



International Journal of Humanities & Social Science Studies (IJHSSS)
A Peer-Reviewed Bi-monthly Bi-lingual Research Journal
ISSN: 2349-6959 (Online), ISSN: 2349-6711 (Print)
Volume-III, Issue-III, November 2016, Page No. 100-107
Published by Scholar Publications, Karimganj, Assam, India, 788711
Website: <http://www.ijhsss.com>

An Assessment of the Reliability and Validity of the Servqual Scale in the Higher Education Context of Tanzania

Prof. Ahmed M. Ame

School of Business Studies and Economics, University of Dodoma, Tanzania

Dr. Amani G. Tegambwage

School of Business Studies and Economics, University of Dodoma, Tanzania

Abstract

Although service quality has received increasing attention in higher education, there is lack of a standardized instrument to measure it. The most popular and widely used instrument to measure service quality across industries is the SERVQUAL scale. However, since the SERVQUAL scale was developed in the retail commercial sector, its reliability and validity for measuring service quality in the higher education sector has been questioned. This is because higher education sector differs significantly from the retail commercial sector where the SERVQUAL scale was developed and widely applied. This paper assesses the applicability of the SERVQUAL scale for measuring service quality in higher education, particularly in the context of Tanzania. The data were obtained from 500 students in two public universities in Tanzania. The study found out that, the SERVQUAL scale, with appropriate modifications, is reliable and valid for measuring service quality in the higher education sector.

Key Words: Reliability, Validity, Higher education, SERVQUAL

1.0 Introduction: In order to achieve superiority in service quality and realize its benefits, higher education institutions must have the means to measure it. One of the most valid and reliable measurement instruments used in the evaluation of service quality across service sectors is the SERVQUAL scale. Despite its wide application across service sectors, there have been arguments about its applicability to all service sectors, that is, whether the instrument is universal or not (Atrek & Bayraktaroglu, 2012). Since the SERVQUAL scale was developed in the retail industry, it may not be able to measure service quality in different service sectors adequately. The reason is that, each service sector has unique and distinguished characteristics (Soliman & Alzaid, 2002).

Thus, the ability of SERVQUAL to measure service quality in higher education is unclear (Hoe, 2005). This is because higher education differs from the services used in the development of SERVQUAL (the retail commercial sector) in terms of: (1) the nature of the service act, (2) the length and nature of relationship with the client, (3) levels of

customization and judgement, (4) nature of demand relative to supply, and (5) method of service delivery. Therefore, the ability of this scale to measure service quality in higher education service sector requires further investigation.

Few studies have tested the applicability of the SERVQUAL scale in the higher education sector (Brochado & Marques, 2009; Hoe, 2005). Though the attempt has been made to extend the knowledge to higher education sector, these studies were conducted in the developed countries which are likely to differ from the developing countries in terms of the number and composition of service quality dimensions (Palmer, 2011). Studies have shown that customer expectations, values, beliefs and perceptions change from one culture to another and from one country to another (Lee, 2007). Furthermore, there appears to be a positive relationship between economic development of a country and its service sector; developed economies being increasingly more service-orientated than developing economies (Palmer, 2011). Thus, standardized instruments developed and tested in the west may miss attributes important in developing countries. These two studies, in addition, have utilized 22 items of the original SERVQUAL scale.

The main objective of this study was to assess the reliability and validity of the SERVQUAL scale for measuring service quality in higher education, particularly in the context of Tanzania.

2.0 Theoretical Base

2.1 Reliability of Measurement Instruments: Reliability is defined as the extent to which a questionnaire, test, observation or any measurement procedure produces the same results on repeated trials (Cook & Beckman, 2006). In other words; it is the stability or consistency of scores over time or across raters. There are three aspects of reliability, namely: equivalence, stability and internal consistency (Creswell, 2008).

The equivalence aspect considers how much error may get introduced by different investigators or different samples of the items being studied (Patten, 2005). Equivalence is measured through a parallel forms procedure or by assessing interrater reliability (Creswell, 2008). The stability aspect of reliability is concerned with securing consistent results with repeated measurements of the same researcher and with the same instrument (Patten, 2005). Stability is assessed by comparing results of repeated measurements (Polit & Beck, 2004). The internal consistency aspect concerns the extent to which items on the instrument are measuring the same thing (Patten, 2005). The appeal of an internal consistency index of reliability is that, it is estimated after only one test administration, and therefore avoids the problems associated with testing over multiple time periods (Creswell, 2008). Internal consistency is estimated via the split-half reliability index, coefficient alpha index or the Kuder-Richardson formula 20 (KR-20) index. However, the widely employed method is coefficient alpha index (Creswell, 2008).

