# Testing the Validity, Reliability and Independence of the Constructs in the SERVQUAL Tool for Measuring Student Satisfaction in Universities in Uganda

*Martha Kyoshaba[1], Fred E. K. Bakkabulindi[2]
[1]Mbarara University of Science and Technology, Uganda
[2]East African School of Higher Education Studies and Development, College of Education and External Studies, Makerere University, Uganda
*Corresponding author: Email: mkyoshaba@must.ac.ug
ORCID:0000-0002-4663-950X
https://doi.org/10.58653/nche.v11i1.02
**(Accepted: 17th October 2023/Published: 25th October 2023)**

## Abstract

The service quality (SERVQUAL) tool is a tool used to measure customer satisfaction (CS) in general and student satisfaction (SS) in particular. It consists of five constructs: tangibles (T), reliability (Rel), responsiveness (Res), assurance (A) and empathy (E). Although it has been widely used in a number of CS/SS studies, it has attracted a lot of criticism, one of which relates to the validity, reliability and independence of its constructs. In this paper we adapted the SERVQUAL tool for measuring SS in universities in Uganda and then tested the validity, reliability and independence of its constructs. Our focus was on two objectives: (i) to establish the validity and reliability of each of the five constructs of SERVQUAL; and (ii) to test if the five constructs were independent. We analysed our data using: (i) confirmatory factor analysis (CFA) and Cronbach's alpha ($\alpha$); and (ii) the Pearson linear correlation. Our findings were to the effect that: (i) Except for the "A" which was valid without dropping any items, the other constructs (T, Rel, Res, and E) were valid only after we dropped some of their items. We also established that the reliabilities of the five constructs of SERVQUAL were high. (ii) The constructs (T, Rel, Res, A and E) of SERVQUAL were interrelated. The SERVQUAL tool is a good measure of student satisfaction (SS). We recommend the SERVQUAL tool for continuous testing in other contexts with the objective of improving it.

***Keywords***: *Cronbach's alpha; Correlation; Customer; Factor analysis; Measurement tool.*

## Introduction

Student satisfaction (SS) is a multi-faceted construct, characterised by its multidimensional nature which stems from the various elements that comprise and influence it (Bell, 2022; Wong & Chapman, 2022; Abu-Rumman & Qawasmeh, 2021). Kanwar and Sanjeeva (2022) define SS as a short-term attitude arising from an evaluation of a student's educational experiences. Wong and Chapman (2022) have a similar definition but add that the student's short-term attitude arises from that student's subjective evaluation of the extent to which their expectations of an educational experience have been met or exceeded. In the light of these definitions, we may say that SS is a good attitude that a student develops after encountering services offered at an education institution. This attitude stems from the student's expectation of a given service and that student's perception of the service after its receipt. SS is important to a university because it is a "crucial index for performance of higher education institutions in today's world" (Wong & Chapman, 2022, p.958). It not only serves as a determinant of a university's overall success (Abu-Rumman & Qawasmeh, 2021) but also has the potential to influence academic achievement, student retention and motivation (Wong & Chapman, 2022). Furthermore, SS may lead to student loyalty (Osman & Saputra, 2019) and enhance a university's reputation, thereby giving the university a competitive edge (Karna & Julin, 2015) which, in turn, may lead to profitability (Guilbault, 2017; Mihanovic, Batinic & Pavicic, 2016). Given the importance of SS in higher education, universities should measure it to ascertain the level of SS, identify critical services and allocate resources accordingly (Eres & Clothey, 2013).

How, then, is student satisfaction (SS) measured? A number of researchers have made attempts at developing tools which measure customer satisfaction (CS) in general and SS in particular. Some of the tools are: service quality (SERVQUAL) (Parasuraman, Zeithaml & Berry, 1985, 1988, 1991), service performance (SERVPERF) (Cronin & Taylor, 1992), higher education performance (HEdPERF) (Abdullah, 2006), student satisfaction inventory (SSI) (Elliot & Shin, 2002) and Shahsavar and Sudzina's (2017) tool for testing SS. In this paper our focus is on the SERVQUAL tool, the validity and reliability of whose use, in spite of its being extensive in CS/SS studies (Ladhari, 2009), have been questioned by scholars (Landrum, Prybutok, Zhang & Peak, 2009). In addition, there are few publications on studies which sought to test the validity, reliability and independence of the five SERVQUAL constructs, particularly in Uganda. With that background, we have two objectives for this paper: (i) to establish the validity and reliability of each of the five constructs (i.e. tangibles [T], reliability [Rel], responsiveness [Res], assurance [A] and empathy [E]) of SERVQUAL; and (ii) to test if the five constructs of SERVQUAL are independent.

## Literature Review

*The Development of the SERVQUAL Tool.* Parasuraman, Zeithaml and Berry (1985) developed the SERVQUAL concept in 1985 based on a gap model, i.e. perceptions minus expectations (P-E). They suggested that customers form expectations and perceptions of service quality (SQ) using 10 dimensions, namely reliability, responsiveness, competence, access, courtesy, communication, credibility, security, understanding/knowing the customer and tangibles. Each dimension had eight to 10 items. The total number of items was 97 (Parasuraman et al., 1985). Each item had two statements; one statement was about a customer's expectations and the other a customer's perceptions of a service. Noting that there had been overlaps in the 10 dimensions, Parasuraman et al. (1988) developed the SERVQUAL scale, reducing the 10 dimensions with 97 items to five with 22 items. Parasuraman et al. (1991) carried out what they termed as "scale purification" and then revised the SERVQUAL instrument and thereafter tested its validity and reliability.

Using the data they collected from 406 respondents from five companies (i.e. a telephone company, two insurance firms and two banks) in the USA, they reported that a "high degree of convergence between the revised SERVQUAL scale and a separate measure of service quality (OSQ) support[ed] the scale's construct validity" (Parasuraman et al., 1991, p. 433). With regard to

reliability, the five dimensions scored high Cronbach's alpha coefficients as measures of reliability, ranging from a minimum of 0.80 for tangibles (T) to a maximum of 0.93 for responsiveness (Res). Thus the "five SERVQUAL dimensions [were] consistently high across the various samples, thereby indicating high internal consistency among items within each dimension" (Parasuraman et al., p. 424). In terms of independence, Parasuraman et al. (1991) reported that "though the SERVQUAL dimensions represent five conceptually distinct facets of service quality, they are interrelated … the responsiveness items while exhibiting some degree of distinctiveness, have sizeable cross-loadings on … assurance and reliability" (p. 442). Though Parasuraman et al. (1991) did not point out any gaps in their study, they suggested two areas for further study: one, "the nature and causes of these interrelationships [between the dimensions]" (p.442) and, two, "the measurement of expectations and the related issues of computing perception minus expectation gap scores" (p. 443).

