Evaluation to effectively support stakeholders centered design and continuous quality improvement of higher education services

Dr. T. Lupo
Dept. of Chemical, Management, Informatics and Mechanical Engineering
Università degli Studi di Palermo,
Viale delle Scienze, Edificio 8
90128 Palermo,
Italy

Abstract: In the academic contest, evaluation and self-evaluation procedures constitute commonly adopted tools to take under control education services and research performance. In this chapter the main opportunities that evaluation based approaches can produce also to meaningfully support processes of design and continuous quality improvement in higher educational institutions are described, even in the light of the recent Italian normative measures addressed to quality aspects of education services. In addition, a new methodology based on the ServQual conceptual model able to evaluate education services is described and considered to carry out a strategic service analysis related to education services of the Management Engineering program at the University of Palermo (Italy).

Key words: Education services performance; education services management; student satisfaction; ServQual; Fuzzy Analytic Hierarchy Process (FAHP);
1. Introduction

The deep attention to quality aspects of higher educational institutions represents a significant part of the transformation process which is characterizing the entire European area of higher education. A fundamental trait of such transformation process is represented by systematic implementation of evaluation and self-evaluation procedures addressed to quality, effectiveness/efficiency aspects, for several reasons. The first one is related to the fact that evaluation represents an essential part of the continuous quality improvement process (De Borger et al. 2002). In fact, the latter is based on a systematic approach to collecting and reviewing data in order to identify opportunities to improve service performance with the end result of delivering better services to customers. More in detail, the continuous quality improvement process emphasizes an ongoing or continual process of improvement and evaluation/self-evaluation that involves the Plan, Do, Check and Act (PDCA) cycle, an ongoing cycle composed by the following steps (see Fig. 1):

- **Plan:** establish the objectives and processes necessary to deliver results in accordance with objectives;

- **Do:** implement the plan step, execute processes and deliver the service. Collect evaluation data for charting and analysis in the following “check” and “act” steps;

- **Check:** study the actual evaluation results measured and collected in “do” phase and compare against fixed objectives to ascertain any differences;
- **Act**: request corrective actions on significant differences between actual and planned results. Analyze the differences to determine their root causes.

Figure 1: Continuous quality improvement with PDCA cycle.

In particular, evaluation allows the comparison between fixed objectives vs achieved results, that represents the capability to set out objectives and to subsequently verify whether, and eventually at what level, they have been reached. Therefore, evaluation returns useful indications and elements to well support the decision whether to continue undertaken activities or to modify a part or even to redefine all or part of settled objectives.

The second reason arises from fundamental importance assumed for Universities by stakeholders satisfaction, which is considered a central factor to face competitive pressures from different sides. First of all, as pointed out by Helgesen and Nesset, (2007), stakeholders satisfaction is related to recruitment, retention and academic success. Moreover, Universities are currently characterized by the so called “rankings
“war” that is meaningful influenced by the achieved stakeholders satisfaction level (Letcher and Neves, 2010).

A quality University is that able to warranty certainties to all its stakeholders on its own reliability and capability to obtain over time suitable results, with respect to the fixed and promised objectives. In particular, considering the external stakeholders, which are composed by students, families and related social and working community, as well as Public Authorities and scientific community, the related quality concept arises from University services capability to fully satisfy their expectations and needs. On the contrary, considering the internal stakeholders, which are composed by professors, researchers and University staff, as well the MIUR (Ministry of Education, University and Scientific Research) and related public authorities, the quality concept is strictly related to the management capability to obtain excellent results for all the stakeholders. In this sense, the implementation of a proper quality management system can be considered a valid approach to improve management and control capabilities to deliver excellent services for the reference market. Such a capability can assume a strategic value for the market success of the service provider, especially in a very competitive contest such as the academic one.

Finally, the last reason concerns the Universities necessity to provide clear and correct informations about quality of delivered education services and their capability to reach the promised results. Such a possibility, that represents a valid support for an aware student choice, constitutes a very important aspect related to evaluation that highlight deep sense of competence, responsibility and commitment of Universities.
The remainder of the present chapter is organized as follows: in the next Section the Italian higher education sector is analyzed in the light of the recent normative measures addressed to education services quality aspects; the herein considered evaluation methodology is theoretically introduced in Section 3 and described in Section 4; in Section 5, the education services of the Management Engineering program at the University of Palermo are strategically analyzed with more details and the related practical considerations are given. Finally, conclusions with a suitable strategy for the overall quality improvement of the analyzed education services close the work.

2. Evaluation and self-evaluation in Italian higher education area

In recent years, the European Community has addressed considerable attention to educational institutions, since they play a critical role in national and global development (Wu et al. 2012). In particular, educational institutions support global development strategies by providing the highly qualified manpower and research necessary for further growth (Al-Turki and Duffuaa, 2003). Among the different levels of education, higher education is especially conducive to fostering high-tech talent, which is the key factor in increasing national economy and the main path to improving a nation’s competitiveness (Fairweather, 2000; Meek, 2000). Thus, higher education has a great impact on the development of a nation’s competitive advantage (Dill, 1997; Dill and Teixeira, 2000). Such a context explains the growing interest in academic institutions (Sanchez and Elena, 2006), and a great number of Universities
are going through important transformations in order to increase their performance level in education and research (Sorbonne Joint Declaration, 1998; Bologna Declaration, 1999; Prague Declaration, 2001, as cited in Sanchez and Elena, 2006). Furthermore, intellectual capital has become crucial in order to reinforce Universities’ roles in the new economy for two reasons: firstly, Universities’ main inputs and outputs are largely intangible, and only a small portion of them have a great effect on the Universities’ operation processes (Canibano and Sanchez, 2004); secondly, Universities are being forced to be more transparent and to disseminate more information to stakeholders such as students, public authorities that fund Universities, labor markets, etc. (European Commission, 2003; Wu el al, 2010).