In this study, the internal consistency aspect of reliability was investigated due to the following reasons: (1) It is the most important aspect of reliability in connection with the multiple item scale (Cook & Beckman, 2006) like the ones employed in this study, (2)

Since the present study adopts a cross-sectional strategy, internal consistency reliability is feasible because it is estimated after only one test administration (Creswell, 2008), (3) The process of validating measurement instruments requires among other things, an investigation of internal consistency reliability (Nunnally & Bernstein, 1994).

2.2 Validity of Measurement Instruments: Validity is defined as the extent to which the instrument measures what it purports to measure (Cook & Beckman, 2006). This implies that, the results of any psychometric assessment have a meaning (validity) only in the context of the construct they purport to assess. Validity is not a property of the instrument, but of the instrument's scores and their interpretations (Cook & Beckman, 2006). Thus, validity must be established for each intended interpretation.

Scholars have distinguished a number of kinds of statements about the validity of a measure, including: (a) content validity which refers to the extent to which the measure adequately samples the content of the domain that constitutes the construct; (b) criterion validity which refers to the extent to which a measure is empirically associated with relevant criterion variables, which may either be assessed at the same time (concurrent validity), in the future (predictive validity), or in the past (postdictive validity), and (c) construct validity, an overarching term now seen by most to encompass all forms of validity, which refers to the extent to which a measure adequately assesses the construct it purports to assess (Nunnally & Bernstein, 1994). However, contemporary thinking on the subject suggests that, these distinctions are arbitrary and that all validity should be conceptualized under one overarching framework, "construct validity" (Cook & Beckman, 2006). This approach underscores the reasoning that an instrument's scores are only useful in as much as they reflect a construct and that, evidence should be collected to support this relationship. According to this conceptualization, the distinct concepts of content and criterion validity are preserved as sources of validity evidence within the construct validity rubric.

Construct validation is always theory dependent (Cook & Beckman, 2006). That is, a statement about the validity of an instrument is a statement about the extent to which its observed associations with measures of other variables, match theoretical predictions about how it should be associated with those variables. In fact, if the theory is wrong, the pattern of correlations will appear to invalidate the measure. The aim of construct validation basically is to establish its relation to other variables with which it should theoretically be associated positively, negatively, or practically not at all (Nunnally & Bernstein, 1994). Thus, construct validity is typically established by presenting correlations between a measure of a construct and a number of other measures that should, theoretically, be associated with it (convergent validity) or vary independently of it (discriminant validity) (Cook & Beckman, 2006). In addition, when assessing simultaneously convergent and discriminant validity, one confirms the presence of construct validity (Cook & Beckman, 2006). This implies that the best way to investigate construct validity is to check convergent and discriminant validity altogether (Ame, 2005). This study has adopted this approach.

3.0 Methods

3.1 Data Collection: The study employed the SERVQUAL scale, as modified to fit the higher education setting (Tegambwage & Ame, 2016), to collect data. The survey was conducted in two purposively selected public universities in Tanzania. Respondents (students) were systematically selected and a total of 250 students from various degree programmes in each of the two universities were selected for participation in this study. The questionnaire took about 20 minutes to complete. It is important to note that the names of the universities under study have not been mentioned in connection to the data collected because it was agreed as a condition during data collection.

3.2 Data Analysis: To assess reliability of the SERVQUAL scale for measuring service quality in higher education, internal consistency reliability (Cronbach's alpha coefficients) of the scale was investigated using SPSS (version 19.0). Cronbach's alpha, which is based on the internal consistency of a scale, is a commonly used index for a scale's reliability (Tavakol & Dennick, 2011). It measures the extent of commonality among the items (Cohen & Swerdlik, 2010), suggesting that, if the scale's items are measuring the same construct, the items will be positively correlated with each other. The value of this coefficient ranges from 0 to 1 implying that, the higher the alpha coefficient, the higher the internal reliability. However, a value of 0.70 or greater indicates good scale reliability (Tavakol & Dennick, 2011).

Specifically, the Cronbach's alpha coefficients of the SERVQUAL scale were computed both at the dimension level (to measure the internal consistency of individual items within each dimension) and at the overall scale level (to measure the commonality of the dimensions to the perceived service quality construct), across all the three data sets: university A, B, and the combined sample. Factor analysis was also conducted to assess the dimensionality of the modified SERVQUAL scale by utilizing the factor module of SPSS version 19.0 (Thompson, 2004).