A number of scholars have criticised SERVQUAL on issues such as "use of gap scores, the overlap among the five dimensions, poor … convergent validity, the ambiguous definition of the 'expectation' construct, and unstable dimensionality" (Landrum et al., 2009, p. 20). Parasuraman et al. (1991) critiqued themselves when they observed that some studies were finding a challenge with the "appropriateness of using difference scores in multivariate analyses … on grounds that such scores might suffer from low reliability and validity" (p. 444). Silva, Moraes, Makiya and Ceasar (2017) reported that some scholars had questioned the application of the SERVQUAL tool in higher education settings.

*Empirical Studies on the Validity, Reliability and Independence of the Constructs of the SERVQUAL Tool*. Abu-Rumman and Qawasmeh (2021) tested the validity and reliability of the constructs of the SERVQUAL tool as part of their study. They did not explicitly test the independence of the constructs of the SERVQUAL tool but did so as a coincident. They used a self-administered questionnaire (SAQ) which they administered to international students at Amman-Jordan University in Jordan. They used structural equation modelling (SEM) to analyse the validity of the SERVQUAL scale and reported that "… the SERVQUAL scale is an acceptable model." They did not, however, provide specific information on the validity of each of the constructs.

Regarding measures of reliability, all the constructs (T, Rel, Res, A and E) had high Cronbach's alpha coefficients, ranging from a minimum of 0.823 for A to a maximum of 0.880 for Res. These high reliabilities, all of them greater than 0.7, implied that the items were reliable measures of the constructs. Regarding the independence of the constructs, they used confirmatory factor analysis (CFA) and reported that the results showed good fitness, which indicated that the constructs were independent. They reported three limitations to their study. One was on generalisability. They carried out the study on international students in one university in Jordan. Secondly, they used a cross-sectional design, which limits the ability to establish causality between service quality dimensions and student satisfaction. Thirdly, they relied on self-reported data, which may be subject to response bias.

Cerri (2012) tested the validity and reliability of the constructs of the SERVQUAL tool as part of a study. However, testing the independence of the constructs was done inadvertently. Cerri used a self-administered questionnaire (SAQ) which she administered to 261 students from five public universities in Albania. Hence the constructs (tangibles [T], reliability [Rel], responses [Res], assurance [A] and empathy [E]) scored high average variances extracted (AVEs) as measures of validity, ranging from a minimum 0.79 for Res to a maximum of 0.920 for Rel. "All AVEs exceed[ed] 0.50 indicating solid construct validity" (p. 673).

Regarding measures of reliability, all the constructs (T, Rel, Res, A and E) had high Cronbach's alpha coefficients, ranging from a minimum of 0.790 for A to a maximum of 0.940 for T. These high reliabilities, all of them greater than 0.7, implied that the items were reliable measures of the constructs. With regard to correlations among constructs, without citing actual figures, Cerri stated that "correlation coefficients [were] all significant at 0.01" (p. 673). The study reported two limitations. One was to the effect that Cerri carried out the study in only public universities and in

only one country, which could curtail generalisability of results to private universities and other countries respectively. The second was that the study used a small sample size which could affect the validity of the results of the study.

Green (2014) tested the reliability of the SERVQUAL tool as part of a main study. He used a SAQ to collect data from 280 respondents. With regard to reliability under expectations, all the constructs (T, Rel, Res, A and E) had high Cronbach's alpha coefficients, ranging from a minimum of 0.757 for E to a maximum of 0.844 for T and Rel. Regarding the reliability under perceptions, all the constructs (T, Rel, Res, A and E) had high Cronbach's alpha coefficients, ranging from a minimum of 0.790 for T to a maximum of 0.899 for Res. Hence "all Cronbach's alpha coefficients [were] higher than 0.7, indicating a reasonably high level of reliability of the measurement instrument" (p.135). Green neither reported about the validity of his instrument nor about the independence of the dimensions of SERVQUAL. He reported a limitation to the effect that he had used a small sample, which could affect generalisability of his findings.

Leonnard (2018) tested the validity and reliability of the SERVQUAL tool as part of his main study. Using a SAQ, he gathered data from 319 students. With regard to validity, although Leonard reported that he tested the construct validity of his tool, he did not reveal the results of the validity of each construct. Regarding measures of reliability, all the constructs (T, Rel, Res, A and E) had high Cronbach's alpha coefficients, ranging from a minimum of 0.728 for E to a maximum of 0.822 for Rel. Thus, the results "indicated that all the item values were > 0.600 and were reliable to measure each construct" (p. 17). With regard to correlation among the constructs, Leonnard did not reveal any results. His limitation was that he carried out his study in only private universities, which could curtail generalisability of his findings to public universities.

Nyandwi and Bakkabulindi (2018) tested the reliability of the SERVQUAL tool as part of their main study. They used a SAQ to gather data from 200 international students. With regard to validity, without citing actual figures, they stated that "the validity was taken for granted basing on the observation that an instrument that is reliable is also likely to be valid" (p.9). Regarding measures of reliability, all the constructs (T, Rel, Res, A and E) scored high composite reliabilities (CRs), ranging from a minimum 0.710 for E and a maximum of 0.910 for T. These high reliabilities, all of them greater than 0.7, implied that they used a reliable SERVQUAL instrument. Regarding independence of the constructs, they did not reveal any results. They reported a limitation to the effect that their study had used only international undergraduate students at one university, hence curtailing generalisability of their results to other students.

Tegambwage and Ame (2017) tested the validity and reliability of the constructs of the SERVQUAL tool as their main study. However, they did not test the independence of the constructs. They used a SAQ to collect data from 250 students. Regarding validity, without citing actual figures, they reported that "the presence of a positive and significant correlation coefficient (p=0.01) between service quality… and the overall service quality rating, across all three data sets (university A, B and the combined samples) [was] an indication of the convergent validity" (p. 90). Regarding reliability, they did not cite actual scores for each construct but stated that "the results … indicate[d] high internal consistency among items within each dimension, ranging from 0.74 (tangibles) to 0.96 (non-tangibles)" (p. 89). They did not reveal any results for the independence of the constructs. They did not point out any limitations to their study.