In Italy, the idea of evaluating services performance in academic contest is relatively new. First attempts date back to the nineties; in particular, in 1991 the CRUI, i.e. the Conference of Italian University Rectors, instituted a committee of Rector’s delegates and considered approaches and tools in order to start an experimental activity related to performance evaluation of education services. Over the years, service performance evaluation became the subject of legislative bills. More in detail, with the Law n. 537/1993 is instituted, inside each University, the "Internal Evaluation Team" with the aim of verifying the proper public resources management, as well as the performance level of both education and scientific research activities. Moreover, such law also institutes the “Observatory for Evaluation of the Academic System”, to which the Internal Evaluation Team has to submit the annual performance assessment report. Subsequently, with the Law n. 59/1997 are introduced specific regulations to control, among other things, services performance of the academic system and, with the Law n. 370/1999 it is introduced as mandatory activity for Universities, the systematic
Recentely, the Legislative Decree n. 19/2012 introduces for Universities as mandatory the following activities:

- implementation of an initial and periodical accreditation system of higher education programs for both the Bachelor and Master Degrees;
- periodical assessment of quality and effectiveness-efficiency of education services;
- development of an effective internal and external communication system with the aim to involve stakeholders in the continuous quality improvement of education services;

In particular, ANVUR, the National Agency for the Evaluation of Universities and Research, in the implementing of the Legislative Decree n. 19/2012 started in March 2012, institutes the AVA integrated system (Self-evaluation, evaluation and accreditation), that is based on the following steps (see Fig. 2):

- initial an periodical accreditation system of education programs and Universities sites;
- University quality, efficiency and obtained results periodical evaluation;
- implementation of self-evaluation procedures related to quality and effectiveness of education services and research.
In the processing and development of AVA system, ANVUR considered three main principles which have to characterize, in a progressive way, University behaviors: autonomy, accountability and performance evaluation.

Very recently, with the implementation of the Ministerial Decree 47/2013, which recognized ANVUR requests related to AVA integrated system, Universities have progressively began the implementation of AVA system related activities starting from 2013. Such activities are also crucial for the implementation of the Quality Management System, in accordance to requirements specified in the ISO 9001:2008.

3. Conceptual approach of education services performance evaluation
In the literature, there is a variety of methods regarding performance measures of education services. Such methods can be basically classified as stated importance methods, in which students are asked to rate each service characteristic on a Likert-type importance scale, or derived importance methods, in which importance measures are statistically derived considering relationships among individual sub-criteria with overall satisfaction (Ory, 2000). Derived importance methods are widely considered in the recent decade since, although stated importance methods are more intuitive and simple to use, they require a significant increase in the length of the survey and can sometimes yield insufficient differentiation among mean importance ratings. Several recent applications of derived importance methods based on student satisfaction (SS) survey are described in: Sultan and Wong, (2013); Ardi et al (2012); Hosseini (2012). From stated preference surveys, the most recent applications are described in: Kuzmanovic M. et al (2013); Vaz and Mansori, (2013); Hur and Pak, (2007).

Recently, many Authors focus their attention on the heterogeneity of students’ perceptions and expectations of different aspects of teaching services (Kuzmanovic M. et al 2012). In particular, such heterogeneity, mainly related to certain aspects of education services, the different students attitudes have toward the use of education services, the social and economic characteristics of students and the different ways of viewing aspects of services, can represent a serious problem for many techniques that intend to measure service quality (Eboli and Mazzulla, 2011).

The level of SS represents an indirect measure of the reached level of education services performance, since it is related to proper service aspects whose performance levels, quantified by means of the so called “manifest variables”, are intended as “latent manifestations” of service performance (De Battisti et al., 2010). The
relationship between manifest variables and service performance can be formalized by means of specific service conceptual models (Ding, 2006).

In the literature, several service conceptual models have been introduced (Schvaneveldt et al. 1991; Cronin and Taylor, 1992; Teas, 1993; Franceschini and Rossetto, 1998; Ghobadian et al. 1994) and, to date, the Service Quality model (ServQual) (Parasuraman et al. 1985) is one of the most established conceptual models for determining customer satisfaction in services (Büyüközkan et. al. (2011); Lupo, 2013a). Several recent applications of the ServQual model in different service fields are described in: Liu et. al. (2009); Lin, (2010); Lupo (2013b).