To evaluate validity of the SERVQUAL scale for measuring service quality in higher education, the two most widely accepted forms of validity - convergent and discriminant validity were assessed (Cook & Beckman, 2006). Specifically, the degree of convergent and discriminant validity of the SERVQUAL scale was assessed using SPSS (version 19.0) by computing the average correlations between students perceived service quality (as measured by the modified SERVQUAL scale) and students' responses to measures of conceptually related variables: the overall service quality; the overall satisfaction; and the recommend intentions (Hair et al., 2005; Parasuraman et al., 1993).

Convergent validity assesses the degree to which two measures of the same concept are correlated (Hair et al., 2005). For the purpose of this study, convergent validity was examined by computing the average correlations between service quality (as measured by the modified SERVQUAL scale) and the directly measured overall service quality. Discriminant validity, on the other hand, gauges the extent to which measures of two different constructs are comparatively distinctive from each other (Hair et al., 2005). In this

study, discriminant validity was examined by correlating service quality (as measured by the modified SERVQUAL scale) with conceptually distinct measures, namely; “the overall satisfaction” and “the intention to recommend the university to a friend” (Hair et al., 2005; Parasuraman et al., 1993).

It is contended that, for the evidence of discriminant validity, the correlation between two different measures of the same variable must be higher than correlation between the measure of that variable and those of any other variable (Hair et al., 2005). The values of the Pearson correlation coefficient ranges from 0 to 1, and the higher the value the greater is the correlation between the variables (Cook & Beckman, 2006).

4.0 Results and Discussion

4.1 Reliability of SERVQUAL Scale in Higher Education: The results shown in Table 1 indicate high internal consistency among items within each dimension, ranging from 0.74 (tangibles) to 0.96 (non-tangibles), all are above the recommended minimum score of 0.70 (Tavakol & Dennick, 2011). Non-tangibles and tangibles are two factors exhibited by the modified SERVQUAL scale after factor analysis. The findings support the internal cohesiveness of the SERVQUAL items forming each dimension, and are comparable to those reported by Parasuraman et al. (1991) that ranged from 0.80 (tangibles) to 0.93 (responsiveness).

Table-1: Reliability Coefficients

Dimension	Number of Items	Reliability Coefficients			
		University A	University B	Combined Sample	Average Values
Non-tangibles	29	0.94	0.96	0.93	0.94
Tangibles	16	0.91	0.83	0.74	0.83
Overall Scale	45	0.95	0.94	0.92	0.94

Note: Non-tangibles and tangibles are two factors exhibited by the modified SERVQUAL scale

At the overall scale level, the Cronbach’s alpha reliability coefficients for the SERVQUAL scale ranged from 0.92 to 0.95 across all the three data sets (university A, B, and the combined sample) examined in this study. These coefficients are comparable to those obtained in other studies that ranged from 0.76 to 0.97 (Brochado & Marques, 2009; Vanpariya & Ganguly, 2010). Overall, the SERVQUAL scale and its dimensions have high internal consistency suggesting that, the items used to measure the constructs are appropriate. The results provide support for the use of the modified SERVQUAL items to examine perceived service quality in higher education.

4.2 Validity of SERVQUAL Scale in Higher Education: The presence of a positive and significant correlation coefficients (p= 0.01) between service quality (as measured by the SERVQUAL scale) and the overall service quality rating, across all the three data sets (university A, B, and the combined sample) is an indication of the convergent validity of the SERVQUAL scale (the average values was 0.52), as depicted in Table 2.

Table-2: Correlation Analysis Results

Measurement Scale	Standardized Correlation Coefficients			
	University A	University B	Combined Sample	Average Values
OSQ	0.23**	0.71**	0.61**	0.52**
Satisfaction	0.15**	0.70**	0.61**	0.49**
Recommend Intentions	0.10*	0.61**	0.56**	0.42**

Note: OSQ stands for Overall Service Quality.

** . Correlation is significant at the 0.01 level (2-tailed)

* . Correlation is significant at the 0.05 level (2-tailed)

On the other hand, findings have shown that, the correlations between perceived service quality and the overall service quality were consistently higher across all the three data sets, compared to the correlations between perceived service quality and the conceptually related but distinct variables: customer satisfaction and recommend intentions. For example, the average correlation coefficient between perceived service quality and the overall service quality was 0.52 and that between perceived service quality and other variables: customer satisfaction and recommend intentions were 0.49 and 0.42 respectively, as depicted in Table 2. The findings imply the presence of discriminant validity in respect of the SERVQUAL scale. The results are quite in conformity with those established by other scholars (e.g., Brochado & Marques, 2009; Hoe, 2005). These researchers have reported that, the SERVQUAL measure has strong correlations with the overall service quality, customer satisfaction, and behavioural intentions.

