TESF Methodology for Statistics Education Improvement

Stefano Barone
Chalmers University of Technology
University of Palermo

Eva Lo Franco
University of Palermo

Journal of Statistics Education Volume 18, Number 3 (2010),
www.amstat.org/publications/jse/v18n3/barone.pdf

Key Words: Course quality evaluation; Teaching experiments; Student satisfaction index; Measuring improvement in education; SERVQUAL.

Abstract

The need for universities to achieve excellence in the services they provide has been the subject of research for several decades. The idea of involving students and recognizing the importance of their opinions has led to the creation of various models and tools. This paper focuses on teaching, a central service from which improvement actions of an academic institution should always begin. The article reviews and updates the previously developed Teaching Experiments and Student Feedback methodology. The methodology, which is primarily addressed to statistics teachers, allows practical aspects to be organized and decisions to be made based on data that has been collected from students and scientifically analyzed.

The steps for building a student satisfaction index are also described. This index, in its most complete version, takes into account possible correlations between importance of the evaluated aspect and scores, both of which are provided by the students. The paper presents an application of the methodology to a statistics course taught by one of the authors.

1. Introduction

At most universities, education (defined here as "knowledge/competence transfer") is primarily pursued through the design and implementation of degree programs and the courses they include. Courses form the basic service of an academic institution (Fram & Camp, 1995), which makes them a good starting point to achieve an overall program improvement. Such an improvement

should be pursued on two fronts: content and modes of delivery (Snee, 1990). The Teaching Experiments and Student Feedback - TESF (Barone & Lo Franco, 2008, 2009) is a methodology designed to manage a course from a perspective of continuous improvement. In its current state of development, proper application of TESF offers a way to improve some aspects of teaching, such as the suitability of the recommended study material or the teacher's ability to interact with students.

A university is responsible for the offered courses to different stakeholders, who may not always have the same interests. The most relevant of these stakeholders include the state (the guarantor of the development of culture and research), the labor market (demanding high level skills), students (the future productive resources) and their parents (Owlia & Aspinwall, 1996). Students, as the direct and largest users of the courses that a university provides and manages, have a prominent role (Quinn, Lemay, Larsen & Johnson, 2009). In fact, most of the theoretical models and feedback tools developed with reference to the academic environment are based on the importance of student perceptions (Alves & Raposo, 2007; Rowley, 2003).

The present work focuses on a single course and the potential to improve it through the direct actions of its teacher as a result of student feedback (Panasuk & LeBaron, 1999). A teacher can improve the quality of his or her course by monitoring several aspects. These include not only student satisfaction levels but also new knowledge and competences acquired by students. For a discussion on how to implement continuous quality improvements in statistics courses and programs, see Hogg, Newton, Cobb & Bryce (1999).

In the earliest version of the TESF methodology, Barone & Lo Franco (2008, 2009) showed how a teacher can improve some aspects of his teaching by using Design of Experiments (DOE). The experiment results were collected through an *ad hoc* feedback tool inspired by the SERVQUAL model on SERvice QUALity (Parasuraman, Zeithaml & Berry, 1988) and submitted to a sample of students attending the course.

Since then further study was conducted to expand the TESF methodology by framing it in a longer time period, and involving not only "experimental" activities. For example, it has been defined which activities a teacher should carry out before and after the phase when he performs teaching experiments. These activities are aimed either to set a benchmark for the improvement or to monitor the progress of improvement over time.

The procedures to be followed in absence of experiments are some of the novelties presented in this article as well as a student satisfaction index purposely formulated to deal with the type of data arising in this case. For the readability of the article such formulation is placed in Appendix A.

This paper comprises four main sections. Section 2 discusses the role and the feedback of the student in respect to a generic university course. Section 3 reviews the TESF methodology and its feedback tool. Section 4 presents an application in a statistics course taught by one of the authors. Concluding remarks follow in Section 5.

2. Roles and limits of students and their feedback

There is an ongoing debate regarding the role that should be attributed to students in relation to

the services provided by a university, especially regarding a course (see, e.g., Maguad, 2007; Redding, 2005). **Table 1** reports a summary of the findings of an extensive bibliographic search carried out by the authors, concerning the role of students as characterized by researchers in the academic environment.

Table 1. Some opinions on students' role at university.

