



# Modern Practices in Psychometric Test Construction and Validation: The Indian Scenario

**Datta S\***

St. Xavier's University, India

**\*Corresponding author:** Sumona Datta, Assistant Professor, St. Xavier's University, Kolkata, India, Email: [sumonadatta.psych@gmail.com](mailto:sumonadatta.psych@gmail.com)

**Editorial**

**Volume 7 Issue 3**

**Received Date:** August 25, 2022

**Published Date:** September 09, 2022

**DOI:** [10.23880/pprij-16000309](https://doi.org/10.23880/pprij-16000309)

## Abstract

Psychometric test construction and validation has come a long way in the past few decades. With the development of modern approaches to testing and applications of statistical softwares, this field of research has become highly sophisticated and enriched. The Indian scenario is however, a little different, pertaining to the acceptability of modern psychometric methods and their use in test development and validation. Although several new tests and questionnaires are being developed or validated in local cultural contexts, Indian researchers are yet to reach the state where psychometric test development and validation becomes a usual practice. Building a common platform for exchange of knowledge and expertise in psychometrics is highly recommended in order to facilitate such practices at a greater extent.

**Keywords:** Psychometric Test construction; Test Validation; Indian context; Item Response Theory; Modern Psychometric Techniques

Almost 150 years back, Francis Galton had defined "Psychometry" as "the art of imposing measurement and number upon operations of the mind" [1]. Undoubtedly, measurement and quantification are omnipresent in all branches of social and behavioural sciences. Nehra, Sarkar & Mahajan [2] stated that "psychometry involves research related to the construction of instruments and procedures for measurement, and the development and refinement of theoretical approaches to mental measurements" (p. 33). In a broader sense, psychometry encompasses the assessment of all observable behaviours and latent psychological constructs that describes one's attitude, personality, cognition, intelligence, beliefs, interests, creativity and dysfunctional behavioural outcomes. This not only includes psychometric testing but also the techniques of such test construction and validation. Psychometric models currently comprise of all statistical techniques involved in all sort of psychological assessments ranging from handling missing

data, validating newly developed test, adapting a test, combining secondary information with new data, construct validation of psychological constructs and many more.

The past couple of decades witnessed a humungous change in psychometric methodology and the ways in which data are analysed and interpreted. Earlier, testing and measurement techniques followed classical theories and use of basic statistical methods. However, the advent of more sophisticated statistical softwares like STATA, MPLUS, R Studio, JAMOVI etc. has revolutionized the field of psychometric test development and validation. One instance can be the development of different polytomous IRT models for treating scale data [3-5]. Although Rasch had developed the Rasch Model during 1960s [6] and subsequent 2-PL and 3-PL models for both dichotomous and polytomous data were developed, the use of the models for test validation became popular only in the recent times. This trend can be

noticed in the validation processes employed for the widely popular tests for different cultures and communities. For instance, Makhubela & Mashegoane [7] validated the Beck Depression Inventory II for a South African population by checking its factorial validity and longitudinal measurement invariance. More recent researches reporting the construction and validation of tests are employing modern techniques in psychometrics including the application of Item Response Theory, methods of enquiring Differential Item Functioning (DIF) and testing invariance among items, even for translations and adaptations. Peterson et al. [8] tested the applicability of DIF technique in test validation and found it to be a useful technique. Walker [9] also, stressed the importance of testing for DIF in test validation process. The past one decade witnessed the application of DIF in test validation process to a great extent. For instance, Pilatti, Read & Caneto [10] checked the psychometric properties of the Spanish version of the Young Adult Alcohol Consequences Questionnaire using DIF. Cascella & Pampaka [11] constructed and validated a measure of attitude towards gender roles in family using a polytomous IRT Partial Credit Model (PCM) and by checking DIF and item fit statistics. Datta & Dutta Roy [12] constructed and validated a new measure of mental rotation ability using a dichotomous 2PL IRT Model and Confirmatory Factor Analysis. Needless to say, the field of psychometric testing has metamorphosed from the application of old school classical test theory approaches during the past century to the employment of sophisticated statistical and machine intelligence techniques and statistical soft wares.