The ServQual model in its original formulation consists of 22 statements measuring 5 critical to quality dimensions of service quality. The required data for the service quality assessment are quantitative in nature, which can be expressed in terms of exact numbers by linguistic-numerical evaluation scales, and 5 major service Gaps are considered by the Authors to evaluate service performance. More in detail, the algebraic discrepancies (Gaps) between customers’ perceptions and their expectations can be considered direct indicators of the customer satisfaction degree (Parasuraman et al., 1985; Parasuraman et al., 1988; Zeithaml et al. 1990). However, many criticalities are associated to the employment of the ServQual model. Some difficulties are related to the use of linguistic-evaluation scales: the well-documented tendency of respondents to select central linguistic categories to express her/his judgments, influence of the linguistic categories number in the evaluation process, the form and the type of the adopted linguistic variables and, finally, the transformation from cardinal to ordinal data. Other critical factors are related to ambiguity of
expectations evaluation (Babakus and Boller, 1992) and the difficulties arising from the use of differential psychometric score (Brown et al., 1993).

In the light of the previous considerations, in the present chapter the ServQual discrepancy paradigm is considered to evaluate the education services performance, i.e. SS. However, to estimate service expectations’ levels, the Analytic Hierarchy Process (AHP) method is herein considered (Saaty 1980) in order to overcome the previously described ServQual criticalities. AHP is a multi-criteria decision making (MCDM) method that helps the decision-maker facing a complex problem with multiple conflicting and subjective criteria. It is based on three principles that determine the procedure steps of the method: (Forman and Gass, 2001; Saaty, 2008):
the principle of hierarchical decomposition, which is considered to reduce the problem complexity; the principle of comparison judgments, used to perform the comparisons between all elements pairs belonging to the same level and finally, the principle of the synthesis, considered to aggregate partial results in order to obtain the global one.

AHP presents several advantages as: full differentiation among importance ratings, seeking consistency in judgments by means of the inconsistency ratio IR, easiness to use, etc. It also allows to structure complex problems and can be combined with operations research techniques to handle more difficult problems. Several papers have compiled AHP success stories in very different fields. Recently, Certa et al. (2013), summarize the various engineering fields in which AHP has been recently considered. On the contrary, in the field of service quality assessment, relatively few works adopting AHP have been proposed and most of them only recently. However, AHP in
its original formulation can be unreliable in handling ambiguity of the concepts associated with the use of the human knowledge. In particular, since the human knowledge to express judgments and preferences can be very often incomplete, inconsistent and even vague or imprecise, as consequence, individuals cannot be able to express, with suitable reliability level, their opinion by means of an exact numerical value related to a linguistic variable as required by AHP. Such a situation can introduce uncertainty in service performance analyses (Hsiu-Yuan Hu et al., 2010).

The choice of the technique to be used to minimize uncertainty effects is usually based on the type and nature of uncertainty (Ferdous et al. 2012). Therefore, the Fuzzy Set Theory (FST) (Zadeh, 1965) is herein considered to deal with uncertainty related to service performance analyses. In fact, uncertainty ascribed to service performance analyses is of epistemic type: it arises from three relevant uncertainty aspects, i.e. vagueness, imprecision and subjectivity in customers’ service quality judgments.

The innovative contributions that came from the use of the FST derive from its representation and processing capability of information affected by some imperfection typically due to the use of the natural language (Zimmermann, 1985). In particular, the FST allows the mathematical representation of uncertain knowledge and provide formalized tools for dealing with intrinsic imprecision of real life problems: it is particularly useful in the quantification of linguistic categories since it allows the representation for different “membership degrees” of a concept (Neigota 1985). This feature is well represented by the idea expressed by Zadeh (Zadeh, 1975; Zadeh, 1996) on the fuzzy set as tool "to compute with words", which highlights the need for
an effective interface between the numbers crisp world and linguistic categories to improve the understanding and utilization capability of real-life information. The FST has been applied in many fields of the management science, but it is still quietly used in the field of the service quality assessment (Chien et al. 2011; Hsiu et al. 2010). For further details and considerations on the considered conceptual approach related to education services performance evaluation, the reader may refer to Lupo (2013 c).

In the light of the previous considerations, it is herein described a methodology based on the ServQual model to evaluate performance of education services. In particular, a combined procedure based on integration between the AHP method and the FST is considered to effectively handle epistemic uncertainty in service performance analyses. Subsequently, the strategic analysis of education services related to the Management Engineering program at the University of Palermo (Italy) is performed by using the proposed method.

4. Fuzzy ServQual based methodology for reliable service performance measurements

In the next Section a brief overview about FST and its theoretical principles useful for the aim of the present work are given. Subsequently, the considered methodological approach for the measurement of students’ importance weight and perception levels of service criteria, sub-criteria and items is described.
4.1 Fuzzy set theory and linguistic-fuzzy evaluation scales

In FST, the concept of convexity of a set differs from that applied in the classical set theory: a fuzzy set is said convex if and only if the degree of membership \( \mu_A \) of an element \( x \) between two elements \( x_1 \) and \( x_3 \) is not less than the minimum value among the membership degrees of \( x_1 \) and \( x_3 \). Then, a fuzzy number \( \tilde{A} \) is a convex fuzzy set defined in \( \mathbb{R} \) and such that:

- \( \exists x_0 | \mu_A(x_0) = 1 \)
- the membership function \( \mu_A(x) \) is continue.