Vaz and Mansori (2013) tested the validity and reliability of SERVQUAL constructs as part of a main study but did not test the independence of the constructs. They used a SAQ which they administered to 431 undergraduate students. The constructs (Rel, T, Res, E and A) scored acceptable average variances extracted (AVEs) as measures of validity, ranging from a minimum 0.400 for T to a maximum of 0.680 for E. Regarding reliability, the constructs scored high composite reliabilities (CRs) as measures of reliability, ranging from a minimum 0.818 for T and a maximum of 0.884 for Res. The constructs scored even higher Cronbach's alpha values, ranging from a minimum 0.722

for A to a maximum of 0.836 for Rel. Hence all the "constructs … met the minimum threshold required" (p. 166). With regard to the correlation among the constructs, they did not give any figures. They reported their limitation as being a small sample size from three private university colleges.

In summary, we reviewed seven empirical studies which tested the quality of the SERVQUAL tool. Two of these studies (i.e. Cerri, 2012; Tegambwage & Ame, 2017) sought to test the validity and reliability of the SERVQUAL tool as the main purpose for their study but did not test the independence of the constructs, while the rest tested the validity, reliability and independence of the same tool inadvertently. All the seven studies modified the SERVQUAL items to suit the higher education settings, which is acceptable by Parasuraman et al. (1991), who suggested that "… minor modifications in the wording of items to adapt them to a specific setting are appropriate …" (p. 445). However, Green (2014) not only re-worded some items but also re-categorised the dimensions of SERVQUAL, which may not be a minor change to the SERVQUAL tool and could have affected the integrity of the scale. Regarding the unit of analysis, only Green had both students and staff as their units of analysis, while the rest of the researchers' units of analysis were students and the samples ranged from 200 respondents (Nyandwi & Bakkabulindi, 2018) to 431 (Vaz & Mansori, 2013), which they reported as small samples.

With regard to analysing the validity, four (i.e. Cerri, 2012; Leonnard, 2018; Tegambwage & Ame, 2017; Vaz & Mansori, 2013) out of the seven studies analysed the validity of their instrument using two methods, namely average variance extracted scores (AVE) (i.e. Cerri, 2012; Vaz & Mansori, 2013) and confirmatory factor analysis (CFA) (Abu-Rumman & Qawasmeh, 2021; Leonnard, 2018; Tegambwage & Ame, 2017). All the studies reported that their items were valid measures of the SERVQUAL constructs. Green (2014) and Nyandwi and Bakkabulindi (2018) did not test the validity of their instrument.

Regarding reliability, all the seven studies analysed the reliability of the constructs of their SERVQUAL tool, with Cerri (2012) and Vaz and Mansori (2013) using both Cronbach's alpha and composite reliability, with Abu-Rumman and Qawasmeh (2021), Green (2014), Leonnard (2018) and Tegambwage and Ame (2017) using only Cronbach's alpha, and with Nyandwi and Bakkabulindi (2018) using only composite reliability. The constructs scored high reliabilities exceeding the threshold of 0.7 (Tavakol & Dennick, 2011), suggesting that the items were reliable measures of constructs of the SERVQUAL tool. With regard to the independence of the constructs, Abu-Rumman and Qawasmeh (2021) reported that the results showed good fitness, which indicated that the constructs were independent. Cerri (2012) reported that the "correlation coefficients [were] all significant at 0.01 level" (p. 673). The rest of the studies did not analyse the independence of the SERVQUAL constructs, hence a gap existed.

We, therefore, noted three gaps from our empirical studies. One was that five out of the seven studies we cited did not primarily set out to test the validity and reliability of the SERVQUAL tool; they tested it as a coincidence. This means that, first, few studies have been published which have specifically sought to test the validity and reliability of the SERVQUAL tool in spite of its wide usage. Second, only two of seven studies examined whether the constructs of SERVQUAL were independent, which also points us to the gap resulting from the inadequate number of publications on testing the independence of the five constructs of SERVQUAL. Third, all the studies had small sample sizes, which could influence the validity of the results. Our study, therefore, attempted to narrow the gaps by (i) testing the validity and reliability of each of the five constructs of SERVQUAL; and (ii) testing if the five constructs were independent. In addition, we narrowed the gap of small sample size by using a larger sample of 704 students. We achieved objectives (i) and (ii) by analysing our data using confirmatory factor analysis (CFA) and the Pearson linear correlation (PLC), respectively.

# Methodology

## Sample

Our sample comprised 704 students whom we selected using convenience sampling from seven universities in Uganda (i.e. Bishop Stuart University, Kabale University, Kampala International University, Makerere University, Mountains of the Moon University, Mbarara University of Science and Technology and Uganda Christian University). Our typical respondent was a male (51.3%) Ugandan student (97%), from the Western Region (64.2%) aged 20 but below 25 (72.3%). He was from Makerere University (36.2%) undertaking a bachelor's degree (85.2%) and in his first year of study (46.9%).

## Data collection instrument

We used a self-administered questionnaire (SAQ) (Appendix A) for data collection. We adapted Parasuraman et al.'s (1991) SERVQUAL tool to measure our main construct, student satisfaction (SS) (see Table 1). We operationalised SS based on SERVQUAL's five constructs: tangibles (T), reliability (Rel), responsiveness (Res), assurance (A) and empathy (E). We used the perception section of the SERVQUAL instrument and left out the expectation section similar to what Landrum et al. (2009) did in their study on measuring IS system service quality with SERVQUAL. We re-worded the items in order to suit university contexts, i.e. questions under each of the constructs T, Rel, Res, A and E about academic staff, and another about administrative staff. That increased the items for each of the constructs, hence, overall, the number of items increased to 41 from 22 (Appendix A). We scaled the items using the five-point Likert scale, ranging from 1 for "strongly disagree" (SD) to five for "strongly agree" (SA).