### Indian Scenario

A review of the existing literature on test construction in the Indian context revealed two approaches: (a) construction and validation of a new measure and (b) adaptation/translation/shortening of an existing tool.

Verma [13] enlisted a number of drawbacks which necessitates the need for constructing or validating new or existing measures respectively in the Indian context. Major issues that emerged from his work were (a) On-spot translation of English scales into local vernacular language; (b) Inability of test takers to understand and respond due to lack of literacy; (c) Lack of local norms; (d) Too long questionnaires with complicated words difficult to be interpreted by test takers; (e) Unavailability of translated questionnaires; (f) Lack of revision of scoring systems and norms; (g) Lack of acceptability of test construction work as a worthwhile research; and (h) Lack of expertise in guiding a test construction and validation work (p. 17).

Practices such as on-spot translation, test administration despite the inability of test takers to respond and not using

local norms for test score interpretation, vehemently challenges the inferences drawn from an assessment, especially if it is high stakes testing. Another type of complication concerns the unavailability of proper norms, and translated/adapted versions. The Indian constitution includes 22 languages while Census data reveals more than 19,500 different languages spoken across the country [14]. However, none of these languages include any version of English. Hence it is a technically challenging to administer tests developed in English language.

Additionally, cultural inequity is another concern when it comes to testing in Indian context. Geert Hofstede (p. 3) [15] defines culture as “the collective programming of the mind that distinguishes the members of one group or category of people from others”. For the test scores to be reliable and valid, it is important that the reactions of the responders are stimulated or elicited by the same psychological attribute or context. Hence, any differences with respect to one’s thought processes, perceived impression, individual differences and social practices, might impact the validity of the scores. Indian society follows a collectivist approach and therefore the perception of events, social practices and interpersonal interactions are substantially different from that of the western societies. Such differences are often reflected in the test items included in the tests developed in the western countries. For instance, personality assessment is one of the major areas where cultural adaptation becomes quite significant. Projective tests like Thematic Apperception Test [16] card pictures are highly westernized and hence often subjects find it difficult to connect with the stories. This undoubtedly impacts the quality of response obtained. Similar issues arise when it comes to assessment of attributes like attitude, self-esteem, interpersonal relationships, adjustment levels and others.

The third issue concerns the lack of local norms and local translations. These days, psychometricians around the world communicate among themselves to produce translated versions of the same scale or questionnaire. For instance, Peters et al. [17] developed a redrawn version of the Vandenberg & Kuse Mental Rotation Test [18] and translated it to 25 different languages worldwide by collaborating with researchers around the world. However, the only Indian language in which it was translated is Hindi and hence the applicability stayed limited. Verma [13] (2015) points out that even the Hindi versions of scales like Hindi Personality Trait Inventory [19] contains words that are complicated, for example, “*ashrupat*” denoting ‘*shedding tears from the eyes*’. Nonetheless, there are many scales and questionnaires that were translated into Hindi and hence are applicable in most parts of northern India. Only a few tests have been translated in South Indian languages or have been adapted in the local culture. Notable among those is the NIMHANS

Neuropsychological Test Battery [20]. Researchers and professionals from other areas of India still rely on on-spot translations and unstandardized translations of tests and questionnaires in English language.

Finally, Verma [13] pointed out another extremely important issue that is, the lack of acceptability of the practice of developing psychometric tests and lack of expertise in guiding the same. Until late 1980s in India, Ph.D. in test construction was not considered to be sufficient for doctoral thesis level. Moreover, most researchers and academicians lacked the skills and expertise to develop and validate psychometric tests. This unacceptability and lack of expertise in psychometric test development and validation might be because: (a) Lack of understanding of modern psychometric techniques; (b) Lack of expertise in using statistical softwares; and (c) Limited knowledge about test equating and adaptation techniques. Till today, the psychological fraternity of India lacks an association or society dedicated to discussions pertaining to psychometric test construction and validation.