The fuzzy assessment approach takes into account that a linguistic variable can be seen as a variable whose values are words or structured combinations defined by semantic rules. In particular, a linguistic variable is characterized by five elements \( (X, T(X), U, G, M) \), where: \( X \) is the name of the variable; \( T(X) \) is the set of its linguistic categories, \( U \) is the universe of discourse, \( G \) is a syntactic rule that generates the terms in \( T(X) \) and \( M \) a semantic rule that associates to each linguistic category \( x \) of \( T(X) \) the related meaning \( M(x) \). Such semantic rule may be defined by a fuzzy number \( \tilde{M}(x) \) in \( U \). Thus, the meaning of \( \tilde{M}(x) \) of a linguistic category \( x \) is defined by a membership function \( \mu_M : U \rightarrow [0,1] \) that associates to each \( u \) of \( U \) its compatibility with \( x \) (Klir and Yuan, 1999).

A positive triangular fuzzy number (TFN), denoted as \( \tilde{A} = (x_L, x_M, x_U) \) (see Fig. 3), where \( x_L \leq x_M \leq x_U \), has the following triangular-type membership function:
\[
\mu_A(x) = \begin{cases} 
\frac{x-x_L}{x_M-x_L} & \text{for } x_L \leq x \leq x_M \\
\frac{x-x_U}{x_U-x_M} & \text{for } x_N \leq x \leq x_U \\
0 & \text{otherwise}
\end{cases}
\]  

Alternatively, by defining the interval of confidence level \( \alpha \) (\( \alpha \)-cut), a TFN can be characterized as:

\[
\forall \alpha \in [0,1] \\
\tilde{A}_\alpha = [a^\alpha_L, a^\alpha_U] = [(x_M - x_L) \alpha + x_L, -(x_U - x_M) \alpha + x_U]
\]

Figure 3: Positive triangular fuzzy number

Service performance analyses often articulate students’ knowledge/judgments in term of linguistic variables such as: very bad, poor, average, good, excellent, etc, and the use of TFNs as a way to compare fuzzy judgments has been proposed by van Laarhoven and Pedrycz (1983) and it is largely adopted in more recent works (Cheng 1996; Kwong and Bai, 2002). In particular, Ayyub and Klir (2006) provided a chart to define the lower and upper boundary for such linguistic variables based on experts’ assessments. Considering the most likely value as an average of these two boundaries,
TFNs can be used to represent such linguistic variables. The fuzzy boundaries of a TFN may also be defined by means of the Fuzzy Delphi method that is a typical multi-experts procedure for combining views and opinions (Kaufmann and Gupta, 1988).

Finally, the FST allows the extension of arithmetic operations from crisp numbers to fuzzy numbers. By considering the membership degree $\alpha$ ($\alpha$-cut) of positive fuzzy numbers, some main operations useful for the aim of the present work are given by the following expressions (Klir and Yuan, 1999):

$$\forall \alpha \in [0,1], \quad \forall a_L, a_U, b_L, b_U \in R^+, \quad A_\alpha = [a_L^\alpha, a_U^\alpha], \quad B_\alpha = [b_L^\alpha, b_U^\alpha]$$

$$A_\alpha \oplus B_\alpha = [a_L^\alpha + b_L^\alpha, a_U^\alpha + b_U^\alpha]$$

$$A_\alpha \odot B_\alpha = [a_L^\alpha - b_L^\alpha, a_U^\alpha - b_U^\alpha]$$

$$A_\alpha \otimes B_\alpha = [a_L^\alpha \times b_L^\alpha, a_U^\alpha \times b_U^\alpha]$$

$$A_\alpha \div B_\alpha = [a_L^\alpha / b_L^\alpha, a_U^\alpha / b_U^\alpha]$$

Finally, in the present work linguistic variables are used to represent students’ assessments and TFNs are considered for their evaluations. The methodological approach is described as follow.

### 4.2 Importance weights measurement of service aspects

As before said, in this chapter a fuzzy extensions of the AHP method (FAHP) is considered to effectively handle uncertainty related to the measurement of students’ judgments. In the literature, FAHP has been widely adopted whenever individuals cannot use an exact numerical value to express their opinion about a situation and a
linguistic variable is used to represent that specific numerical value. Often, a proper linguistic variable is created to explain the ambiguity and vagueness associated with the domain of the problem. Therefore, the concept of the linguistic expression can be quantified by fuzzy numbers using a proper membership function. Several recent applications of the FAHP approach in different fields are described in: Cebeci (2009); Celik et al. (2009); Chamodrakas et al (2010). However, the FAHP approach is still scarcely used in the field of service quality assessment (Lupo, 2013c; Büyüközkan 2011).

In order to adopt the FAHP method for the aim of the present work, the first stage concerns the description of the analyzed service quality structure, also mentioned as “value tree” or “value hierarchy” (Kirkwood, 1997). Such structure consists of several hierarchical levels: the first one includes the general objective or goal of the analysis, i.e. overall customer satisfaction. In the second level service quality criteria, i.e. service main characteristics that are deeply related with customer satisfaction are reported. Subsequently, in the third level service sub-criteria for each service criterion are identified. Finally, in the fourth level the service items for each service sub-criterion are considered. These satisfaction dimensions should assure a consistent family of criteria, with the following properties: monotonicity, exhaustiveness, and non-redundancy (Roy and Bouyssou. 1993; Keeney and Raiffa, 1976). Keeney et al, (1976), proposed also that the set of criteria and the formulated value hierarchy should be operational, decomposable, and minimal. Fig. 4 shows a general hierarchical service quality structure composed by w quality criteria, $C_1, C_2, \ldots, C_w$. 
Subsequently, the four step-procedure described below has to be considered:

- **Step 1: Assessment of pairwise comparisons coefficients:**

Linguistic terms are used to indicate the relative importance of each elements pair in the same hierarchal level (Fig. 4) and TFNs are considered to quantify concepts of linguistic expressions. In particular, considering the service sub-criteria of the generic service criterion $C_k$, the generic coefficient $\tilde{a}_{i,j}$ represents the relative importance
weigh, expressed in fuzzy form, of the service sub-criterion \( i^{th} \) vs the \( j^{th} \) one.

