A comparative analysis of the different measurement scales of service quality


Abstract
This research analyses the quality of university services measurement at Málaga University. This approach leads to determine the quality service dimensions in this area. Measuring the quality and identifying the components that make it up, we have used the most generally accepted measurement tools: Servqual, Servperf, Modified Servqual, Revised Servqual, Performance Evaluation and Standardised Quality.
Under these premises, we have issued the following steps: First, hypothesis, objectives and methodology are exposed. Second, University quality services dimensions are set up, based on the comparative analysis of the different quality measurement scales already quoted. Finally, we aim to highlight the most important conclusion we have reached.

Keywords: Quality, service quality, measurement tools, Servqual, Servperf.


Josefa García-Mestanza y Rocío Díaz-Muñoz
Departamento de Economía y Administración de Empresas. Facultad de Ciencias Económicas y Empresariales. Universidad de Málaga. c/ El Ejido s/n 29071 Málaga. e-mail: jgm@uma.es / romu@uma.es
1. Introduction
Nowadays, it is essential for organizations to be competitive, capable of adapting to constant technological changes and of responding to markets’ ever increasing demands. This can only be achieved if organizations pay a considerable amount of attention to quality, as this is fundamental for customer satisfaction and for obtaining continuity and more competitiveness.

In that case, after the company has realized it needs to evaluate the quality of its services regularly and systematically, it has to be aware of the situation at the outset by means of a measurement system.

As the quality of the service offered needs to be measured, the company has to choose the model it considers most suitable, asking which of the existing systems is most suited for the purpose aimed at, their validity and usefulness, the benefits they provide, among other aspects. The objective of this article is to answer these questions.

For that reason, the process followed in this research is set out under three headings, The first provides details of the structure, objectives and methodology of the work to be carried out. The quality dimensions of the university service are obtained using a comparative analysis of the different measurement scales applied to the University of Malaga in the second, and, in the third, a series of conclusions are deduced from the information obtained.

2. Structure, objectives and methodology
2.1. Theoretical framework
In the process of searching for information prior to starting this research, we compiled the most significant European and North American studies on the issue in question. Despite the fact that the quality measurement systems used for services are different from the ones used for products (since the former are not as quantitative and gradual as the latter), there are methods that enable a company’s situation, one of its sectors or one of its acti-
vities to be examined methodically with regards to quality (Judd, 1968), and which also enable the effect of the service quality on the customer's level of satisfaction and involvement to be isolated from the effect of the service’s other components (product range, prices, distribution, advertising, promotions...). The common characteristic that all these methods need to have is to be fair and to be seen as such. Otherwise, their results are not trusted.

Pursuant to Brogowicz, Delene and Lyth’s outline (1990), two currents are differentiated in measuring quality: the Nordic school based around the Service Research Center of Karlstad University and the North American school led by the Interstate Center for Services Marketing of Arizona State University.

Those who stand out in the Nordic school, which focuses on the service quality issue from the product’s point of view, are Lehtinen and Lehtinen (1982) and Grönroos (1984, 1988 and 1994).

Lehtinen and Lehtinen proposed that there are three dimensions: physical quality, which includes the service’s physical aspects; corporate quality, which is associated with the company image, and interactive quality, which derives from the interaction between the company’s contact personnel and its customers, as well as the interaction between some customers and others. Furthermore, they differentiate between quality associated with the service provision process and quality associated with the outcome of the service.

This differentiation is also used by Grönroos, who suggested two service quality dimensions: process quality and outcome quality. But both are influenced by the corporate image. Process quality is judged by the consumer when the service is being provided and it encompasses technical and functional qualities. On the other hand, outcome quality is judged by the consumer after the service has been provided and it cannot always be controlled by the company, since it is susceptible to external and environmental influences, which act as a filter to judge the quality of an organization.
Those who stand out in the North American school, which has centered on studying service quality from the viewpoint of customers’ expectations and perceptions, are Berry, Parasuraman and Zeithaml (1985 and 1988).

Berry, Parasuraman and Zeithaml developed their work in stages. First they defined service quality, and then they created a measurement model, the so-called Servqual, which they later corrected. For these authors, the main indicator of the quality level of a service is the customer who receives it and this is based on two factors: his experience and his expectations. In other words, a customer thinks that a service is top quality when his specific experience with the company is the same as or exceeds his initial expectations. The challenge for the company is to eliminate everything that is not in line with this.

