The Importance of Validity and Reliability of Psychological Testing Instruments

The two most important features of psychological testing instruments that are used to conclusively prove that the acquired results are correct are reliability and validity. Reliability indicates the consistency of a measured value across different tested items, different investigators, and various time-points. On the other hand, validity refers to the degree to which the measured values represent the variables that were the basis of the research hypothesis. Validity may be determined by considering face validity, content validity, criterion validity, and discriminant validity. Therefore, any psychologist first establishes the reliability and validity of a testing instrument before proceeding to use it for a study (Cook and Beckman, 2006). This paper explores the main reasons that a testing instrument needs to be assessed for validity and reliability, and also discusses the effects on the results of a study if this assessment is not done properly.

Mohajan (2017) has described reliability as ‘stability’ and validity as ‘truthfulness’ of research findings. These two parameters need to be used to introduce research settings in a psychological study. These features increase transparency of research findings and prevent researcher bias in qualitative studies. They also set a strong foundation for analysing scores from psychometric instruments such as questionnaires, symptom scales, and observer ratings. They not only enhance the accuracy of a research study, but also enable the extrapolation of data to a theoretical framework (Mohajan, 2017).

Assessment of validity and reliability of testing instruments are more important for psychological instruments rather than, for instance, physical parameters. This is because physical measurements such as height, weight, length, and temperature already have established global measurement standards. In contrast, it is difficult to establish such measurement standards for psychometric assessments due to the abstractness of values obtained from measures of clinical symptoms and psychological indicators. As these measures are not as tangible as physical measures, they may turn out to be inaccurate and misleading, thereby compromising the quality of research findings. As the field of psychology largely lacks ‘gold standards’ for measures, both

validity and reliability are extremely important to determine the suitability of a psychological testing instrument (Bannigan and Watson, 2009).

The assessment of testing instruments for measuring psychological distress and other indicators of mental health problems is extremely important before proceeding with any study based on the characteristic features of reliability and validity. One reason is that the instruments that may be ideal for a particular population may not be suitable for another thereby making them invalid for a study. This is true for instruments that are developed in one country or for one specific population group, and which may not be suitable for people living in a different cultural, social, or financial context. The primary explanation for this is that testing instruments are developed to assess symptoms and parameters that are present in the local population. These symptoms may be absent in another population and therefore, the testing instrument that may be perfect for one population may not be valid and reliable for another population. Using testing instruments without assessing them for validity and reliability for the specific population group can lead to misrepresentation of key clinical indicators and erroneous conclusions with respect to the study objectives (Hall et al., 2015).

A study by Catalan and Gordon (2020) has explored the importance of reliability and validity of testing instruments in poverty measurement studies. Parameters such as poverty need to be accurately assessed because the results of such studies are used for informing policies and programs. There are several different instruments to measure poverty and each of them is based on different definitions, assumptions, indicators, thresholds, weights, and dimensions of poverty. It is, therefore, extremely essential to evaluate the reliability and validity of the testing instrument and see if it is relevant and suitable to the study objectives (Catalan and Gordon, 2020).

If the assessment of reliability and validity of testing instruments is not done properly, it may result in errors in the measurement process, thereby compromising the accuracy of the study findings. It is the duty of every researcher to evaluate the potential sources of errors in measurement, otherwise it may interfere with interpretation of the values obtained. A further complication arises when qualitative attributes of a psychological study need to be quantified. In such a scenario, if testing instruments

used to measure qualitative data are not tested for validity and reliability, there may be inconsistency in quantitative extrapolations of the research findings (Kimberlin and Winterstein, 2008).

A study conducted by Shirali et al. (2018) has thrown light on the ramifications of not evaluating the reliability and validity of testing instruments for measuring a multi-dimensional parameter, which is a safety management system. According to them, most testing instruments focus on only one dimension of a safety management system, whereas safety is a multi-dimensional construct. Therefore, measuring the level of safety in a system using instruments that have low validity and reliability can result in a false value of safety levels being obtained in the study. Also, the safety level of a system is largely context-dependent and testing instruments need to be assessed for suitability and relevance to the specific study parameters as well. Another parameter is workplace hazards, which if not measured using valid and reliable testing instruments can provide false results thereby affecting the quality of the work environment (Shirali et al., 2018).

In conclusion, the importance of evaluating the validity and reliability of testing instruments for psychometric assessments cannot be overstated. This is because the testing instruments that may be suitable for a particular system or population may not be relevant to another system or population. For instance, measures of psychological symptoms of distress that may be present in a particular cultural group may be context-dependent and may not apply to another population. Therefore, it is extremely important to assess and present the results of validity and reliability assessments of psychological testing instruments so that the results of the study findings may be inferred accurately and extrapolated at a global level.

References

Bannigan, K., & Watson, R. (2009). Reliability and validity in a nutshell. *Journal of Clinical Nursing, 18*(23), 3237-3243. doi:10.1111/j.1365-2702.2009.02939.x

Catalán, H. E., & Gordon, D. (2019). The Importance of Reliability and Construct Validity in Multidimensional Poverty Measurement: An Illustration Using the Multidimensional Poverty Index for Latin America (MPI-LA). *The Journal of Development Studies, 56*(9), 1763-1783. doi:10.1080/00220388.2019.1663176

Cook, D. A., & Beckman, T. J. (2006). Current Concepts in Validity and Reliability for Psychometric Instruments: Theory and Application. *The American Journal of Medicine, 119*(2). doi:10.1016/j.amjmed.2005.10.036

Hall, B. J., Puffer, E., Murray, L. K., Ismael, A., Bass, J. K., Sim, A., & Bolton, P. A. (2014). The Importance of Establishing Reliability and Validity of Assessment Instruments for Mental Health Problems: An Example from Somali Children and Adolescents Living in Three Refugee Camps in Ethiopia. *Psychological Injury and Law, 7*(2), 153-164. doi:10.1007/s12207-014-9188-9

Kimberlin, C. L., & Winterstein, A. G. (2008). Validity and reliability of measurement instruments used in research. *American Journal of Health-System Pharmacy, 65*(23), 2276-2284. doi:10.2146/ajhp070364

Mohajan, H. K. (2017). Two Criteria For Good Measurements In Research: Validity And Reliability. *Annals of Spiru Haret University. Economic Series, 17*(4), 59-82. doi:10.26458/1746

Shirali, G., Shekari, M., & Angali, K. A. (2018). Assessing Reliability and Validity of an Instrument for Measuring Resilience Safety Culture in Sociotechnical Systems. *Safety and Health at Work, 9*(3), 296-307. doi:10.1016/j.shaw.2017.07.010