Moreover, not all the \( C_k^2 \) pairwise comparison coefficients have to be directly detected, but only \( C_k (C_k - 1)/2 \), given that it is valid the comparisons reciprocity property, that is:

\[
\begin{align*}
\tilde{a}_{i,j} &= \frac{1}{\tilde{a}_{j,i}} \\
\tilde{a}_{i,i} &= 1
\end{align*}
\]

\( \forall i, j = 1,2,...,C_k \)
\( \forall i \neq j \) \[4\]

- **Step 2: Constructing of fuzzy comparison matrices:**

Collected pairwise fuzzy comparison coefficients are used to construct the pairwise comparison matrices, which are squared, reciprocal and positive matrices.

For the aggregation of multiple customers’ judgments, a number of methods, e.g., max-min. arithmetic averaging, symmetric sum, t-norm, etc., are available. However, the geometric mean is the herein considered as aggregation operator since, as pointed out by Enea and Piazza, (2004), it allows the respect of the AHP constraint expressed by Eq. (4). In fact, if \( \tilde{p}_{i,j,k} \) is the fuzzy preference of the generic \( k^{th} \) customer and \( t \) the number of judgments to be aggregated, it is possible to write:

\[
\begin{align*}
\tilde{a}_{i,j} &= \left( \prod_{k=1}^{t} \tilde{p}_{i,j,k} \right)^{1/t} \\
\tilde{a}_{j,i} &= \left( \prod_{k=1}^{t} \frac{1}{\tilde{p}_{i,j,k}} \right)^{1/t}
\end{align*}
\]

\[5\]
and consequently:

\[
\begin{align*}
\tilde{a}_{i,j} &= 1/\tilde{a}_{j,i} \\
\tilde{a}_{i,i} &= 1
\end{align*}
\]  \hspace{1cm} \text{[6]}

- **Step 3: Computing the local importance weights:**

Considering the service sub-criteria of the service criterion \( C_k \), such step is related to computing of the maximum fuzzy eigenvalue \( \tilde{\lambda}_{\max} \) of the related fuzzy comparison matrix \( \tilde{A}_k \):

\[
\tilde{A}_k \cdot \tilde{w} = \tilde{\lambda}_{\max} \cdot \tilde{w}
\]  \hspace{1cm} \text{[7]}

in which \( \tilde{w} \) is a fuzzy vector (\( C_k \times 1 \)) composed by \( C_k \) fuzzy numbers representing the local importance weights of the considered service sub-criteria.

In particular, considering the relationships reported in Eq. (2), for the generic service sub-criterion \( i^{th} \), the Eq. (7) can be written as:

\[
\left[ \left( a_{L}^a \right)_i \cdot \left( w_{L}^a \right)_i \right] \oplus \left[ \left( a_{U}^a \right)_i \cdot \left( w_{U}^a \right)_i \right] \ldots \oplus \left[ \left( a_{L}^a \right)_n \cdot \left( w_{L}^a \right)_n \right] \oplus \left[ \left( a_{U}^a \right)_n \cdot \left( w_{U}^a \right)_n \right] = \left[ \tilde{\alpha}_{i} \cdot \left( w_{L}^a \right)_i \right] \oplus \left[ \tilde{\alpha}_{U} \cdot \left( w_{U}^a \right)_i \right]
\]  \hspace{1cm} \text{[8]}

in which:
The $\alpha$-cut is known to include the customer’s confidence over her/his preferences. In the case herein considered it incorporates the students’ confidence and uncertainty over their judgments. Therefore, by considering the index of optimism $\mu$ (Chang 1996; Lee 1999), the crisp pairwise comparison coefficient of the importance between the service sub-criteria $i$ and $j$ at the confidence level $\alpha$ can be written as:

$$\tilde{a}_{ij}^{\alpha} = \mu \cdot a_{i,j}^{\alpha} + (1 - \mu) \cdot a_{i,j}^{\beta}$$

where $a_{i,j}^{\alpha}$ and $a_{i,j}^{\beta}$ are the crisp pairwise comparison coefficients at confidence levels $\alpha$ and $\beta$, respectively. The expression is valid for $\forall \alpha \in [0,1]$.

When $\alpha$ is fixed, after setting the index of optimism $\mu$ the following matrix (11) can be obtained and considered to estimate the local importance weights of the considered service sub-criteria.

$$\tilde{A}_k = \begin{bmatrix}
1 & \tilde{a}_{1,2}^{\alpha} & \cdots & \tilde{a}_{1,n}^{\alpha} \\
\tilde{a}_{2,1}^{\alpha} & 1 & \cdots & \tilde{a}_{2,n}^{\alpha} \\
\vdots & \ddots & \ddots & \vdots \\
\tilde{a}_{n,1}^{\alpha} & \cdots & \cdots & 1
\end{bmatrix}$$

The Eqs. (7), (9) and (11) correspond to the fuzzification of the Lambda-Max method, initially introduced by Saaty (1980) in crisp term with the AHP method, which has been introduced by Csutora and Buckley, (2001).
• **Step 4: Computing the global importance weights of service items:**

The last step is related to determination of the global importance weights of the service items. The latter can be obtained by multiplying the local importance weight of each service item by the local importance weights of the related service criterion and sub-criterion (Saaty 1980).

