Application of Likert Scale’s Type and Cronbach’s Alpha Analysis in an Airport Perception Study

Adeniran Adetayo Olaniyi

Abstract

This study explains the basis of Likert scale’s type and Cronbach’s Alpha analysis in determining the reliability and inter-item consistency of questions regarding the assessment of passengers’ satisfaction and service quality of Murtala Muhammad Airport 2. The essence of adopting Cronbach’s Alpha coefficient for checking the reliability and internal consistency of Likert-type scales cannot be overemphasized as it does not provide reliable estimates for single items or individual items, but for summated scales or subscales so as to have a known reliability of the items. Likert scale’s type is suitable with studies in social and behavioural sciences that have to do with perceptions, attitudes, emotions, opinions, personalities, and descriptions of people’s environment. It was revealed from the survey of 114 respondents (air passengers) that the reliability test for statements of thirty-nine airport services blended into five service quality attributes for service quality and passengers’ satisfaction was 0.893, and 0.861 respectively. The two Alpha values indicate high reliability of questionnaire instrument and internal consistencies of the five-point Likert-type scales. It was recommended that for every perception study particularly in airport study, the reliability and consistency of the questionnaire instrument enhance the reliability of results.

Keywords: Likert scale’s type, Cronbach’s Alpha analysis, Perception study.

Introduction

Background to the study

Gathering information in social and behavioural sciences has to do with perceptions, attitudes, emotions, opinions, personalities, and descriptions of people’s environment. This can be achieved with the use of Likert-type scales. According to Joseph and Rosemary [1], as researchers attempt to quantify constructs which are not directly measurable, they oftentimes use multiple-item scales and summated ratings to quantify the construct(s) of interest. The Likert scale’s invention is attributed to Rensis [2], who described this technique for the assessment of perceptions.

Reliability is the tendency towards consistency found in repeated measurements of the same phenomenon, while internal consistency is the extent to which all of the items in a scale measure the different aspects of the same attribute. Reliability is the degree to which data collection (tools and techniques) produces consistent results when the unit being measured has not changed. There are numerous synonyms to describe reliability such as ‘dependable’, ‘consistent’, ‘stable’, ‘trustworthy’ and ‘predictable’. Saunders et al. [3] reveals that reliability is the extent to
which data collection techniques or analysis procedures will yield consistent findings. It can be assessed by posing the questions regarding if the measure will yield the same result on the occasions? If the measure of similar observations be reached by other observers? If there is transparency in the result generalized from the raw data? The Cronbach’s alpha helps to determine the reliability as it measures the internal consistency of a set of items comprising a scale. The closer the Cronbach’s alpha coefficient is to 1.0, the greater the internal consistency of the items in the scale will be.

The aim of this study is to explain the basis for applying Likert scale’s type and Cronbach’s Alpha analysis in a perception study. The objectives explain Likert type scales and explicate Cronbach’s Alpha measurement of internal consistency reliability of airport questionnaire instrument.

**Literature Review**

**Likert type scales**

Likert scales are described as the set of items, composed of approximately an equal number of favourable and unfavourable statements concerning the attitude object that is given to a group of subjects [4]. Respondents were asked to respond to each statement in terms of their own degree of agreement or disagreement. Typically, there are different responses among are; one of seven, one of five, one of four, and one of three responses which respondents are instructed to select: strongly agree, agree, undecided, disagree, or strongly disagree. According to Joseph and Rosemary [1], the specific responses to the items are combined so that individuals with the most favourable attitudes will have the highest scores while individuals with the least favourable (or unfavourable) attitudes will have the lowest scores, and not all summated scales are created according to Likert’s specific procedures; all such scales share the basic logic associated with Likert scaling.

**Conditions for regarding a scale as a summated rating scale**

Spector [5] identified four characteristics that make a scale a summated rating scale which is as follows:

- A scale must contain multiple items. This implies that multiple items will be combined or summed.
- Each individual item must measure something that has an underlying quantitative measurement scale. In other words, it measures a property of something that can vary quantitatively rather than qualitatively.
- Each item has no “right” answer, which makes the summed rating scale different from a multiple-choice test. Thus, summed rating scales cannot be used to test for knowledge or ability.
- Each item in a scale is a statement, and respondents are asked to give rating about each statement. This involves asking subjects to indicate which of several response choices best reflects their response to the item.