Nevertheless, the past couple of decades witnessed a number of tests developed and/or translated/adapted for Indian population. These include a wide range of tests applicable in various contexts. These included: (a) cognitive tests like Bhatia's battery of performance tests of intelligence [21], WAIS-R verbal scales in Hindi [22], PGI Battery of Brain Dysfunctions [23], Mental Rotation test (Datta & Dutta Roy, 2021), PGI Locus of Control scale [24], Hindi version of Montreal Cognitive Assessments [25]; (b) Projective tests like Sack's Sentence Completion test in Hindi [26], Indian adaptation of Thematic Perception Test [27], Indian adaptation of Resenzweig Picture-frustration test [28], Indian adaptation of Rorschach Inkblot Test, Semantic Inkblot Series [29]; (c) Other scales like Hindi version of Eysenck Personality questionnaire, Hindi Self-report Altruism Scale [30], Hindi adaptation of Gratitude Questionnaire [31], Value Conflict Scale [32-34], Learned Helplessness scale [35], Home Environment Scale [36], Self-efficacy scale [37], Hindi version of MMSE [38], Emotion Regulation Scale [39] and many more.

### Going Forward

In the past few decades, Indian researchers and scholars have contributed significantly in the field of psychometric testing in Indian context. This included developing and validating new tests/questionnaires, translating and adapting tests/questionnaires and also shortening tests/questionnaires. Local norms have been developed for a number of these tests and applied in different clinical and research settings.

Nonetheless, there is still a lack of acceptability of psychometric test development as a regular practice among Indian researchers. This is a dire need of an initiation of discourse among academicians and researchers regarding the advanced techniques of psychometric test development and validation. A common platform where scholars and psychometricians can share their knowledge and concerns, will promote higher acceptability among them for such practices. In a nut shell, there is still miles to go for Indian researchers and academicians to make a niche in the field of psychometry at the global level.

### References

1. Galton F (1879) Psychometric experiments. *Brain* 2(2): 149-162.
2. Nehra R, Sarkar S, Mahajan S (2015) Psychometry as an Adjunct to Psychiatry: An Indian Perspective. In: *Developments in Psychiatry in India*. Springer, New Delhi, pp: 33-49
3. Samejima F (1969) Estimation of latent ability using a response pattern of graded scores *Psychometrika* 34: 1-97.
4. Samejima F (1972) A general model for free-response data. *Psychometrika* 37(1): 68.
5. Muraki E (1992a) A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement* 16(2): 159-176.
6. Rasch G (1960) Probabilistic models for some intelligence and attainment tests. Copenhagen: Nielsen & Lydiche.
7. Makhubela MS, Mashegoane S (2016) Validation of the Beck Depression Inventory-II in South Africa: Factorial validity and longitudinal measurement invariance in university students. *South African Journal of Psychology* 46(2): 203-217.
8. Petersen MA, Groenvold M, Bjorner JB, Aaronson N, Conroy T, et al. (2003) Use of differential item functioning analysis to assess the equivalence of translations of a questionnaire. *Quality of Life Research* 12(4): 373-385.
9. Walker CM (2011) What's the DIF? Why differential item functioning analyses are an important part of instrument development and validation. *Journal of Psychoeducational Assessment* 29(4): 364-376.
10. Pilatti A, Read JP, Caneto F (2016) Validation of the Spanish version of the young adult alcohol consequences questionnaire (S-YAACQ). *Psychological assessment* 28(5): e49.