4.3 Perceptions measurement of service aspects

The crisp perception related to the service sub-criterion \( i \) of the generic service criterion \( k \), at the confidence level \( \alpha \) (\( \alpha \)-cut), \( \bar{p}_{\alpha_{k,i}} \), can be obtained with refer to the judgments satisfaction degree estimated by the index of optimism \( \mu \):

\[
\bar{p}_{\alpha_{k,i}} = \mu(p^\alpha_{U_{k,i}}) + (1-\mu)\cdot(p^\alpha_{L_{k,i}})
\]

\[\forall \alpha \in [0,1]\]

in which in Eq. (12) \( p^\alpha_{U_{k,i}} \) and \( p^\alpha_{L_{k,i}} \) are the upper and lower bounds of fuzzy aggregated judgments at the confidence level \( \alpha \) (\( \alpha \)-cut), considering as aggregator operator the arithmetic mean. While \( \alpha \) is fixed, after setting the index of optimism \( \mu \), Eq. (12) gives the related crisp value of students’ perception for the considered sub-criterion.

5. Performance evaluation of Palermo Management Engineering program education services

22
The Management Engineering program at the University of Palermo has both Bachelor and Master Degrees and it is characterized by a total students’ basin of about 500 units. The quality cycle tools (see Fig 5) are typically considered to control and to improve education services and related research activities performance.

![Quality Cycle Tools](image_url)

**Figure 5: Quality cycle tools**

In addition, currently further self-evaluation and quality assurance procedures requested by AVA integrated system have been implemented. Such further recent implementations represent deep sense of competence, responsibility and commitment of the higher education program. In particular, the strategic education services analysis hereafter reported is a meaningful step toward the direction to be pursued, in which stakeholders, and in particular students, are the main drivers for education services continuous quality improvement.
5.1 Value tree of education services

As before said, to adopt the developed evaluation methodology the value tree of the under analysis service has to be described. Considering education services, there is not a unique and widely accepted service value hierarchy: such a situation highlights the deep complexity of teaching, learning and support services in academic context (Kuzmanovic et al., 2013). However, starting from the hierarchical structure of the education services dimensions, sub dimensions and items developed by Lupo and Passannanti, (2008), the value tree of criteria, sub-criteria and items for the considered education services was pointed out by using the Critical Cases Approach (Cronin and Taylor, 1992), on the basis of preliminary focus groups with both service experts (decision makers group) and a limited number of stakeholders. The structure reported in Fig.6 describes the obtained value tree of Management Engineering program education services.
Figure 6: Services value tree of the Palermo Management Engineering program.
5.2 Questionnaire structure and fuzzy-linguistic evaluation scales

The value tree reported in Fig. 6 is considered to develop the questionnaire for the SS survey. In particular, such questionnaire is composed by two parts; in the first one, students are asked to indicate the relative importance of all the pairwise comparisons of education services criteria, sub-criteria and items. On the contrary, in the second one, students are asked to assess their perception levels related to education services items. In both the questionnaire parts, students point out levels of their judgments by using suitable fuzzy-linguistic evaluation scales. Table 1 shows an extract of the first and the second part of the developed questionnaire related to the service sub-criterion *Classrooms* (*SC₂₁*) belonging to the service criterion *Infrastructures* (*C₂*).
Table 1: Extract of the developed questionnaire for the *Classrooms* service items

**First part:**

| Compared with | How important is: | | |
|---------------|-------------------|---|---|---|
| | Availability of individual places | Classrooms lighting | Classrooms air ventilation and conditioning |
| | A B = a b | A B = a b | A B = a b |
| Classrooms acoustic | C D = c d | C D = c d | C D = c d |
| Classrooms air ventilation and conditioning | A B = a b | A B = a b | A B = a b |
| C D = c d | C D = c d | C D = c d |
| Classrooms lighting | A B = a b | A B = a b | A B = a b |
| C D = c d | C D = c d | C D = c d |

Please, mark in the questionnaire form the letters related to your judgements:

- **D:** Extremely more important
- **C:** Very strongly more important
- **B:** Strongly more important
- **A:** Moderately more important
- **=:** Equally important
- **a:** Moderately less important
- **b:** Strongly less important
- **c:** Very strongly less important
- **d:** Extremely less important

**Second part:**

<table>
<thead>
<tr>
<th>Classrooms</th>
<th>Mark the performance level of the following service attributes:</th>
<th>Very bad</th>
<th>Poor</th>
<th>Average</th>
<th>Good</th>
<th>Excellent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Availability of individual places</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td></td>
</tr>
<tr>
<td>Classrooms lighting</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td></td>
</tr>
<tr>
<td>Classrooms air ventilation and conditioning</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td></td>
</tr>
<tr>
<td>Classrooms acoustic</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td></td>
</tr>
</tbody>
</table>
Finally, for both the index of optimism $\mu$ and the confidence level $\alpha$ ($\alpha$-cut) it is considered a value equal to 0.5 and the linguistic-fuzzy evaluation scales reported in Table 2 have been considered.