**Reasons for using multi-item measures instead of a single item**

According to Joseph and Rosemary [1], Nunnally and Bernstein [6], Spector [5], and McIver and Carmines [4], the reasons for using multi-item measures instead of a single item for measuring psychological attributes or perceptions are:

- Individual items have considerable random measurement error, i.e. are unreliable.
- An individual item can only categorize respondents into a relatively small number of groups, i.e. they lack precision.
- It is very unlikely that a single item can fully represent a complex theoretical concept or any specific attribute for that matter” i.e. they lack scope.

The most fundamental problem with single item measures is not merely that they tend to be less valid, less accurate, and less reliable than their multi-item equivalents. It is rather that the social scientist rarely has sufficient information to estimate their measurement properties. Thus, their degree of validity, accuracy, and reliability is often unknowable [7]. With a single measure of each variable, one can remain blissfully unaware of the possibility of measurement (error), but in no sense will this make his inferences more valid. Given this brief background on the benefits of Likert-type scales with their associated multi-item scales and summated rating scores, many individuals consistently invalidate research findings due to improper data analysis.

**Cronbach’s alpha analysis**

Cronbach’s Alpha was developed to meet the need of finding an objective way of measuring the internal consistency reliability of an instrument used in a research work [8]. It is mostly used when the research being carried out has multiple-item measures of concept [9].

Joseph and Rosemary [1] concluded in their paper that when using Likert-type scales, it is imperative to calculate and report Cronbach’s Alpha coefficient for internal consistency reliability for any scales or subscales that a study is adopting. The analysis of the data then must use these summated scales or subscales and not individual items. If otherwise done, the reliability of the items is at best probably low and at worst unknown; hence Cronbach’s Alpha does not provide reliable estimates for single items.

The value of Cronbach’s Alpha is usually expressed as a number between 0.00 and 1.0. A value of 0.00 means no consistency in measurement while a value of 1.0 indicates perfect consistency in measurement [10]. The acceptable range is between 0.70 and 0.90 or higher depending on the type of research. Cronbach’s Alpha of 0.70 is acceptable for exploratory research while 0.80 and 0.90 are acceptable for basic research and applied scenarios respectively [10,11].

Furthermore, the number of items used on a scale usually affects the estimated reliability. A low value (e.g. <0.5) could be as a result of factors such as a low number of questions
or poor interrelatedness between items, while a high value of alpha (e.g. >0.90), maybe as a result of some redundant items in the instrument [11].

Methodology

Study area

Murtala Muhammed Airport Terminal 2 (MMA2) is the first and only privately-funded, as well as the preeminent terminal in Nigeria. It was conceived after the fire gutted the domestic terminal of the Murtala Muhammed Airport, Lagos, on May 10, 2000. The terminal had been built in the pre-independence era, and before the construction of the International terminal, to cater for both international and regional flights.

After the inferno, the Federal Government of Nigeria made a decision to redevelop the airport using private sector investment under a Public-Private Partnership Scheme. The plan completely transferred all development and operating risks to the private sector, specifically on a Build-Operate-Transfer (BOT) arrangement. There were competitive bids from several companies for the project. A company named Royal Sanderson emerged the preferred bidder, while Bi-Courtney Limited, a wholly-indigenous conglomerate and the parent company of Bi-Courtney Aviation Services Limited (BASL), was the reserved bidder. But owing to prolonged delays in commencing the project, the Federal Government invited Bi-Courtney to take up the responsibility.

Consequently, in 2003, the Federal Government awarded the concession to design, build and operate MMA2 and ancillary facilities to Bi-Courtney Limited. Bi-Courtney set out to work promptly with the goal of building a world-class Airport Terminal that would be the pride of Nigerians and promote Lagos as the major hub in Africa. The company, however, suffered the pains of being the pioneer of the BOT arrangement in Nigeria, given the fact that the idea was novel. The attendant challenges associated with funding huge projects on a long-term financing in Nigeria also required the strength and determination often associated with Bi-Courtney.

Despite all these challenges, the company remained undeterred, even when it was forced to scale down the scope of the project considerably. While syndication was being processed, Bi-Courtney commenced work due to the belief, patriotism, and determination of the promoters of the company. The project was funded from two sources:

- Equity of the owners/proprietor and
- The loans from the banks which were syndicated from six banks.
- The terminal eventually commenced operations on May 7, 2007 [12].

Research design

This is a survey research which explores only primary data in reliability and internal consistency of passengers’ satisfaction and service quality in the international and domestic terminals of Murtala Muhammed Airport (MMA2), Lagos, Nigeria. The target populations of this research study were domestic passengers. For reliability and internal consistency, the reliability analysis of thirty-nine questions for passengers’ satisfaction and thirty-nine questions for service quality was carried out.