*Table 1:* Constructs in the Instrument

| Construct | Number of items adapted | Source of instrument, number of items and their reliability ($\alpha$ value) |
|---|---|---|
| Tangibles | 04 | Parasuraman et al. (1991), 04 items ($\alpha = 0.60$) |
| Reliability | 05 | Parasuraman et al. (1991), 05 items ($\alpha = 0.85$) |
| Responsiveness | 04 | Parasuraman et al. (1991), 04 items ($\alpha = 0.61$) |
| Assurance | 04 | Parasuraman et al. (1991), 04 items ($\alpha = 0.81$) |
| Empathy | 05 | Parasuraman et al. (1991), 05 items ($\alpha = 0.66$) |

# Results

**Objective one**: To establish the validity and reliability of each of the five constructs of SERVQUAL

The five constructs of SERVQUAL were tangibles (T), reliability (Rel), responsiveness (Res), assurance (A) and empathy (E). To test for validity of the constructs, we carried out a confirmatory factor analysis (CFA). We adopted the Kaiser-Guttman rule (Schmidt, Baran, Thompson, Mishra, Koelher & Shin, 2009). The rule posits that a factor is significant if it has an eigenvalue of a magnitude of at least one. To decide on the most valid items, we based ourselves on the rotated component matrix because, according to Mvududu and Sink 2013), those results can be readily interpreted. The items we considered as the most valid were those that loaded highly on the first factor (Kahn, 2006), with loading of at least 0.5 (Matsunaga, 2010). We determined reliability using Cronbach's alpha, and considered as reliable items with a reliability index of at least 0.7 (Tavakol & Dennick, 2011).

*Tangibles (T).* The results in Table 2 show that our confirmatory factor analysis (CFA) reduced the five items of "T" (T1-T5) to two significant factors, with 2.540 and 1.011 as their eigenvalues. Hence the two factors 1 and 2, respectively, accounted for 2.540/5*100 = 50.795% and 1.011/5*100 = 20.225% of the total variance among the five items. Table 2 also shows that items T3 - T5 loaded highly with a loading of over 0.5 on the first factor, hence they were the most valid items of "T". Items T1 and T2 loaded highly on the second and less significant factor and, therefore, we did not consider them as valid items of construct "T". The reliability index of the three most valid items (T3- T5) using Cronbach's alpha ($\alpha$) was 0.716, which, being large (greater than 0.7), suggested that the items T3-T5 were also reliable measures of "T".

***Table 2:*** Loadings and Cronbach's Alpha on the Factors on T

| | | Loadings | | |
|---|---|---|---|---|
| **Items** | **Description** | **Factor 1** | **Factor 2** | **Cronbach's Alpha ($\alpha$)** |
| T1 | My university has modern looking equipment | | 0.875 | 0.716 |
| T2 | The physical facilities in my university are visually appealing | | 0.864 | |
| T3* | The academic staff in my university appear neat | 0.844 | | |
| T4* | The administrative staff in my university appear neat | 0.874 | | |
| T5* | The printed materials (e.g. admission letters, registration forms, brochures) in my university are visually appealing | 0.613 | | |
| Eigenvalue | | 2.540 | 1.011 | |
| % variation explained | | 59.795 | 20.225 | |

*Valid items of "T"

*Reliability (Rel).* The results in Table 3 show that our confirmatory factor analysis (CFA) reduced the 10 items of "Rel" (Rel1-Rel10) to two significant factors, with 5.009 and 1.304 as their eigenvalues. Hence the two factors 1 and 2, respectively, accounted for 5.009/10*100 = 50.091% and 1.304/10*100 = 13.037% of the total variance among the 10 items. Table 3 also shows that items Rel1-Rel8 loaded highly with a loading of over 0.5 on the first factor, thus they were the most valid items of "Rel". Items Rel9 and Rel10 loaded highly on the second and less significant factor and, therefore, we did not consider them as valid items of construct "Rel". The reliability index of the eight most valid items (Rel1- Rel8) using Cronbach's alpha ($\alpha$) was 0.894, which, being large (greater than 0.7), suggested that the items Rel1-Rel8 were also reliable measures of "Rel".

***Table 3:*** Loadings and Cronbach's Alpha on the Factors on Rel

| | | Loadings | | |
|---|---|---|---|---|
| **Items** | **Description** | **Factor 1** | **Factor 2** | **Cronbach's alpha ($\alpha$)** |
| Rel1* | When academic staff in my university promise to do something for me by a certain time, they do it | 0.746 | | 0.894 |
| Rel2* | When administrative staff in my university promise to do something for me by a certain time, they do it | 0.741 | | |
| Rel3* | When I have a problem, an academic staff in my university shows interest in solving it | 0.703 | | |
| Rel4* | When I have a problem, an administrative staff in my university shows interest in solving it | 0.690 | | |

| | | | | |
|---|---|---|---|---|
| Rel5* | Academic staff in my university perform services for me right the first time | 0.735 | | |
| Rel6* | Administrative staff in my university perform services for me right the first time | 0.761 | | |
| Rel7* | Academic staff in my university provide services at the time they promise me to do so | 0.757 | | |
| Rel8* | Administrative staff in my university provide services at the time they promise me to do so | 0.773 | | |
| Rel9 | The academic records given to me by academic staff in my university are error free | | 0.896 | |
| Rel10 | The academic records given to me by administrative staff in my university are error-free | | 0.913 | |
| Eigenvalue | | 5.009 | 1.304 | |
| % variation explained | | 50.091 | 13.097 | |

*Valid items of "Rel"

Responsiveness (Res). The results in Table 4 show that our confirmatory factor analysis (CFA) reduced the eight items of "Res" (Res1-Res8) to two significant factors with 3.941 and 1.213 as their eigenvalues. Hence, the two factors 1 and 2, respectively, accounted for 3.941/8*100 = 49.266% and 1.213/8*100 = 15.162% of the total variance among the eight items. Table 6 also shows that items Res1-Res4 loaded highly with a loading of over 0.5 on the first factor, hence they were the most valid items of "Res". Items Res5-Res8 loaded highly on the second and less significant factor and, therefore, we did not consider them as valid items of construct "Res". The reliability index of the four most valid items (Res1- Res4) using Cronbach's alpha ($\alpha$) was 0.830, which, being large (greater than 0.7), suggested that the items Res1-Res4 were also reliable measures of "Res".