Table 2: Linguistic-fuzzy evaluation scales.

<table>
<thead>
<tr>
<th>Perception evaluation scale</th>
<th>Importance evaluation scale</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Linguistic category</strong></td>
<td><strong>TFN</strong></td>
</tr>
<tr>
<td>Very bad</td>
<td>(0, 1, 3)</td>
</tr>
<tr>
<td>Poor</td>
<td>(2, 3, 5)</td>
</tr>
<tr>
<td>Average</td>
<td>(3, 5, 7)</td>
</tr>
<tr>
<td>Good</td>
<td>(5, 7, 9)</td>
</tr>
<tr>
<td>Excellent</td>
<td>(7, 9, 9)</td>
</tr>
</tbody>
</table>

5.3 Perception and importance weight levels of education services

The SS survey has been conducted for three months, between February and April 2013, and about 200 students equally distributed between Bachelor and Master Degrees have been face to face interviewed. Just for an example, considering the service items of the sub-criterion Classrooms, Table 3 shows both the obtained aggregated fuzzy comparison matrix ($a$) and the related crisp comparison one ($b$).
Table 3: Fuzzy (a) and crisp (b) aggregated comparison matrices for the Classrooms service items ($\mu = 0.5; \alpha = 0.5$).

<table>
<thead>
<tr>
<th>$S_{1,1,1}$</th>
<th>$S_{1,1,2}$</th>
<th>$S_{2,1,2}$</th>
<th>$S_{2,1,4}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(0.215; 0.299; 0.475)</td>
<td>(0.181; 0.249; 0.424)</td>
<td>(0.540; 0.763; 1.599)</td>
</tr>
<tr>
<td>$S_{2,1,2}$</td>
<td>(2.189; 3.342; 4.832)</td>
<td>1</td>
<td>(1.645; 2.382; 3.518)</td>
</tr>
<tr>
<td>$S_{2,1,3}$</td>
<td>(2.552; 4.096; 5.822)</td>
<td>(0.640; 0.963; 1.462)</td>
<td>1</td>
</tr>
<tr>
<td>$S_{2,1,4}$</td>
<td>(0.939; 1.309; 2.779)</td>
<td>(0.308; 0.419; 0.659)</td>
<td>(0.309; 0.481; 0.798)</td>
</tr>
</tbody>
</table>

(a) (b)

The global importance weighs of education services items have been obtained by solving the Eq. (7) for each considered fuzzy comparison matrix and, subsequently, by applying the step 4 of the developed four step-procedure. On the contrary, perception levels of education services items have been obtained considering Eq. (12).

For example, with refer to the service items of the sub-criterion Classrooms, Tables 4 reports both the obtained fuzzy aggregated measurements of students’ perceptions (a) and the related crisp ones (b).

Table 4: Fuzzy (a) and crisp (b) aggregated measurements of students’ perceptions for the Classrooms service items ($\mu = 0.5; \alpha = 0.5$).
Finally, SS levels have been obtained by calculating for each education services item the Gap between the related levels of importance weight and perception (Parasuraman et al., 1985). Therefore, student’s dissatisfaction is collected for the service items in which a negative Gap value is obtained. Table 5 summarizes the obtained levels of students’ importance weight, perception and satisfaction.
Table 5: Students’ expectation, perception and satisfaction levels.

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Importance Weight</th>
<th>Local Importance Weight</th>
<th>Service Item</th>
<th>Local Importance Weight</th>
<th>Global Importance Weight</th>
<th>Perception</th>
<th>SS level</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>0.360</td>
<td></td>
<td>SC1,1</td>
<td>0.502</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI1,1,1</td>
<td>0.171</td>
<td>0.031</td>
<td>0.033</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI1,1,2</td>
<td>0.215</td>
<td>0.039</td>
<td>0.029</td>
<td>-0.010</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI1,1,3</td>
<td>0.360</td>
<td>0.065</td>
<td>0.022</td>
<td>-0.043</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI1,1,4</td>
<td>0.130</td>
<td>0.023</td>
<td>0.035</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI1,1,5</td>
<td>0.177</td>
<td>0.032</td>
<td>0.025</td>
<td>-0.007</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI1,2,1</td>
<td>0.191</td>
<td>0.034</td>
<td>0.035</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI1,2,2</td>
<td>0.088</td>
<td>0.016</td>
<td>0.029</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI1,2,3</td>
<td>0.539</td>
<td>0.097</td>
<td>0.039</td>
<td>-0.058</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI1,2,4</td>
<td>0.182</td>
<td>0.033</td>
<td>0.033</td>
<td>0.000</td>
</tr>
<tr>
<td>C2</td>
<td>0.107</td>
<td></td>
<td>SC2,1</td>
<td>0.867</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI2,1,1</td>
<td>0.444</td>
<td>0.041</td>
<td>0.029</td>
<td>-0.012</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI2,1,2</td>
<td>0.125</td>
<td>0.012</td>
<td>0.027</td>
<td>0.015</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI2,1,3</td>
<td>0.126</td>
<td>0.012</td>
<td>0.019</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI2,1,4</td>
<td>0.305</td>
<td>0.028</td>
<td>0.026</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI2,2,1</td>
<td>0.667</td>
<td>0.009</td>
<td>0.040</td>
<td>0.031</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI2,2,2</td>
<td>0.201</td>
<td>0.003</td>
<td>0.045</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI2,2,3</td>
<td>0.132</td>
<td>0.002</td>
<td>0.034</td>
<td>0.032</td>
</tr>
<tr>
<td>C3</td>
<td>0.334</td>
<td></td>
<td>SC3,1</td>
<td>0.057</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI3,1,1</td>
<td>0.268</td>
<td>0.005</td>
<td>0.014</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI3,1,2</td>
<td>0.243</td>
<td>0.005</td>
<td>0.015</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI3,1,3</td>
<td>0.135</td>
<td>0.003</td>
<td>0.014</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI3,1,4</td>
<td>0.353</td>
<td>0.007</td>
<td>0.015</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI3,2,1</td>
<td>0.299</td>
<td>0.059</td>
<td>0.018</td>
<td>-0.041</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI3,2,2</td>
<td>0.351</td>
<td>0.069</td>
<td>0.021</td>
<td>-0.048</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI3,2,3</td>
<td>0.351</td>
<td>0.069</td>
<td>0.014</td>
<td>-0.055</td>
</tr>
<tr>
<td>C4</td>
<td>0.199</td>
<td></td>
<td>SC4,1</td>
<td>0.539</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI4,1,1</td>
<td>0.233</td>
<td>0.025</td>
<td>0.038</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI4,1,2</td>
<td>0.218</td>
<td>0.023</td>
<td>0.041</td>
<td>0.018</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI4,1,3</td>
<td>0.364</td>
<td>0.039</td>
<td>0.039</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI4,1,4</td>
<td>0.184</td>
<td>0.020</td>
<td>0.038</td>
<td>0.018</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI4,2,1</td>
<td>0.137</td>
<td>0.013</td>
<td>0.036</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI4,2,2</td>
<td>0.427</td>
<td>0.039</td>
<td>0.035</td>
<td>-0.004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SI4,2,3</td>
<td>0.435</td>
<td>0.040</td>
<td>0.053</td>
<td>0.013</td>
</tr>
</tbody>
</table>