To determine the appropriate sample size for large (infinite) population and an uncertain number of populations, judgment was made about the confidence level and the maximum error allowance. The equation below was applied [13]. The sample size for each terminal was determined.

\[ n = \frac{Z^{2} \times \bar{p} \times (1 - \bar{p})}{E^{2}} \]

where;

\( n = \text{Sample size for MMA2} \)
\( Z = \text{Z score for the 92 percent level of confidence is 1.75} \)
\( E = \text{Maximum acceptable error=0.08} \)

92 percent Confidence level at 0.08 maximum error was chosen because of the time consciousness of air passengers. When inserting the above values into the sample size equation, it resulted in a sample size of 120 questionnaires for distribution (Table 1).

The sampling technique is a purposive (non-probability) sampling. According to Henry [14]; Saunders et al. [3], convenience sampling is also referred to as grab sampling, accidental sampling, opportunity sampling, or availability sampling. It is a type of non-probability sampling that involves the sample being drawn from part of the population that is close to hand or easy to reach. According to Adeniran [15,16], there are no other criteria for the sampling method except that people or respondents are available and willing to participate.

This is appropriate for this study because of time consciousness of air passengers in the airport, and limitation/constraint of resources regarding questionnaire distribution and response. Primary data were collected for one week and three days (19th August to 28th August 2017). The study sought to gather information from airport passengers.

Table 2 shows that a total of one hundred and twenty (120) questionnaires were distributed to domestic passengers in MMA2 and 114 questionnaires were returned valid having been filled completely. According to Mugenda and Mugenda [17], a response rate of 50 percent is adequate for data analysis and reporting; a rate of 60 percent is good and a response rate of 70 percent and over is excellent. This implies that a 95 percent response rate for this study was excellent for data analysis and reporting.

<table>
<thead>
<tr>
<th>Terminal</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Airport passengers in MMA2</td>
<td>120</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Questionnaires</th>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Questionnaires returned</td>
<td>114</td>
<td>95</td>
</tr>
<tr>
<td>Questionnaires not returned</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>120</td>
<td>100</td>
</tr>
</tbody>
</table>
Benchmarking airport operational performance

SKYTRAX uses a ranking system for its passengers’ satisfaction surveys based on the following thirty-nine (39) product and service factors or indicators. In order to measure the reliability and internal consistency of the questionnaire instrument, a list of thirty-nine (39) airport service factors were determined for passengers’ satisfaction and service quality respectively. The thirty-nine services blended into service quality attributes are summarized below:

**Tangibles:** These are the physical facilities and equipment available in the airport, the appearance of airport staff; how easy it is to understand communication materials.

**Reliability:** This is the ability of the airport to perform the promised airport service dependably and accurately.

**Responsiveness:** This is the willingness of the airport employees to help airport passengers and providing a prompt service.

**Assurance:** This is the ability of airport employees to convey trust and confidence in the passengers, such as; competence to perform the service, politeness, and respect for the passengers.

**Empathy:** This is the act by which the airport provides caring, individualized attention provided to airport customers [18].

There are thirty-nine airport services blended into five service quality attributes, they are;

- For reliability attributes, the airport services are Efficiency for availability of public transport options; Taxi availability and prices; Immigration and queuing times; Prevent lost luggage services; Security and safety standards; Ease of transit through the airport; Smoking policy and standard of smoking lounges; Standard of physically impaired facilities; and Priority baggage delivery efficiency.

- For assurance attributes, the airport services are Immigration staff attitude; Courtesy and attitude of security staff; Waiting times at security screening; and Friendliness of airport staff.

- For tangible attributes, the airport services are Getting to and fro airport with ease; Availability of luggage trolleys; Terminal comfort, ambiance, general designs and appearance; Seating facilities throughout terminal; Washroom and shower facilities; Television and entertainment facilities; Quiet areas, day rooms, rest area, hotel facilities; Children play area facilities; Check-in, and queuing facilities; Location of airline lounges; Internet facilities and WIFI availability; Telephone and fax location; Bureau de change facility; and ATM facility.

- For empathy attributes, the airport services are Cleanliness of terminal, floor, seating, and public area; Flight information, screen clarity and quality of information; Clarity of boarding calls, and airport public announcement; Cleanliness of washroom facilities; and Terminal signage facilities, boarding gates, transfer, and arrivals.