Servqual Model: The Servqual model, developed by Parasuraman, Zeithaml and Berry (1988), is the most widely known in the services sector. These authors’ initial aim was to create a scale of multiple items in order to measure service quality, as well as to discuss their properties and potential applications.

Therefore, once the service quality had been described as the discrepancy between expectations and perceptions, they identified ten dimensions forming the frame of reference for this concept. They used 97 items (approximately 10 per dimension) to describe these dimensions, which were rated by customers based on their expectations and perceptions, using a Likert scale, which varied from “strongly agree” (7) to “strongly disagree” (1).

Later, after market research of users of five different types of services, they proceeded to refine the scale, decreasing the number of items used, as well as the number of dimensions, which went down to five. This version of the model was tested by means of a mail survey, which enabled the instrument to be refined even further down to 22 items and to confirm its reliability and validity.

Servperf Model: Cronin and Taylor (1992) researched the measuring capacity of a more concise scale than the above-mentioned one, based...
exclusively on performance. They did this by taking the items directly from the Servqual model. This scale was called the Servperf model.

According to these authors, the conceptualization of the Servqual scale is inadequate, as it has very little theoretical support. The literature in services marketing seemed to offer considerable support to the idea that service quality measurements based on performance are superior. In addition, they made an empirical comparison\(^7\) between the superiority of the model they proposed and other alternative ones\(^8\).

In view of these criticisms, Parasuraman, Zeithaml and Berry (1994) responded by stating that there is strong theoretical support for the general notion that a customer’s assessments after a stimulus invariably occur in connection with some standard. They also questioned some methodological aspects of the empirical work they developed.

**Modified Servqual model:** Teas (1993) made a series of comments on the Servqual model, basically centered on the interpretation of the “expectations” standard and how this works, to later develop and evaluate alternative models.

In connection with the first aspect, this author suggests that the increase in the P-E difference may not necessarily reflect an increase in perceived quality levels.

Zeithaml, Parasuraman and Berry (1993) recognize that the P-E concept can be problematic for some types of attributes and under certain conditions. Nevertheless, the P-E measurement that Servqual uses is valid if vector attributes are considered, with ideal infinite points, for which high performances are always better. The problem arises when the attribute has a classic ideal point, in other words, the customer’s ideal point is a finite level.

Teas suggests two interpretations of the ideal standard:

- Using the classical attitudinal model: which he considers to be the service level (I) after which the customer experiences non-usefulness. This is not Servqual’s interpretation, since if P>E, the higher the P-E
is, the higher the quality level will be. In the event that \( I = E \), when \( P \) is more than \( I \), the quality observed by the customer will decrease.

- Feasible ideal point: which represents the level of service the customer considers possible for the best company. But in appropriate conditions, the company can exceed this standard up to a level marked by the classic ideal point. Following this meaning, Teas proposed the Modified Servqual model, according to which \( MQ = -1 \cdot [|P - I| - |E - I|] \). Given that \( E \) (feasible ideal point) can never be more than \( I \) (classic ideal point) and considering that \( P < I \), the equation would be \( MQ = P - E \) which is equivalent to the Servqual model. The problem would arise when \( P > I \), in which case the equation would be \( SQ = (I - E) - (P - I) \).

**Revised Servqual model:** Meanwhile, Parasuraman, Zeithaml and Berry (1994) considered the revised expectations (\( E^* \)), which they defined as the extension (rating) with which customers think that a particular attribute is essential for a company with excellent services. In this way, by including this consideration of the attributes, this model was defined:

**Evaluated Performance Model (EP):** Based on the above, Teas proposed the standard interpretation of the “expectation” standard as the classic ideal point (finite), so the quality was indicated by \( Q = -\left[ \sum_{i=1}^{m} W_i |I_i - 1| \right] \), where \( W_i \) was a factor of importance of the corresponding attribute as a determining factor of the quality perceived and in which \( I_i \) represented the classic ideal point.