AUTHORS	STUDENT ROLE
Matthews (1993)	Customer
Helms & Key (1994)	Customer and employee
Babbar (1995)	Customer
Fram & Camp (1995)	Customer
Bailey & Bennett (1996)	Product
Owlia & Aspinwall (1996)	Customer
Sirvanci (1996)	Product in progress, internal customer for campus facilities and course material, laborer
Ensby & Mahmoodi (1997)	Product in progress
Felder & Brent (1999)	Acquire knowledge, skills, and values the instructor has set to impart
Kanji & Tambi (1999)	All roles of buyer, user, and partners of education
Karapetrovic, Rajamani & Willborn (1999)	Customer
Wallace (1999)	Primary customer
Canic & McCarthy (2000)	Primary customer
Chinn & Hilgers (2000)	User
Bier & Cornesky (2001)	Product in progress and users of the curriculum
Emery, Kramer & Tian (2001)	Unfinished product/ Product in progress
Clewes (2003)	Consumer of Higher Education
Harvey (2003)	Receive services
Politis & Sisko (2004)	Customer
Polese & Monetta (2006)	Primary client
Maguad (2007)	Internal customer
Quinn et al. (2009)	Customer

According to the ISO 9000 (2005), a customer is any recipient of products or services from a supplier organization. Customers can be people or organizations and can be either external or internal to the supplier organization.

The authors espouse the line of thought that students are the most direct customers of a university course. In addition to students, other subjects (e.g. administrative and teaching staff, labor market, students' parents) can be recognized as customers of a university course, since they receive a service from the university even though they are not direct users (Maguad, 2007).

This is a starting point for an attempt to deepen the often debated, but as yet unresolved, issue of listening to and using the so-called "voice of the student". The positions on this subject in the literature are controversial, ranging from those who do not recognize student feedback as being at all useful (Richardson, 2005), to those who believe that student feedback is an essential element for the improvement of a university (Kember, Leung & Kwan, 2002; Leckey & Neill, 2001).

According to the authors, expectations, perceptions and, in general, opinions of students can be very helpful to identify what, where and when to improve in a university. Student opinions collected through an *ad hoc* tool like a questionnaire can also help a teacher who wants to improve his course. Usually a student feedback tool aims either to verify learning outcomes (see e.g. Brady & Allen, 2002) or to measure student opinions (see e.g. Tudor, 2006; Toland and De Ayala, 2005; Marozzi, 2009).

In the statistics education community, the Survey of Attitudes Toward Statistics (SATS©) is widespread. It aims to measure student attitudes in statistics, and includes a series of statements in which the student respondent is the subject. The student marks a response (the responses range from 1-"strongly disagree", through 4-"neither disagree nor agree", to 7-"strongly agree") according to his degree of agreement with that statement. Conversely, the TESF feedback tool aims to measure student quality perceptions on some aspects of teaching, and includes a series of items in which the teacher or any other course factor (e.g. classroom, study material) is the subject. The student-respondent marks a response according to his degree of satisfaction with that item (see Appendix C). TESF is not designed to be exclusively adopted by teachers of statistics, but the authors believe that they are the natural recipients of this methodology because it is applicable to improvement in statistics instruction.

Based on the contents to be included in a feedback tool for student satisfaction with a particular course, this paper argues that students should not evaluate certain aspects, such as the topics to be included in the syllabus and how they are expanded upon. In fact, these variables depend upon one main factor that is of direct concern for the state rather than for students: the need to transfer the most appropriate knowledge and skills in relation to the scenarios of the labor market. Most students do not have sufficient knowledge to express an authoritative opinion on these issues.

On the other hand, students may be given the opportunity to express opinions on a teacher's ability to transfer knowledge (such as the perceived level of clarity), as well as on the external conditions in which the lessons take place (such as air conditioning and classroom comfort) because these aspects directly affect the learning process (Masjuan & Troiano, 2009). In these cases student feedback can be a valuable source of information for improvement measures (Braskamp, Wise & Hengstler, 1979).

3. TESF methodology, feedback tool and student satisfaction data analysis

TESF is a methodology for the continuous improvement of a course. Accordingly, a teacher should manage a course following the four phases of the Deming cycle: Plan, Do, Study, Act (PDSA). Figure 1 identifies the main activities of the TESF methodology, which can be seen as an iteration of PDSA phases.