- For responsiveness attributes, the airport services are Language skills for airport staff; Choice of shopping, tax-free and other outlets; Prices charged in retail outlets; Choice of bars, cafes, and restaurants, including international options.

**Model specification**

The formula for Cronbach’s Alpha is

\[
\alpha = 1 - \frac{kr}{(1 + (k-1)r)}
\]

Where \( k \) is the number of indicators or number of items;

\( r \) is the mean inter-indicator correlation;

The value that is obtained for \( \alpha \) usually indicates the percentage of the reliable variance. An example is the value of 0.80, which means that 80% of the variance in the scores is reliable variance and that 20% is error variance.

Cronbach’s alpha does not provide reliable estimates for single items or individual items, but for summated scales or subscales. If otherwise done for single items or individual items, the reliability of the items will be unknown. This is suitable with studies in social and behavioural sciences that have to do with perceptions, attitudes, emotions, opinions, personalities, and descriptions of people’s environment.

The Likert scale’s invention is attributed to Rensis [2], who described this technique for the assessment of perceptions. There are different responses among are; one of seven, one of five, one of four, and one of three responses which respondents are instructed to select: strongly agree, agreed, undecided, disagree, or strongly disagree. Cronbach’s Alpha was developed in 1951 by Cronbach Lee to meet the need of finding an objective way of measuring the internal consistency reliability of an instrument used in a research work. The value of Cronbach’s Alpha is usually expressed as a number between .00 and 1.0. A value of .00 means no consistency in measurement while a value of 1.0 indicates perfect consistency in measurement. The acceptable range is between 0.70 and 0.90 or higher depending on the type of research. Cronbach’s Alpha values of 0.7 and above indicates that all the variable indices dimensions demonstrate acceptable internal consistency.

**Results and Discussions**

It was revealed that the reliability test for statements of thirty-nine airport services blended into five service quality attributes for service quality and passengers’ satisfaction was 0.893, and 0.861 respectively. For all service dimensions, the Alpha values indicate high reliability of questionnaire instrument and internal consistencies of the five-point Likert-type scales as shown in Tables 3 and 4.

**Table 3:** Result of a reliability test for airport service quality.

<table>
<thead>
<tr>
<th>Service Dimension</th>
<th>Cronbach’s Alpha Values</th>
<th>No. of Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>R Reliability</td>
<td>0.790</td>
<td>15</td>
</tr>
<tr>
<td>A Assurance</td>
<td>0.757</td>
<td>10</td>
</tr>
<tr>
<td>T Tangibles</td>
<td>0.812</td>
<td>4</td>
</tr>
<tr>
<td>E Empathy</td>
<td>0.823</td>
<td>5</td>
</tr>
<tr>
<td>R Responsiveness</td>
<td>0.881</td>
<td>4</td>
</tr>
<tr>
<td>Overall</td>
<td>0.893</td>
<td>39</td>
</tr>
</tbody>
</table>
Conclusion and Recommendation

The essence of adopting Cronbach’s Alpha coefficient for checking the reliability and internal consistency of Likert-type scales cannot be overemphasized. For all service dimensions, the Alpha values indicate high reliability of questionnaire instrument and internal consistencies of the five-point Likert-type scales. It was recommended that for every perception study, the reliability and consistency of the instrument must be carried out so that the results emanating from such study will be dependable.

References

1. Joseph AG, Rosemary RG (2003) Calculating, interpreting, and reporting Cronbach’s Alpha reliability coefficient for Likert-type scales. Midwest Research to Practice Conference in Adult, Continuing, and Community Education. The Ohio State University, Columbus, OH; pp: 82-88.
12. Bi-Courtney Aviation Services Limited (BASL).

Table 4: Result of the reliability test for customers’ satisfaction.

<table>
<thead>
<tr>
<th>Service Dimensions</th>
<th>Cronbach’s Alpha Values</th>
<th>No. of Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>R Reliability</td>
<td>0.753</td>
<td>15</td>
</tr>
<tr>
<td>A Assurance</td>
<td>0.821</td>
<td>10</td>
</tr>
<tr>
<td>T Tangibles</td>
<td>0.921</td>
<td>4</td>
</tr>
<tr>
<td>E Empathy</td>
<td>0.736</td>
<td>5</td>
</tr>
<tr>
<td>R Responsiveness</td>
<td>0.880</td>
<td>4</td>
</tr>
<tr>
<td>Overall</td>
<td>0.861</td>
<td>39</td>
</tr>
</tbody>
</table>