**Normalized Quality Model (NQ):** This author goes one step further and associated the concept of the quality of the evaluated performance model with the concept of revised expectations (\( E^* \)), which define the standard of excellence on which the revised SERVQUAL model is based. This is the argument of the new model, which is defined based on \( NQ = Q - Q_e \), where \( Q_e \) is the quality perceived by the customer of the service considered as excellent.
If the expectations indicate by revised Servqual coincide with this standard of excellence, the model is
\[
NQ = -\sum_{i=1}^{n} w_i \left( |p_i - I_i| - |E_i - I_i| \right)
\]
If, in addition, the ideal points are considered infinite
\[
Q = \sum_{i=1}^{n} w_i (p_i - E_i)
\]

**Combined Servqual Model (SQ*)**: Teas’s specification (1993), despite its apparent empirical superiority over other models he tested, is said to have shortcomings due to the questionable presumption that all the aspects of the service are classic ideal point attributes. Parasuraman, Zetikhaml and Berry (1993), after recognizing Servqual’s limitations, proposed a new compound specification (SQ*), which presupposes that some aspects can be vector attributes, while others are classic ideal point attributes.

\[
SQ^* = \sum_{j=1}^{m} \left( (p_j - E_j) D_{1j} - (p_j - E_j) D_{2j} + \left\{ (I_j - E_j) - (p_j - I_j) \right\} D_{3j} \right)
\]

Where \(D_{1j}\) is 1 if \(j\) is a vector attribute or if it is a classic ideal point and \(p_j \leq I_j\) and is 0 in any other case; \(D_{2j}\) is 1 if \(j\) is a classic ideal point attribute and \(E_j\) is interpreted as the classic ideal point and \(p_j > I_j\) and is 0 in any other case; \(D_{3j}\) is 1 if \(j\) is an attribute with a classic ideal point and \(E_j\) is interpreted as a feasible ideal point and \(p_j > I_j\) and is 0 in any other case. This expression means that the same attribute \(j\) can be considered differently by different customers and it can be determined if there are homogenous customer segments.

Teas (1994) responded by stating that the specification of the EP model and the NQ model, to make it suitable under the presumption of attributes with classic ideal points, is not a restrictive presumption. On the contrary, it is more flexible than the vector model, in fact the vector model is a special kind of ideal point model, in other words, a restricted version. The difference between the combined model (SQ*) and the EP model is based, exclusively, on the fact that the latter does not include the expectation concept as a standard, which the former does include. We fully agree with this observation, so from now on, we will not take this model into consideration.
Even when there are these two large currents of thought in the study of service quality, the Nordic school and the North American school, the range of authors that have dealt with this issue is very broad\(^9\). In an attempt to synthesize the most important contributions, we could state that:

- The dimensions of service quality are determined by the customer. Therefore, we cannot speak about generic dimensions\(^10\).
- The measurement models of the North American school are the most widespread in the services sector\(^11\).

### 2.2. Objective of the research

The objective of this research is to determine the dimensions of service quality in the University of Malaga using the most widespread service quality measurement models in the services sector: Servqual, Servperf, modified Servqual, revised Servqual, Evaluated Performance and Normalized Quality.

Prior studies, even though they can be used as a guide and are very valid, were not entirely useful in reaching our research objective, since some of them were carried out quite some time ago, others do not consider all the aspects we wished to cover and there were no instances of a differentiated treatment of the service quality offered by the University of Malaga, the geographical area that is the main focus of our study.

In view of this situation, we realized that we needed to do our own research. In fact, even though we are immersed in a market globalization process at the moment, a study of educational bodies from a territorial point of view is not only valid, but also important, as there is a need to guarantee that the particular characteristics of the different regions and areas can be taken into consideration sufficiently in the new European continental area.

### 2.3. Research methodology

The methodology followed to reach the objective proposed consisted of personal interviews using a structured questionnaire, aimed at users of the services of the University of Malaga.
A COMPARATIVE ANALYSIS OF THE DIFFERENT MEASUREMENT SCALES
OF SERVICE QUALITY

With this aim, we proceeded to design the sample and the questionnaire. The specifications summarize the design of the sample made (Table 1).