5.0 Conclusion and Recommendations: Based on the study findings, it can be concluded that, the SERVQUAL scale, as modified for this study, is reliable and valid for measuring service quality in the higher education setting in Tanzania. This gives the implication that, the SERVQUAL scale is applicable across industries and cultures provided that, it is appropriately customized to capture the service attributes of the industry and context in which it is applied. This study, therefore, recommends that, appropriate modifications must be made to SERVQUAL before applying it to a new setting.

6.0 Acknowledgements: We would like to extend our sincere gratitude to our employer the University of Dodoma for providing funding to conduct this study. Furthermore, we would like to take this opportunity to thank the management of the two Universities for giving us access to conduct the research in their institutions without which the study would be a nightmare. Last but not least, we thank, Ms. Khalila Shariff for going through this manuscript.

7.0 References:

1. Ame, A.M. (2005). *The effect of quality on satisfaction and its consequences on customers' behavioural intentions: A study of selected service firms in Tanzania*. PhD Thesis, University of Dar es Salaam.
2. Atrek, B., & Bayraktaroglu, G. (2012). Is there a need to develop a separate service quality scale for every sector?: Verification of SERVQUAL in higher education services. *Journal of Faculty of Economics and Administrative Sciences*, 17(1), 423-440.
3. Brochado, A. O., & Marques, R.C. (2009). Comparing alternative instruments to measure service quality in higher education. *Quality Assurance in Education*, 17(2), 174-190.
4. Cavana, R. Y., Corbett, L. M., & Lo, Y. L. (2007). Developing zones of tolerance for managing passenger rail service quality. *International Journal of Quality and Reliability Management*, 24(1), 7–31.
5. Cohen, R., & Swerdlik, M. (2010). *Psychological testing and assessment*. Boston: McGraw-Hill.
6. Cook, D.A., & Beckman, T.J. (2006). Current concepts in validity and reliability for psychometric instruments: Theory and application. *American Journal of Medicine*, 119, 116-166.
7. Creswell, J.W. (2008). *Educational research: Planning, conducting and evaluating quantitative and qualitative research*. New Jersey: Pearson Prentice Hall.
8. Fink, A. (1995). *The survey hand book*. California: Sage Publications.
9. Ghauri, P. N., & Gronhaug, K. (2010). *Research methods in business studies* (4th ed.). New York: Prentice Hall.
10. Hair, J., Black, B., Anderson, R.E., & Tatham, R.L. (2005). *Multivariate data analysis*, (6th ed.). New Jersey: Prentice Hall.
11. Hoe, T.C. (2005). *Measuring student perception of service quality in higher education*. PhD Thesis, University of South Australia.
12. Lee, J. (2007). SERVQUAL versus SERVPERF: Round 2 in a multi-cultural setting. *Journal of Academy of Business and Economics*, 7(3), 27-41.
13. Nunnally, J.C., & Bernstein, I.H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.
14. Palmer, A. (2011). *Principles of marketing* (6th ed.). Maidenhead: McGraw-Hill Education.
15. Parasuraman, A., Zeithaml, V.A., & Berry, L.L. (1993). Research note: More on improving service quality measurement. *Journal of Retailing*, 69(1), 140-147.
16. Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1991). Refinement and reassessment of the SERVQUAL scale. *Journal of Retailing*, 67(3), 420-450.
17. Patten, M.L. (2005). *Understanding research methods: An overview of the essentials*. New York: Pyczak Publishing.

18. Polit, D.F., & Beck, C.T. (2004). *Nursing research: Principles and methods* (7th ed.). Philadelphia: Lippincott, Williams and Wilkins.
19. Soliman, A. A., & Alzaid, A. A. (2002). Service quality in Riyadh's elite hotels: Measurement and evaluation. *Journal of King Saud University*, 14(2), 83-103.
20. Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 1(2), 53-55.
21. Tegambwage, A.M.G., & Ame, A.M. (2016). Measuring service quality in higher education. *Asian Research Journal of Business Management*, 4(1), 85-101.
22. Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington, DC: American Marketing Association.
23. Tsoukatos, E.K. (2007). *Customer behaviour, service quality and the effects of culture: A quantitative analysis in Greek insurance*. PhD Thesis, Lancaster University.
24. Vanpariya, B., & Ganguly, P. (2010). SERVQUAL versus SERVPERF: An assessment from Indian banking sector. *QUEST Journal of Management and Research*, 1 (1), 123-137.