***Table 4:*** Loadings and Cronbach's Alpha on the Factors on Res

| | | Loadings | | |
|---|---|---|---|---|
| **Items** | **Description** | **Factor 1** | **Factor 2** | **Cronbach's alpha ($\alpha$)** |
| Res1* | Academic staff in my university tell me exactly when a service will be performed | 0.828 | | 0.830 |
| Res2* | Administrative staff in my university tell me exactly when a service will be performed | 0.857 | | |
| Res3* | Academic staff in my university give me prompt response whenever necessary | 0.694 | | |
| Res4* | Administrative staff in my university give me prompt response whenever necessary | 0.720 | | |
| Res5 | Academic staff in my university are willing to help me whenever need arises | | 0.739 | |
| Res6 | Administrative staff in my university are willing to help me whenever need arises | | 0.725 | |
| Res7 | Academic staff in my university are not too busy to respond to my requests | | 0.787 | |

| | | | | |
|---|---|---|---|---|
| Res8 | Administrative staff in my university are not too busy to respond to my requests | | 0.40 | |
| Eigenvalue | | 3.941 | 1.213 | |
| %variation explained | | 49.266 | 15.162 | |

*Valid items of "Res"

*Assurance (A).* The results in Table 5 show that our confirmatory factor analysis (CFA) reduced the eight items of "A" (A1-A8) to one significant factor with 4.184 as its eigenvalue, implying that the factor accounted for 4.184/8*100 = 52.303% of the total variance among the eight items. All the eight items of "A" (A1-A8) loaded highly on the first factor (all loadings above 0.5), hence all of them were valid items of "A". The eight items had a Cronbach's alpha ($\alpha$) of 0.869, which, being larger than 0.7 (greater than 0.7), suggested that the items were also reliable measures of "A".

***Table 5:*** Loadings and Cronbach's Alpha on the Factors on A

| Items* | Description | Factor Loadings | Cronbach's alpha ($\alpha$) |
|---|---|---|---|
| A1 | The behaviour of academic staff in my university instils confidence in me | 0.711 | 0.869 |
| A2 | The behaviour of administrative staff in my university instils confidence in me | 0.742 | |
| A3 | I feel safe dealing with academic staff in my university | 0.760 | |
| A4 | I feel safe dealing with administrative staff in my university | 0.787 | |
| A5 | Academic staff in my university are polite to me | 0.693 | |
| A6 | Administrative staff in my university are polite to me | 0.738 | |
| A7 | Academic staff in my university have knowledge to answer my inquiries | 0.672 | |
| A8 | Administrative staff in my university have knowledge to answer my inquiries | 0.674 | |
| Eigenvalue | | 4.184 | |
| %variation explained | | 52.303 | |

*Valid items of "A"

*Empathy* (E). The results in Table 6 show that our confirmatory factor analysis (CFA) reduced the 10 items of "E" (E1-E10) to two significant factors, with 5.082 and 1.200 as their eigenvalues. Hence, the two factors 1 and 2 accounted for 5.082/10*100 = 50.822% and 1.200/10*100 = 12.000%, respectively, of the total variance among the 10 items. Table 6 shows that items E1-E2 and E5-E10 loaded highly with a loading of over 0.5 on the first factor, hence they were the most valid items of "E". Items E3-E4 loaded highly on the second and less significant factor and, therefore, we did not consider them as valid items of construct "E". The reliability index of the eight most valid items (E1-E2 and E5-E10) using Cronbach's alpha ($\alpha$) was 0.897, which, being large (greater than 0.7), suggested that the items E1-E2 and E5-E10 were also reliable measures of "E".

*Table 6:* Loadings and Cronbach's Alpha on the Factors on "E"

| | | Loadings | | |
|---|---|---|---|---|
| **Items** | **Description** | **Factor 1** | **Factor 2** | **Cronbach's Alpha (α)** |
| E1* | I receive attention from academic staff in my university when I need it | 0.593 | | 0.897 |
| E2* | I receive attention from administrative staff in my university when I need it | 0.596 | | |
| E3 | The operating hours of academic staff in my university are convenient for me | | 0.813 | |
| E4 | The operating hours of administrative staff in my university are convenient for me | | 0.823 | |
| E5* | Academic staff in my university are concerned about me | 0.656 | | |
| E6* | Administrative staff in my university are concerned about me | 0.677 | | |
| E7* | Academic staff in my university have my interests at heart | 0.755 | | |
| E8* | Administrative staff in my university have my interests at heart | 0.785 | | |
| E9* | Academic staff in my university understand my specific needs | 0.823 | | |
| E10* | Administrative staff in my university understand my specific needs | 0.822 | | |
| Eigenvalue | | 5.082 | 1.200 | |
| % Variation explained | | 50.822 | 12.000 | |

*Valid items of "E"

**Objective two**: To test if the five constructs of the SERVQUAL tool were independent.

In order to achieve objective two, we calculated the average indexes for the valid items of our five constructs and thereafter correlated them using the Pearson linear correlation (PLC). The results of this (see Table 7) show that all the five constructs were significantly interrelated, with the highest correlation between E and A (r = 0.691) and the lowest between Res and T (r = 0.396). Table 7 also shows that T had the lowest correlation with other constructs while Rel had the strongest correlation with other constructs.

*Table 7:* Inter-correlations of the SERVQUAL Constructs

| Constructs | T | Rel | Res | A | E |
|---|---|---|---|---|---|
| T | | 0.550** | 0.396** | 0.505** | 0.405** |
| Rel | | | 0.615** | 0.640** | 0.665** |
| Res | | | | 0.600** | 0.583** |
| A | | | | | 0.691** |
| E | | | | | |

** Correlation is significant at the 0.01 significance level

## Discussion

Objective one of our study was to establish the validity and reliability of each of the five constructs of SERVQUAL. The five constructs of SERVQUAL, as developed by Parasuraman et al. (1991), are tangible (T), reliability (Rel), responsiveness (Res), assurance (A) and empathy (E). We tested for validity using confirmatory factor analysis (CFA) and established that the items under construct T were categorised under two factors: factor 1 (see Table 2, i.e. T3, T4, T5), which was the most significant factor, and factor 2 (Table 2, i.e. T1, T2), which was the less significant factor. This implies that if T were an independent variable in the SERVQUAL model, then we would have to regress the dependent variable on two different factors on T and not one, as suggested in the SERVQUAL tool. This finding supports Parasuraman et al.'s (1991) rather dated source which found that "tangibles split into two subdivisions … one pertaining to physical facilities/equipment and another pertaining to employees/communication materials" (p. 431, para 1). If we were to look for valid items on T, then, as was done earlier, the factors which loaded highly (T3, T4, T5) on the first and most significant factor were the most ideal set since their reliability is high at $\alpha$ 0.716. Hence, dropping two items (T1, T2) from the tool in favour of only three valid ones could mean that the construct T in SERVQUAL as operationalised by Parasuraman et al. (1991) may be unnecessarily long.