Table 1. Specifications

<table>
<thead>
<tr>
<th>Specifications</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Universe</td>
<td>Students at the University of Malaga enrolled in the Faculty of Economics (4061).</td>
</tr>
<tr>
<td>Sample size</td>
<td>383 individuals.</td>
</tr>
<tr>
<td>Sample error</td>
<td>± 10%.</td>
</tr>
<tr>
<td>Confidence level</td>
<td>90%.</td>
</tr>
<tr>
<td>Sampling procedure</td>
<td>Proportional to the number of students enrolled for each course (Centre Total: 383; Degree in Economics: 125; Degree in Business Management: 214; Diploma in Management and Public Administration: 10; and Degree in Actuarial Science and Finance: 34). The sample unit was selected at random.</td>
</tr>
<tr>
<td>Timing of the field work</td>
<td>From 15 April to 31 May 2005.</td>
</tr>
</tbody>
</table>

Source: our own.

For the design of the questionnaire, we took the 22 items in the Servqual model and other aspects required to apply the six measurement scales of the service quality as a basis. In this way, the survey was structured in two parts. The first includes the rating students give the university service for each and every one of the twenty-two aspects that define it, taking the different standards required to apply each method into account, and, the second part includes a global rating of the university service, that is, the level of satisfaction, level of future use of the service, global quality and the extent to which they would recommend the service, as well as any comments they deemed appropriate.

After the initial questionnaire had been designed, we did the pilot tests. We tested it in one and in the other we did a pretest. The first of the tests was carried out on a very small group of people, ten, to be precise, with the aim of refining the questionnaire in its formal aspects. In the second, with twenty people, we piloted the questionnaire in real conditions, testing its fundamental aspects (suitability of the framework proposed to select the sample, suitability of the data collection method...). In the meantime, LeBlanc and Nguyen (1988) identified four components on which customers base their ratings of the service: satisfaction, performance, service aspects and key aspects of the business. On the other hand, Armistead (1989) established a list of dimensions that he grouped into two: firm (where he includes the temporary system, the lack of faults and flexibility) and soft (which encompasses style, management and...).
3. Determining the quality dimensions of the university service using a comparative analysis of the different measurement scales

In order to establish the quality dimensions of the service in the University of Malaga objectively, we compared the different measurement models used beforehand (Servqual, Servperf, modified Servqual, revised Servqual, Evaluated Performance and Normalized Quality) and analyzed their psychometric properties.

3.1. Psychometric properties

The psychometric properties of the measurement scales of the quality of the university service used were analyzed using their reliability and validity properties.

A) Reliability:

A scale must have the property of reliability, in other words, it must be capable of generating constant results in successive applications to the same groups and in similar situations (Fernández Barcala, 2000). Cronbach’s alpha coefficient allows this property to be measured (Table 2).

<table>
<thead>
<tr>
<th>MEASUREMENT TOOLS</th>
<th>Alfa de Cronbach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Servqual</td>
<td>.944</td>
</tr>
<tr>
<td>Servperf</td>
<td>.927</td>
</tr>
<tr>
<td>Modified Servqual</td>
<td>.954</td>
</tr>
<tr>
<td>Revised Servqual</td>
<td>.945</td>
</tr>
<tr>
<td>Evaluated Performance</td>
<td>.940</td>
</tr>
<tr>
<td>Normalized Quality</td>
<td>.955</td>
</tr>
</tbody>
</table>

Table 2. Reliability

Source: our own.

Based on the results obtained, coefficients above 0.9, and bearing in mind that the minimum advised is 0.7, we can confirm that the six scales are reliable.
A) Validity:

As far as the validity property is concerned, it can be analyzed from several perspectives: content validity, predictive validity, convergent validity and discriminant validity.

The six scales have content validity since the determinants of the quality of the services analyzed were selected taking the items contained in the Servqual scale as a reference. In addition, the opinions of the experts who reviewed the questionnaire confirmed this and later, after the refining process, we discovered that the correlation between the points of each item and the total of the points of the rest of the items was high (Díaz Martín et al., 1997), thus confirming their reliability and validity.

In order to analyze the predictive validity, we measured the correlation between each scale and customers’ global satisfaction with respect to the university service they are provided with, since there should be a positive association between satisfaction and perceived quality (Anderson et al., 1994; Cronin and Taylor, 1992).

According to the results obtained (Table 3), the six scales have a correlation with satisfaction, which is considerably more for the measurement obtained directly from perceptions (Servperf).

