,1975, Holtzman, *et al* 1961) [15]

The tradition of inkblot testing was further extended by Cassell by developing Somatic Inkblot Series (SIS) mainly to understand psychological problems through Somatization based on conceptualizations of body perception, body symbolism, somatic repression, and inner cry (Cassell, 1960, 1980, 1984; Cassell and Dubey, 2003). At present four forms of SIS are available: Cards Form (SIS-I); Booklet Form (SIS-

II, 62 Images); Video Forms. SIS-II- video); and Living Images Form. SIS-II Booklet Form (62 images) is considered as more improvised and comprehensive Somatic Inkblot Technique consisting A & B series, each with 31 images with a combined set of 62 images alongwith a sample image in the beginning. But in case of SIS also, no serious efforts have been made to establish the psychometric properties (reliability, validity) in the true sense in term of test theory.

Purpose of present study is to ascertain the reliability estimates of fourteen scales of SIS-II (Booklet Form).The concept of reliability and the methods of estimating it have always occupied a prominent place in psychometric theory. Reliability is a necessary and minimum qualification of a standardized test must have as is clear in Guilford's view that "Reliability is the minimum information one should have concerning a test" (Guilford, 1954) [14]. Reliability is an umbrella term under which different types of score stability are assessed implying that reliability index of a test score indicates its stability (Kline, 2005) [9]. It means stability of test score over time (test-retest), stability of item scores across items (internal consistency), or stability of rating across judges or raters of a person, object, event and so on (inter-rater reliability). These approaches are based on Classical Test Theory (CTT), which characterizes the reliability to be solely concerned with random measurement error. Classical Test Theory assumes that the raw score (X) on a test consists of a true component (T), and a random error component (E) i.e. X = T+E. implying that less the random error score more it represents the true score. In this perspective, formula for calculating the reliability is: Reliability= 1-[VAR-(E)/VAR(X)] meaning thereby, the reliability is equal to the ratio of random error variance to total score variance. In addition to the true and error score, there is also a domain sampling model of Classical Test Theory (CTT), which assumes that a test is made up of a sample parallel item from the universe of possible items. Thus, conceptually reliability is the degree of accuracy of a test in measuring whatever it measures; and statistically it is the proportion of non- error variance in test scores varying somewhere between 0 and 1 as estimated by reliability coefficient. Conceptually, it is an issue that what degree of reliability of a must have; yet practically psychometrician agree that a test individual or clinical whatever it is called in general must have a degree of reliability falling in the range of.70 to.90.

There are different ways of assessing reliability in terms of Classical Test Theory, namely, Test-retest, Alternate Form (Split-half), Internal Consistency, Reliability of a Composite, Difference Score Reliability and Inter-rater (inter-scorer) Reliability. The present study intends to estimate the inter-rater (inter-scorer) reliability of fourteen scale scores of SIS-II. Reliability index which describes consistency of judgments or ratings across individuals or raters/judges are called inter-rater (inter-scorer) reliability. There are many different types of inter-rater (inter-scorer) reliability index, and one to be used depends upon the type of data collected. Main types of inter-rater reliability estimates are: observer agreement percentage, inter observer correlations, Kendall's coefficient of concordance, Cohen's Kappa coefficient, and Average Deviation Index and Intra-Class correlation Coefficient. The present study has made use of inter observer correlation in the estimation of reliability coefficient of fourteen scales of SIS-II. In addition to CTT, Modern Test Theory (Item Response Theory; IRT), uses somewhat different index of reliability based on pattern of response to the test items and item parameters. IRT provides information for each item and with the test as a whole in terms of its accuracy in estimating an individual's trait level (Klein, 2005) [9].

Method

One hundred SIS-II protocols were selected randomly from ongoing research by the investigator. Each protocol was rated or scored by two independent scorers. One scorer was investigator herself and another was having the experience of administration, scoring, and interpretation of about 1000 SIS-II protocols, as she used it in her doctoral work. 100 protocols consisted of about 50 from male and 50 from female subjects. Obtained two sets of scores by two independent raters were analyzed by applying Descriptive Statistics including frequency distribution and Pearson's Correlations.

Results

First of all, Frequency Distributions were set up and Descriptive Statistics (mean, SD, skewness, kurtosis) were computed, to ascertain the normalcy of distribution of scores.