With regard to the second construct, reliability (Rel), the confirmatory factor analysis (CFA) showed that items under construct Rel were categorised under two factors: factor 1 (see Table 3, i.e. Rel1– Rel8), which was the most significant factor, and factor 2 (Table 3, i.e. Rel9 and Rel10), which was the less significant factor. This implies that if Rel were an independent variable in the SERVQUAL model, then we would have to regress the dependent variable on two different factors on Rel and not one, as suggested by the SERVQUAL tool. If we were to look for valid items on Rel, then, the factors which loaded highly (Re1– Rel8) on the first and most significant factor were the ideal set since their reliability is high at $\alpha$ 0.894. Hence, dropping two items (Rel9 and Rel10) from the tool in favour of eight valid ones could mean that the construct Rel in the SERVQUAL tool as operationalised by Parasuraman et al. (1991) may be unnecessarily long.

We carried out a confirmatory factor analysis (CFA) for the construct responsiveness (Res) and established that items under construct Res were categorised under two factors: factor 1 (see Table 4, i.e. Res1– Res4), which was the most significant factor, and factor 2 (Table 4, i.e. Rel5 and Rel8), which was the less significant factor. This implies that if Res were an independent variable in the SERVQUAL model, then we would have to regress the dependent variable on two different factors on Res and not one, as suggested in the SERVQUAL tool. If we were to look for valid items on Res, then the factors which loaded highly (Res1– Res4) on the first and most significant factor were the most ideal set since their reliability is high at $\alpha$ 0.830. Hence, dropping four items (Rel5 and Res8) from the tool in favour of four valid ones could mean that the construct Res in SERVQUAL as operationalised by Parasuraman et al. (1991) may be unnecessarily long.

The confirmatory factor analysis (CFA) for the construct assurance (A) showed that items under construct A were categorised under one factor (see Table 5, i.e. A1– A8), which was the significant factor. Implying that if A were an independent variable in the SERVQUAL tool, then we would have to regress, the dependent variable on one factor A as suggested in the SERVQUAL model. If we were to look for valid items on A, then since all items (A1– A8) loaded highly on only one factor and their reliability is high at $\alpha$ 0.869 thus not dropping any item from the tool it could mean that the construct A in SERVQUAL as operationalised by Parasuraman et al. (1991) is the right length.

We carried out a confirmatory factor analysis (CFA) for the construct empathy (E) and established that items under construct E were categorised under two factors: factor 1 (see Table 6, i.e. E–E2 and E5– E10), which was the most significant factor, and factor 2 (Table 8, i.e. E3– E4), which was the less significant factor. This implies that if E were an independent variable in the

SERVQUAL model, then we would have to regress the dependent variable on two different factors on E and not one, as suggested in the SERVQUAL tool. If we were to look for valid items on E, then the factors which loaded highly (E1– E2 and E5– E10) on the first and most significant factor were the ideal set since their reliability is high at $\alpha$ 0.897. Hence, dropping two items (E3– E4) from the tool in favour of eight valid ones could mean that the construct E in the SERVQUAL tool as operationalised by Parasuraman et al. (1991) may be unnecessarily long.

We, therefore, conclude that our CFA suggested that, except for construct A for which no item was dropped in order for A to be valid, the other four constructs (T, Rel, Res and E) of the SERVQUAL tool were valid, after we dropped a number of items. This, therefore, calls for continuous research on the SERVQUAL tool in the area of testing its validity and reliability until the four constructs (T, Rel, Res and E) achieve the ideal length. In comparison to the findings of the seven studies we have cited in the literature review concerning validity, our findings support those of Abu-Rumman and Qawasmeh (2021), Cerri (2012), Leonnard (2018) and Tegambwage and Ame (2017), who reported their items as valid measures of the different constructs of SERVQUAL. With regard to reliability, our findings show high Cronbach's alpha coefficients as measures of reliability with the minimum alpha of 0.716 for tangibles and a maximum alpha of 0.897 for empathy. These results lend support to studies we cited which reported that the items were reliable measures of their respective constructs in the SERVQUAL tool.

Objective two of our paper was to test if the five constructs of SERVQUAL were independent. The results of the correlation analysis (Table 7) suggested that all the five constructs were significantly interrelated, confirming Parasuraman et al.'s (1991) conclusion regarding the same. Although Parasuraman et al. (1991) did not critique themselves on the interrelatedness of SERVQUAL dimensions, they raised it as a critical area for further study, i.e. the need for other researchers to investigate "the nature and causes of these interrelationships [between the dimensions]" (p.442). This finding calls into question whether the constructs measure different concepts. This finding may suggest that if we were to undertake a study in which the SERVQUAL constructs were explanatory variables, we could use some and drop some to avoid multi-collinearity (Sweet & Grace-Martin, 2003). On the other hand, Vaz and Mansori (2013) reported that there were no serious multi-correlations among independent variables, hence differing with our findings.

Table 7 also shows that T had the lowest correlation with other constructs while Rel had the strongest correlation with other constructs. These results support Parasuraman et al.'s (1991) finding that "reliability ha[d] the strongest coefficients, … and tangibles ha[d] the weakest coefficients" (p. 433).

## Conclusion

In this study we set out to test the validity and reliability of SERVQUAL as a tool which measures student satisfaction (SS) in universities in Uganda. Our paper is among the few papers which have done so in the context of a developing country. We found the SERVQUAL tool to be valid and reliable. Our correlation results showed that the five constructs of SERVQUAL were not independent and, thus, they may not be measuring different concepts. This implies that if the five SERVQUAL constructs were explanatory variables, we would not need to use all of them. Using all the five constructs may lead to multi-collinearity. The SERVQUAL tool is a good measure for student satisfaction (SS). We, therefore, recommend its usage in SS studies. Scholars should continuously test it in other contexts with the objective of improving it.