Table 3. Predictive validity of the scales

<table>
<thead>
<tr>
<th>Correlation coefficient</th>
<th>Global satisfaction</th>
<th>Global quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Servqual</td>
<td>.460</td>
<td>.440</td>
</tr>
<tr>
<td>Servperf</td>
<td>.559</td>
<td>.537</td>
</tr>
<tr>
<td>Modified Servqual</td>
<td>.446</td>
<td>.441</td>
</tr>
<tr>
<td>Revised Servqual</td>
<td>.457</td>
<td>.425</td>
</tr>
<tr>
<td>Evaluated Performance</td>
<td>.472</td>
<td>.425</td>
</tr>
<tr>
<td>Normalized Quality</td>
<td>.442</td>
<td>.422</td>
</tr>
</tbody>
</table>

Significance level = 0.01.

Source: our own.
As far as *convergent validity* is concerned, a high correlation between the six scales is observed, above 0.7 points (Table 4). Servperf and Evaluated Performance are the models with the least convergence with regards to the other models.

### Table 4. Convergent validity of the scales

<table>
<thead>
<tr>
<th>CORRELATION COEFFICIENT</th>
<th>Servqual Model</th>
<th>Servperf Model</th>
<th>Modified Servqual Model</th>
<th>Revised Servqual Model</th>
<th>Evaluated Performance Model</th>
<th>Normalized Quality Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Servqual</td>
<td>1</td>
<td>.744</td>
<td>.986</td>
<td>.964</td>
<td>.780</td>
<td>.949</td>
</tr>
<tr>
<td>Servperf</td>
<td>.744</td>
<td>1</td>
<td>.737</td>
<td>.703</td>
<td>.792</td>
<td>.689</td>
</tr>
<tr>
<td>Modified Servqual</td>
<td>.986</td>
<td>.737</td>
<td>1</td>
<td>.958</td>
<td>.794</td>
<td>.566</td>
</tr>
<tr>
<td>Revised Servqual</td>
<td>.964</td>
<td>.703</td>
<td>.958</td>
<td>1</td>
<td>.794</td>
<td>.966</td>
</tr>
<tr>
<td>Evaluated Performance</td>
<td>.780</td>
<td>.792</td>
<td>.794</td>
<td>.841</td>
<td>1</td>
<td>.844</td>
</tr>
<tr>
<td>Normalized Quality</td>
<td>.949</td>
<td>.689</td>
<td>.966</td>
<td>.988</td>
<td>.844</td>
<td>1</td>
</tr>
</tbody>
</table>

Source: Based on Pruitt (1983).

In order to check the *discriminant validity*, we included an additional question that compiles the recommendations that customers would make of the university service provided. This idea suggests a logical association: the points obtained from the scales used to measure perceived quality must be greater when the service is recommended and the opposite must occur when the service is not recommended.

In Table 5 you can see that in the six instruments measuring quality, the ratings are higher for customers that had recommended the university service than for those that had not done so.

### Table 5. Mean ratings of perceived quality

<table>
<thead>
<tr>
<th>Perceived service quality</th>
<th>Servqual Model</th>
<th>Servperf Model</th>
<th>Modified Servqual Model</th>
<th>Revised Servqual Model</th>
<th>Evaluated Performance Model</th>
<th>Normalized Quality Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Would recommend</td>
<td>-1.5660</td>
<td>4.2929</td>
<td>-1.5809</td>
<td>-1.36</td>
<td>-1.7931</td>
<td>-1.3822</td>
</tr>
</tbody>
</table>

Source: our own.
Furthermore, to analyze this hypothesis, we carried out a contrast analysis of means between the values of quality perceived by customers that would not recommend it, applying Student’s statistic t. In both cases, the results confirmed that the significance level is very low. (0,000<0,001), so the null hypothesis is rejected, referring to the fact that the population means are significantly different and that the six instruments used do not pose any problems for measuring the quality perceived by customers in connection with the university service.

Taking their psychometric properties into account, we can conclude that the six measurement scales of the quality of the university service used are valid and reliable.

3.2. Multidimensional nature of the quality of the university service
After the validity and reliability of the methods used had been checked, we studied the multidimensional nature of the quality of the university service. For that purpose, we used the factor analysis of main components as it is better adapted to the structure of our data, applying the following methodology: analysis of the suitability of the data, selection of factors required to represent the original data and rotating them in order to make interpreting them easier and grouping and interpreting the different factors or dimensions obtained.