Moreover, the sensitivity analysis performed by varying both the index of optimism $\mu$ and the confidence level $\alpha$ ($\alpha$-cut) (see the Appendix 1) shows that, for most of the considered service items, there are not ranking reversal situations. More in detail,
global importance weights are characterized by almost constant values with respect to the confidence level $\alpha$. In addition, considering also transition situations, from pessimistic to optimistic state, the global importance weights values are characterized by not significant variations.

Finally, Fig. 7 summarizes obtained education services performance results. For each graph, “+” and “X” respectively denote levels of the importance weight and perception, for the considered education services items.

For further details and considerations on the developed strategic education services analysis, the reader may refer to Lupo (2013 c).
6. Conclusions

University system reforms occurred at worldwide level in the last two decades are the main manifestation of the deep transformation process of academic contests, starting from the education services design until to all Universities tasks. In particular, it states needs for a suitable quality management of these institutions, taking into account their specific tasks and necessities to promote and stabilize effective relationships between educational aspects with professional needs. One of the fundamental steps of this reorganization process is to promote evaluation and self-evaluation approaches and procedures of achieved performance, effectiveness and efficiency levels. For such reason, in this chapter a methodology able to perform reliable education services performance evaluation has been described. In particular, such methodology is
centred on the ServQual discrepancy paradigm and uses in combined manner AHP and the Fuzzy Sets Theory. The application of such method has been shown in a strategic education services performance analysis related to the Management Engineering program at the University of Palermo (Italy). From such analysis, the student satisfaction has been evaluated for all the main education services items as Gap measures between students’ perceptions and their expectations and, consequently, a suitable “Gaps oriented” strategy for the overall education services quality improvement can be identified. More in detail, in the light of the obtained results, the strategic Gaps oriented implications for services quality improvement should primarily take into account the service criteria Academic staff ($C_1$) and Equipment ($C_3$), and, in particular, the service items Frequency of exams sessions ($SI_{1,2,3}$), Suitability of student support equipment ($SI_{3,2,3}$), Suitability of teaching aids ($SI_{3,2,2}$) and Design of course structure based on job requirements ($SI_{1,1,3}$).
7. References


Appendix 1

By setting the index of optimism $\mu$ as 0.05, 0.5 and 0.95 respectively, thus considering pessimistic, moderate and optimistic situations, the graphs reported below by varying the confidence level $\alpha$ ($\alpha$-cut) from 0.1 (maximum uncertainty) to 1 (maximum certainty) have been obtained.
Table 4.1: Importance weights for different levels of α (α-cut) for different categories

<table>
<thead>
<tr>
<th>Category</th>
<th>Level</th>
<th>Importance Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic Staff</td>
<td>SI1,1</td>
<td>0.05</td>
</tr>
<tr>
<td>Infrastructures</td>
<td>SI2,1</td>
<td>0.06</td>
</tr>
<tr>
<td>Equipment</td>
<td>SI3,1</td>
<td>0.04</td>
</tr>
<tr>
<td>Support Services</td>
<td>SI4,1</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Figure A1: Sensitivity analysis considering $\mu = 0.05$
a): Academic staff $C_1$

b): Infrastructures $C_2$

c): Equipment $C_3$

d): Support services $C_4$

Figure A2: Sensitivity analysis considering $\mu = 0.5$
Figure A3: Sensitivity analysis considering \(\mu = 0.95\)