# References


Abdullah, F. (2006). The development of HEdPERF: A new measuring instrument of service quality for the higher education sector. *International Journal of Consumer Studies, 30*(6), 569–581. doi: 10.1111/j.1470-6431.2005.00480.x

Abu-Rumman, A., & Qawasmeh. R. (2022). Assessing international students' satisfaction of a Jordanian university using the service quality model. *Journal of Applied Research in Higher Education*. doi:10.1108/JARHE-05-2021-0166

Bell, K. (2022). Increasing undergraduate student satisfaction in higher education: The importance of relational pedagogy. *Journal of Further and Higher Education*. *46*(4), 490–503. doi: 10.1080/0309877X.2021.1985980

Cerri, S. (2012). Assessing the quality of higher education services using a modified SERVQUAL scale. *Annales Universitatis Aputensis Series Oeconomica*, *14*(2), 664–679.

Cronin, J. J. J., & Taylor, A. S. (1992). Measuring service quality: A reexamination and extension. *Journal of Marketing, 56*, 55–68. doi: 10.2307/1252296

Elliot, M. K., & Shin, D. (2002). Student satisfaction: An alternative approach to assessing this important concept. *Journal of Higher Education Policy and Management*, *24*(2), 197–209. doi: 10.1080/13600800/22000013518

Eres, F., & Clothey, R. (2013). Perceived service quality in schools of education: A comparative study between Drexel University and Gazi University. *American International Journal of Social Science, 2*(7), 6–15. Retrieved from www.aijssnet.com/journals/vol_2_7_october 2013/2.pdf

Green, P. (2014). Measuring service quality in higher education: A South African case study. *Journal of International Education Research, 10*(2), 1–12.

Guilbault, M. (2017). Students as customers in higher education: The (controversial) debate needs to end. *Journal of Retailing and Consumer Services, 40*, 295–298. doi: 10.1016/j.jretconser.2017.03.006

Kanwar, A., & Sanjeeva, M. (2022). Student satisfaction survey: A key for quality improvement in higher education institutions. *Journal of Innovation and Entrepreneurship*. doi: 110.1186/s13731-022-00196-6

Karna, S., & Julin, P. (2015). A framework for measuring student and staff satisfaction with university campus facilities. *Quality Assurance in Education, 23*(1), 47–66. doi: 10.1108/QAE-10-2013-0041

Ladhari, R. (2009). A review of twenty years of SERVQUAL research. *International Journal of Quality and Social Services, 1*(2), 172–198. doi: 10.1108/175666910971448

Landrum, H., Prybutok, V., Zhang, X., & Peak, D. (2009). Measuring IS system service quality with SERQUAL: User's perceptions of relative importance of the five SERVQUAL dimensions. *Informing Science. The International Journal of an Emerging Transdiscipline, 12*, 17–35.

Leonnard (2018). The performance of SERVQUAL to measure service quality in private university. *Journal on Efficiency and Responsibility in Education and Science, 11*(1), 16–21. doi: 10.7160/eriesj.2018.110103.

Mihanovic, Z., Batinic, B.A., & Pavicic, J. (2016). The link between students' satisfaction with faculty, overall students' satisfaction with life and student performances. *Review of Innovation and Competitiveness, 2*(1), 37–60. Retrieved from www.researchgate.net/publication/320243579

Mvududu, N. H., & Sink, C. A. (2013). Factor analysis in counseling research and practice. *Counseling Outcome Research and Evaluation*, *4*, 75–98. doi: 10.1177/2150137813494766

Nyandwi, E., & Bakkabulindi, F. E. K. (2018). Satisfaction of international undergraduate students at Kampala International University: Application of the SERVQUAL model. *East African Researcher, 7*(1), 1–17.

Osman, R., A., & Saputra, S., R. (2018). A pragmatic model of student satisfaction: A view point of private higher education. *Quality Assurance in Education, 27*(2), 142–165. doi: 10.1108/QAE 05-2017-0019

Parasuraman, A., Zeithaml, A. V., & Berry, L. L. (1985). A conceptual model of service quality and its implication for future research. *The Journal of Marketing*, *49*(4), 41–50. doi: 10.237/1251430

Parasuraman, A., Zeithaml, A. V., & Berry, L. L. (1988). SERVQUAL: A multiple-item scale for measuring customer perceptions of service quality. *The Journal of Retailing*, 64(1), 12–40. Retrieved from https://www.researchgate.net/publication/225083802

Parasuraman, A., Zeithaml, A. V., & Berry, L. L. (1991). Refinement and reassessment of the SERVQUAL scale. *The Journal of Retailing*, *67*(4), 420–450. Retrieved from https://www.researchgate.net/publication/247373898

Schmidt, D. A., Baran, E., Thompson, A. D., Mishra, P., Koehler, M. J., & Shin, T. S. (2009). Technological pedagogical content knowledge (TPACK): The development and validation of an assessment of instrument for pre-service teachers. *Journal of Research on Technology in Education, 42*, 123–150.

Shahsavar, T., & Sudzina, F. (2017). Student satisfaction and loyalty in Denmark: Application of ESPI methodology. *PLoS ONE, 12*(12), 1–18. Retrieved from https://doi.org/10.1371/journal.pone.0189576

Silva, S. D., Moraes, M. S. H. G., Makiya, K. I., & Ceasar, G. I. F. (2017). Measurement of perceived service quality in higher education institutions: A review of HEdPERF scale use. *Quality Assurance in Education*. doi: 10.1108/QAE-10-2016-0058

Sweet, S. A., & Grace-Martin, K. (2003). *Data analysis with SPSS: A first course in applied statistics* (2nd ed.). Boston, US: Pearson Education Inc.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53–55. doi: 10.5116/ijme.4dfb.8dfd

Tegambwage, G. M. A., & Ame, M. A. (2017). An assessment of the reliability and validity of the SERVQUAL scale in the higher education context in Tanzania. *The Pan-African Journal of Business Management, 1*(1) 83–100.