A) Analysis of the suitability of the data:
Prior to checking the validity or relevance of applying this multifactor analysis, the level of correlation between the variables is calculated by means of: the correlation matrix between the different variables, Bartlett’s sphericity test, the anti-image correlation matrix and the Kaiser-Meyer-Olkin index (Table 6).
Table 6. Indicators of the degree of correlation between the variables

<table>
<thead>
<tr>
<th></th>
<th>SERVQUAL Model</th>
<th>SERVPERF Model</th>
<th>Modified SERVQUAL Model</th>
<th>Revised SERVQUAL Model</th>
<th>Evaluated Performance Model</th>
<th>Normalized Quality Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation Matrix</td>
<td>91.34% variables correlated, associations are significant at 0.000 level</td>
<td>91.43% variables correlated, associations are significant at 0.000 level</td>
<td>91.77% variables correlated, associations are significant at 0.000 level</td>
<td>90.47% variables correlated, associations are significant at 0.000 level</td>
<td>86.15% variables correlated, associations are significant at 0.000 level</td>
<td>92.64% variables correlated, associations are significant at 0.000 level</td>
</tr>
<tr>
<td>Bartlett’s sphericity test</td>
<td>2768.010 signif. 0.000</td>
<td>2367.642 signif. 0.000</td>
<td>1331.549 signif. 0.000</td>
<td>2842.277 signif. 0.000</td>
<td>2765.430 signif. 0.000</td>
<td>1441.736 signif. 0.000</td>
</tr>
<tr>
<td>Anti-image correlation matrix</td>
<td>Diagonal values very close to 1 and the rest not very high.</td>
<td>Diagonal values very close to 1 and the rest not very high.</td>
<td>Diagonal values very close to 1 and the rest not very high.</td>
<td>Diagonal values very close to 1 and the rest not very high.</td>
<td>Diagonal values very close to 1 and the rest not very high.</td>
<td>Diagonal values very close to 1 and the rest not very high.</td>
</tr>
<tr>
<td>Kaisser-Meyer-Olkin Index</td>
<td>0.934</td>
<td>0.917</td>
<td>0.898</td>
<td>0.931</td>
<td>0.931</td>
<td>0.889</td>
</tr>
</tbody>
</table>

Source: our own.


In order to analyze the correlation between the different variables the correlation matrix was inspected and we discovered that the number of correlations above 0.5 is considerable, although for different authors the value taken as a reference is 0.3 (Luque Martínez, 2000). The matrix inspection revealed that there is a high percentage of considerable associations above 70% for the six scales and that they are significant at the 0.000 level. Consequently, applying the factor analysis is appropriate.

Bartlett’s sphericity test provides the statistical probability of the correlation matrix being an identity matrix, in other words, it is a case of confirming the null hypothesis that the correlation matrix equals one. In the event that the correlation matrix is equal to the identity matrix, there would not be a correlation between the variables, the variables observed would be mutually non-correlated and, therefore, there would be no point in applying the factor analysis. According to Luque Martínez (2000, p. 44), the null hypothesis could no longer be rejected after a certain significance level (0.05) and we would be mistaken more than 5% of the time. In our research, Bartlett’s sphericity test confirmed that the null hypothesis of non-correlated initial variables is not significant in all of the scales,
with a level of 0.000 (less than the level normally taken as a reference, 0.05). Therefore, applying the factor analysis makes sense.

Another way of checking whether this analysis is good or not is using the anti-image correlation matrix. We can observe very few high values in absolute terms, as well as diagonal values very close to one in this matrix.

The value that the Kaiser-Meyer-Olkin statistic or KMO index takes in all the scales is above 0.88 points, which are values very close to the unit, which indicates that our data is excellently adapted to a factor analysis model.

From the results obtained in the tests carried out, we deduced that these data are perfectly adapted to this type of analysis. We should add that the variables used are metric and form a homogenous whole appropriate for the factor analysis.

B) Selection of factors and rotating them:

Having checked the relevance of the factor analysis, the second step was to identify the number of factors or dimensions to be chosen in order to synthesize all the original data. For that purpose we selected three and four factors, depending on the model used, and which explained 62.5% of the total variance of the original variables for the Servqual scale, 58.34% for the Servperf variance, 63.104% for the modified Servqual scale, 58.68% for the revised Servqual scale, 62.12% for the Evaluated Performance scale and 64.12% for the Normalized Quality scale.