11. Cascella C, Pampaka M (2020) Attitudes towards gender roles in family: A Rasch-based validation study. *Journal of applied measurement*.
12. Datta S, Roy DD (2021) Development and Validation of a New Measure of Mental Rotation for Preadolescent and Adolescent Groups. *Journal of Cognitive Education and Psychology* 20(1): 18-37.
13. Verma SK (2015) Development of psychological testing in India. In: *Developments in Psychiatry in India*. Springer, New Delhi, pp: 15-32.
14. The Indian Express (2022) More than 19,500 mother tongues spoken in India: Census.
15. Hofstede G (2011) Dimensionalizing cultures: The Hofstede model in context. *Online readings in psychology and culture* 2(1): 2307-0919.
16. Murray HA (1943) Thematic apperception test.
17. Peters M, Laeng B, Latham K, Jackson M, Zaiyouna R, et al. (1995) A redrawn Vandenberg and Kuse mental rotations test: different versions and factors that affect performance. *Brain Cogn* 28(1): 39-58.
18. Vandenberg SG, Kuse AR (1978) Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and motor skills* 47(2): 599-604.
19. Verma SK, Pershad D, Mahajan A (1990) Hindi translation of Personality Trait Inventory. *The Creative Psychologist* 2(1): 23-26.
20. Tripathi R. Unpublished Ph.D. Thesis, Department of Clinical Psychology. Bangalore, India: NIMHANS (Deemed University); 2012. Development and standardization of neuropsychological test battery for older adults.
21. Bhatia CM (1955) Performance intelligence test battery. Bombay: Oxford University Press.
22. Pershad D, Verma SK, Randhawa A (1980) Adaptation of the Verbal scale of the WAIS-R. *Psychological Research Journal*.
23. Pershad D, Verma SK (1990) PGI battery of brain dysfunction. Agra Psychological Corporation: Agra.
24. Menon DK, Wig NN, Verma SK, Jain S (1975) PGI Locus of Control Scale: Preliminary data. *Indian Journal of Clinical Psychology*.
25. Gupta M, Gupta V, Buckshee RN, Sharma V (2019) Validity and reliability of Hindi translated version of Montreal cognitive assessment in older adults. *Asian Journal of Psychiatry* 45: 125-128.
26. Verma SK, Khanna BC, Wig NN (1985) Construction of PGI sentence completion tests (in Hindi). *Mind* 11: 7-14.
27. Choudhary U (1974) Thematic Apperception Test. Indian Adaptation of Murray HA (1943) Thematic Apperception Test, USA: Harvard College.
28. Roy SK (1969) Manual of the Indian Adaptation of the Rosenzweig Picture-Frustration Study (Adult Form).
29. Dwivedi CB, Mishra M, Dubey BL (1995) Diagnostic Significance of the Indices of the SIS - II and the Rorschach Inkblot Tests. *SIS J of Proj Psy & Ment Health* 2: 165-170.
30. Khanna R, Singh P, Rushton JP (1993) Development of the Hindi version of a self-report altruism scale. *Personality and Individual differences* 14(1): 267-270.
31. Dixit SK, Sinha J (2021) Adaptation and validation of the gratitude questionnaire (GQ-6) for the Indian context. *Current Psychology*, pp: 1-11.
32. Bharadwaj RL (1999) The development of a value-conflicts scale. *Indian Journal of Psychometry & Education* 30: 79-82.
33. Bharadwaj RL (1999) Value-Conflicts Scale. Agra: Pankaj Mapan.
34. Bharadwaj RL (2017) Revised Manual for Value-Conflicts Scale. Agra: Pankaj Mapan.
35. Dhar U, Kohli S, Dhar S (1987) Manual for the Learned Helplessness Scale (LH Scale), Bhiwani, Haryana Council of Psychological Researches.
36. Mishra KS (1989) A Home Environment Inventory. Ankur Psychology Agency, Lucknow, pp: 7.
37. Shrivastava P, Sonekar L, Janghel G (2018) Development of psychometric properties in self-efficacy scale in Hindi language (SE scale in Hindi version). *International Journal of Indian Psychology* 6(2): 50-55.
38. Ganguli M, Ratcliff G, Chandra V, Sharma S, Gilby J, et al. (1995) A Hindi version of the MMSE: the development of a cognitive screening instrument for a largely illiterate rural elderly population in India. *International Journal of Geriatric Psychiatry* 10(5): 367-377.
39. Bhatnagar P, Shukla M, Pandey R (2020) Validating the factor structure of the hindi version of the difficulties in Emotion Regulation Scale. *Journal of Psychopathology and Behavioral Assessment* 42(2): 377-396.