Wong, W.H., & Chapman, E. (2022). Student satisfaction and interaction in higher education. *Higher Education*. 85, 957–978. Doi: 10.1007/s10734-022-00874-0

Vaz, A., & Mansori, S. (2013). Malaysian private quality: Application of SERVQUAL model. *International Education Studies, 6*(4), 164–170. doi: 10.5539/ies.v6n4p164

# Appendix

## Section A: Background Variables

This section contains items about you. Tick or fill in the correct response against each item.

| | | | | | | |
|---|---|---|---|---|---|---|
| BV1 | Gender | 1. Female | 2. Male | | | |
| BV2 | Nationality | 1. Ugandan | 2. International | | | |
| BV3 | Region if Ugandan | 1. Central | 2. Eastern | 3. Northern | 4. Western | |
| BV4 | Age to the nearest whole year | 1. Below 20 | 2. 20 but below 25 | 3. 25 and above | | |
| BV5 | University of Study | …… | | | | |
| BV6 | Level of study | 1. Undergraduate Diploma | 2. Bachelors | 3. Postgraduate Diploma | 4. Masters | 5. PhD |
| BV7 | Year of Study | 1. First | 2. Second | 3. Third | 4. Fourth | 5. Fifth and above |
| BV8 | Program of Study (e.g., Bachelor of Science with Education) | | | | | |

## Section B: Dependent Variable: Student Satisfaction (SS)

| Strongly Disagree (SD) | Disagree (D) | Undecided (U) | Agree (A) | Strongly Agree (SA) |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |

| | Tangibles (T) | SD | D | U | A | SA |
|---|---|---|---|---|---|---|
| T1 | My University has modern looking equipment | 1 | 2 | 3 | 4 | 5 |
| T2 | The physical facilities in my University are visually appealing | 1 | 2 | 3 | 4 | 5 |
| T3 | The academic staff in my University appear neat | 1 | 2 | 3 | 4 | 5 |
| T4 | The administrative staff in my University appear neat | 1 | 2 | 3 | 4 | 5 |
| T5 | The printed materials (e.g., admission letters, registration forms, brochures) in my University are visually appealing | 1 | 2 | 3 | 4 | 5 |
| | Reliability (Rel) | SD | D | U | A | SA |
| Rel1 | When academic staff in my University promise to do something for me by a certain time, they do it | 1 | 2 | 3 | 4 | 5 |
| Rel2 | When administrative staff in my University promise to do something for me by a certain time, they do it | 1 | 2 | 3 | 4 | 5 |
| Rel3 | When I have a problem, an academic staff in my University shows interest in solving it | 1 | 2 | 3 | 4 | 5 |
| Rel4 | When I have a problem, an administrative staff in my University shows interest in solving it | 1 | 2 | 3 | 4 | 5 |
| Rel5 | Academic staff in my University perform services for me right the first time | 1 | 2 | 3 | 4 | 5 |
| Rel6 | Administrative staff in my University perform services for me right the first time | 1 | 2 | 3 | 4 | 5 |
| Rel7 | Academic staff in my University provide services at the time they promise me to do so | 1 | 2 | 3 | 4 | 5 |
| Rel8 | Administrative staff in my University provide services at the time they promise me to do so | 1 | 2 | 3 | 4 | 5 |
| Rel9 | The academic records given to me by academic staff in my University are error free | 1 | 2 | 3 | 4 | 5 |
| Rel10 | The academic records given to me by administrative staff in my University are error-free | 1 | 2 | 3 | 4 | 5 |
| | Responsiveness (Res) | SD | D | U | A | SA |
| Res1 | Academic staff in my University tell me exactly when a service will be performed | 1 | 2 | 3 | 4 | 5 |
| Res2 | Administrative staff in my University tell me exactly when a service will be performed | 1 | 2 | 3 | 4 | 5 |
| Res3 | Academic staff in my University give me prompt response whenever necessary | 1 | 2 | 3 | 4 | 5 |
| Res4 | Administrative staff in my University give me prompt response whenever necessary | 1 | 2 | 3 | 4 | 5 |
| Res5 | Academic staff in my University are willing to help me whenever need arises | 1 | 2 | 3 | 4 | 5 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Res6 | Administrative staff in my University are willing to help me whenever need arises | 1 | 2 | 3 | 4 | 5 |
| Res7 | Academic staff in my University are not too busy to respond to my requests | 1 | 2 | 3 | 4 | 5 |
| Res8 | Administrative staff in my University are not too busy to respond to my requests | 1 | 2 | 3 | 4 | 5 |
| | Assurance (A) | SD | D | U | A | SA |
| A1 | The behaviour of academic staff in my University instils confidence in me | 1 | 2 | 3 | 4 | 5 |
| A2 | The behaviour of administrative staff in my University instils confidence in me | 1 | 2 | 3 | 4 | 5 |
| A3 | I feel safe dealing with academic staff in my University | 1 | 2 | 3 | 4 | 5 |
| A4 | I feel safe dealing with administrative staff in my University | 1 | 2 | 3 | 4 | 5 |
| A5 | Academic staff in my University are polite to me | 1 | 2 | 3 | 4 | 5 |
| A6 | Administrative staff in my University are polite to me | 1 | 2 | 3 | 4 | 5 |
| A7 | Academic staff in my University have knowledge to answer my inquiries | 1 | 2 | 3 | 4 | 5 |
| A8 | Administrative staff in my University have knowledge to answer my inquiries | 1 | 2 | 3 | 4 | 5 |
| | Empathy (E) | SD | D | U | A | SA |
| E1 | I receive attention from academic staff in my University when I need it | 1 | 2 | 3 | 4 | 5 |
| E2 | I receive attention from administrative staff in my University when I need it | 1 | 2 | 3 | 4 | 5 |
| E3 | The operating hours of academic staff in my University are convenient for me | 1 | 2 | 3 | 4 | 5 |
| E4 | The operating hours of administrative staff in my University are convenient for me | 1 | 2 | 3 | 4 | 5 |
| E5 | Academic staff in my University are concerned about me | 1 | 2 | 3 | 4 | 5 |
| E6 | Administrative staff in my University are concerned about me | 1 | 2 | 3 | 4 | 5 |
| E7 | Academic staff in my University have my interests at heart | 1 | 2 | 3 | 4 | 5 |
| E8 | Administrative staff in my University have my interests at heart | 1 | 2 | 3 | 4 | 5 |
| E9 | Academic staff in my University understand my specific needs | 1 | 2 | 3 | 4 | 5 |
| E10 | Administrative staff in my University understand my specific needs | 1 | 2 | 3 | 4 | 5 |