According to the data obtained from the factor loading matrix, the correlations of the components with the initial variables cannot be interpreted clearly. Therefore, after discarding the idea of decreasing the number of components, we then rotated them using orthogonal Varimax, Quartimax and Equamax rotations and oblique Promax and Oblimin rotations with different values of the delta coefficient.

After rating them, we opted to apply the factor analysis of main components using the Equamax rotation for the first four scales and the Varimax rotation for the last two.
C) Grouping and interpreting of the different factors:

After a satisfactory factor solution had been obtained, the next step is to give that solution a significance (Table 13).

### Table 13. Grouping of variables in each measurement scale

<table>
<thead>
<tr>
<th>Factor</th>
<th>EQUAMAX</th>
<th>VARIMAX</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Servqual Model</td>
<td>Servperf Model</td>
</tr>
<tr>
<td>Factor 1</td>
<td>18,19,20,21,22,14,15,16,17,11,12,13</td>
<td>12,3,4,5,6</td>
</tr>
<tr>
<td>Factor 2</td>
<td>5,6,7,8,9,10,11,12,13</td>
<td>14,15,16,17,11,12,13</td>
</tr>
<tr>
<td>Factor 3</td>
<td>1,2,4,14,15,16,17</td>
<td>1,2,3,4,7,8</td>
</tr>
<tr>
<td>Factor 4</td>
<td>3</td>
<td>18,19,20,21,22</td>
</tr>
</tbody>
</table>

Source: our own.

Between three and four factors were extracted using the factor analysis depending on the service quality measurement scale used. However, there are coinciding subfactors in each factor in the different measurement scales. In order to verify this classification coincidence, we used the hierarchical cluster analysis, which, applied to the six measurement models, enabled the variables forming part of each one of the groups (number of clusters) to be observed.

By means of this analysis, which has used squared Euclidean distance as a measurement and Ward’s linkage as a grouping method, we can observe that all the variables are grouped very far halfway along the dendrogram, thus corroborating the creation of subfactors in the different fac-
tors obtained by the six measurement scales, except variable 3, which is the last one producing aggregation with the others. This means that the employees’ neat appearance is not perceived as yet another attribute to measure the quality of university service. By way of example, the dendrogram of the Servqual measurement scale is offered, as they all have very similar results (Graph 1).

Graph 1. Hierarchical cluster analysis (dendrogram)

For the purposes of simplifying and not fully identifying with the characteristics they encompass, we have called the subdimensions obtained in the six quality measurement models as follows:

- **Tangibles (1).** Centered mainly on the service components that the customer can perceive with his senses, that is, appearance of the equipment, materials and facilities.
In three of the six scales used, item number 3 “the employees have a neat appearance” is separated from the rest of the elements called “tangibles”. Therefore, in the revised Servqual scales and the Evaluated Performance model, it is included with reliability items and in the Servqual model, item 3 becomes a factor in its own right, independent of any other associated with reliability or tangible elements. Therefore, and as long as we are not faced with extreme cases or which do not meet minimal acceptable conditions of manners and decorum, perhaps the appearance point can be classified as dispensable when determining the quality of the university service, which has been corroborated by the numerous comments made by those surveyed and demonstrated by a hierarchical cluster analysis, as previously argued.

- **Responsibility and empathy (2).** It groups variables 5 and 6, which refer to fulfilling promises made by the company, as well as its interest in solving its customers’ problems.
- **Reliability (3).** It includes aspects associated with the ability to perform the service in a reliable and careful way. It includes the following variables 7 and 8.
- **Formality (4).** The variables that explain the content of this dimension are the company’s interest in not making any mistakes and the employees notifying when the provision of the service will end, in other words, variables 9 and 10.
- **Responsiveness (5).** It is basically associated with the employees’ performance, in other words, their willingness to help and to provide customers with prompt service. It includes the variables 11, 12 and 13.
- **Security (6).** It includes different aspects associated with confidence, courtesy and knowledge employees offer customers, as well as how secure customers feel when they work with the company. It is formed by the variables 14, 15, 16 and 17.
- **Personalized service (7).** It includes from variable 18 to 22 and it refers to the individualized way the company treats its customers,
the concern shown for their interests and needs, as well as the ease in which the service can be accessed.

The mean rating attributed to these seven subfactors is shown in Table 14, where coincidences between the values of some of the measurement scales used can be observed.