Table 1: Showing Inter-Scorer Reliability of SIS-II Scales

Variables	Scorer 1		Scorer 2		Inter Scorer r_{12}	t-Values	Sig.
	Mean	SD	Mean	SD			
R	58.95	2.11	59.61	1.94	.88	-2.31	.02
H	20.98	4.01	20.79	3.89	.70	.34	.73
A	8.37	3.26	8.33	3.00	.41	.09	.93
At	12.57	4.58	13.20	4.95	.68	-.93	.35
SEX	1.75	1.70	1.16	1.31	.71	2.76	.01*
M	8.63	3.81	9.60	4.19	.86	-1.71	.09
MT	12.81	1.91	12.21	2.17	.65	2.08	.04
T	21.38	5.37	20.32	5.14	.77	1.43	.16
ATT	24.82	5.99	27.09	5.92	.70	-2.70	.01*
REJ	3.01	2.03	2.39	1.92	.88	2.22	.03
PAS	.43	.74	.54	.79	.77	-1.01	.31
D	1.37	1.16	1.48	1.39	.50	-.61	.54
HAS	1.59	1.20	1.69	1.43	.80	-.53	.59
P	.72	.77	1.04	.96	.40	-2.60	.01*

Inspection of table -1 reveals that inter-scorer reliability coefficients for all the fourteen scores of SIS-II scales are fairly good ranging from.408 to.883.Respective reliability coefficients are.879 for Total No of Responses(R),.700for Human Response(H).408 for Animal Responses(A),.676for Anatomical Responses,(At),.706 for Sex Responses,.858 for Movement Responses(M),.652 for Most Typical Responses(MT),.770for Atypical Responses(AT),.883 for Rejection of Images(Rej),.765for Pathological Anatomy Responses(PAS),.503for Depression Responses(D),.801for Hostility& Aggression Responses(HAS),and.398 for Paranoia Responses(P).Obtained interscorer reliability estimates are almost comparable to those reported by Cassell and Dubey(1991,2003).The reliability coefficient for Animal and Paranoia scores are low, may be because of high subjectivity in the scoring procedure.The high reliability coefficient for Rejection of Images Scores (.883),Total NO of Responses(..879), and Movement (.858) are well understandable in terms of their objective scoring procedure.Inter- scorer reliability coefficients for all the fourteen scale scores of SIS-II are comparatively higher than the split- half reliability coefficients(obtained in the present study)are well in tune with the inference from the earlier studies that interscorer reliability in general is found higher than split-half reliability in projective tests, particularly in Inkblot tests(Lilienfeld etal,2000;Viglione etal,2012) [5, 16]. Early research findings have also demonstrated the interscorer reliability estimates of SIS-II scale scores to be high as compared to others reliability estimates of SIS-II and to the interscorer reliability estimates of other projective tests like Rorschach and TAT(Cassell and Dubey,2003) [4]. Comparison of value of mean and standard deviations of scores of all the fourteen variables by the two scores also testify the authenticity of obtained interscorer reliability coefficients, as the means and standard deviations in data from two scores are almost equal.Obtained t test-values also depict the same inference.

References

1. Anastasia A Psychological Testing.N; Y: MacMillan, 1982.
2. Cassell WA. *Body symbolism and the Somatic Inkblot Series*. Aurora Pvt. Co. Alaska, USA. Aurora Pvt. Co, 1980.
3. Cassell WA. *Somatic Inkblot Series Video.*SIS Center, Anchorage, Alaska, USA, 1984.
4. Cassell WA, Dubey BL. *Interpreting inner world through Somatic Imagery: Manual of the Somatic Ink blot Series*. Somatic Inkblot Centre, Anchorage, Alaska, USA, 2003.
5. Lilienfeld SO, Wood JM, Garb HN. The Scientific Status of Projective Techniques. Psychological Science in the Public Interest. 2000; 1(2):27-66.
6. Watkins CE, Campbell VL, Nieberding R, Hallmark P. Contemporary practice of psychological assessment by clinical psychologist. professional psychology; Research and Practice, 1995; 26:54-60.
7. Piotrowski C, Keller JW, Ogawa T. Projective Techniques: An international perspective. Psychological Report. 1993; 72:179-182.
8. Holtzman WH. New Development in Holtzman Inkblot Technique. In P.M. Reynolds (Ed). Advances in Psychological Assessment. 1975; 3:243-274.
9. Kline TJB. Psychological Testing: A Practical Approach To Design and Evaluation. Sage Publications, Inc, 2005.
10. Lilienfeld SO Projective Measures of Personality and psychopathology: How well do they work ? Skeptical Inquirer. 1999; 23:32-39.
11. Jensen AR A Review of Rorschach, In O.K.Buros(Ed), Sixth Mental Measurement, Handbook (501-509) Highland park, new hanspire; gryphon, 1965.
12. Carter NT, Daniels MA, Zickar MJ. Projective Testing: Historical Foundations and uses for Human Resources Management. Human Resources Management Review. 2012; 23:205-208.
13. Cronbach LJ. Statistical methods applied to Rorschach scores: A Review Psychological., bulletin. 1949; 46:393-429.
14. Guilford JP. Psychometric Methods Second edition, New York: MC Grow Hill, 1954.
15. Holtzman WH. Holtzman Inkblot Technique. In A.I Robin (Eds) Projective Techniques in personality Assessment (136-170). New York: Sprenger, 1968.
16. Vigilione. et. al. An Interscorer Reliability, study for the Rorschach Performance Assessment System. Journal of Personality Assessment. 2012; 94(6):607-612.