Graphically, the similarity of results between the different measurement scales used is even more obvious (Graph 2) and their correlation is very high statistically, above 99%, except in the Servperf method, which is around 85%.

Graph 2. Comparison of subfactors in a decimal scale

Considering test t for associated samples, it can be seen that the greatest association is between: the Servqual measurement model and modified Servqual, followed by revised Servqual and the Normalized Quality model. At a greater distance are modified Servqual and the Evaluated Performance model, as well as the latter and Servqual. Of the six models
analyzed, Servperf is the one that most differs with respect to the mean values it has, although its line has similar ups and downs in its rating.

4. Conclusions
In view of constant technological changes and consumers’ increasing demands, organizations are forced to steer their strategies towards the quality of their services as a means of achieving their continuity in the market and being more competitive. However, after realizing the need to evaluate the quality of their services regularly and systematically, companies have to be aware of the situation at the outset by means of a measurement system.

Two currents are differentiated in service quality measurement systems: the Nordic school based around the Service Research Center of Karlstad University and the North American school led by the Interstate Center for Services Marketing of Arizona State University. From their contributions, as well as those of many other authors, we can deduce that: service quality dimensions are determined by the customer and the measurement models of the North American school are the most widespread in the services sector.

In view of this situation, our work studies the measurement of the quality of the university service by determining its dimensions in the University of Malaga using the most widespread models: Servqual, Servperf, modified Servqual, revised Servqual, Evaluated Performance and Normalized Quality.

After comparing the psychometric properties of the different models used, we demonstrated their validity and reliability, without finding any significant differences between them. We cannot conclude, therefore, that there is a quality measurement model that is more useful than the others.

After the validity and reliability of the methods used had been checked, we studied the multidimensional nature of the quality of the university service. For that purpose, we used the factor analysis of main components, as it is better adapted to the structure of our data, and we checked their validity by means of: the correlation matrix between the different variables,
Bartlett’s sphericity test, the anti-image correlation matrix and the Kaiser-Meyer-Olkin index.

Having checked the relevance of the factor analysis, we identified the number of factors or dimensions to be chosen in order to synthesize all the original data (62.5% of the total variance of the original variables for the Servqual scale, 58.34% for the Servperf scale, 63.104% of the modified Servqual scale, 58.68% for the revised Servqual scale, 62.12% for the evaluated performance scale and 64.12% for the Normalized Quality scale were explained) and we then rotated them using orthogonal Varimax, Quartimax and Equamax rotations and oblique Promax and Oblimin rotations with different values of the delta coefficient to make the dimensions obtained easier to interpret.

Based on the factor analysis and the cluster analysis carried out, we obtained a similar factor structure in the six scales used which, for the purposes of simplifying and not fully identifying with the characteristics they encompass, we call: personalized service, security, tangibles, empathy and responsibility, reliability, formality, and responsiveness. Item number 3 “the employees have a neat appearance” is separated from the rest of the elements. Therefore, it is considered to be dispensable when determining the quality of the university service.

These dimensions proposed are only a partial reflection of the factors identified by the creators of the Servqual scale. New dimensions appeared and some of them were divided. Therefore, bearing the factors extracted in this study in mind and coinciding with the findings of many researchers that have been questioning the generality of service quality dimensions for a long time (Carman, 1990; Buttle, 1995), we can conclude that the factors forming quality in the provision of a service are determined by the customer, in other words, the number of dimensions depends on the type of service provided. Therefore, we cannot speak about generic dimensions.

In the six measurement scales used, not only is the factor structure similar, but its correlation is also above 99% if we speak about the mean values obtained, except in the Servperf method, which is around 85%.
greatest association is between the Servqual measurement model and modified Servqual, followed by revised Servqual and the Normalized Quality model. At a greater distance are modified Servqual and the Evaluated Performance model, as well as the latter and Servqual.

5. References


Chia-Ming, C.; Chin-Tsu, C. y Chin-Hsien, H. (2002): “A Review of Service Quality in Corporate and Recreational Sport/Fitness Pro-


A COMPARATIVE ANALYSIS OF THE DIFFERENT MEASUREMENT SCALES OF SERVICE QUALITY


A COMPARATIVE ANALYSIS OF THE DIFFERENT MEASUREMENT SCALES 
of service quality