A teacher who decides to start implementing the TESF methodology should measure the degree of satisfaction of the students at the first course edition. From that point the teacher may choose to alter aspects of the course through one or more "teaching experiments". The TESF methodology can be used to assess the impact of these experiments on student opinions. Alternatively, the teacher may simply wish to gauge how student opinions change over time without investigating any specific alterations in course content or delivery made in a teaching experiment.

The structure of the feedback tool and the procedures for its application vary in terms of the presence or absence of teaching experiments (Figure 1).

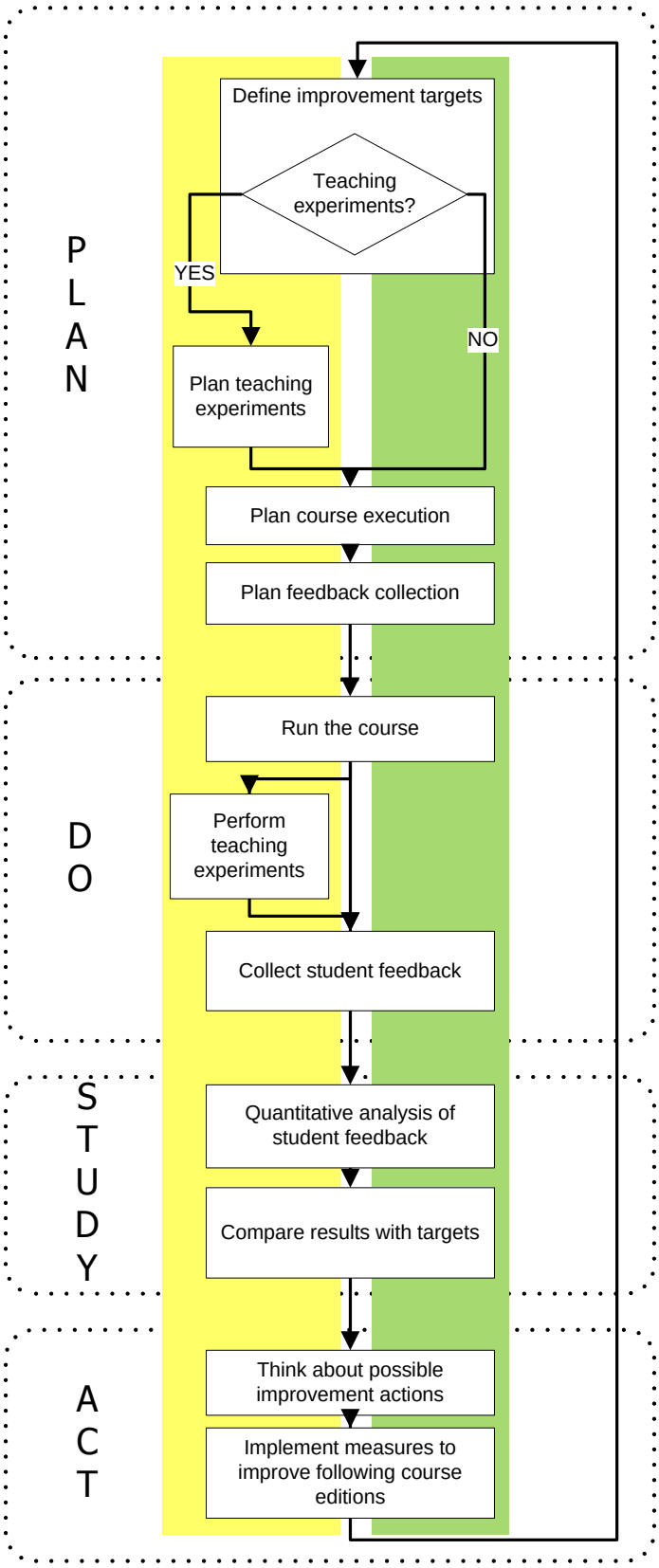

Figure 1. Flow chart of the TESF methodology.

3.1 TESF in absence of teaching experiments

If no experiments are carried out, the Plan phase is characterized by setting a target for the student satisfaction index (Appendix A). This index is a summary statistic that can be calculated for each item of the Evaluation Form (Appendix C). It can be calculated either by taking into account only the assigned scores (student satisfaction index, Γ), or by considering also the assigned weights (weight-adjusted satisfaction index, Γ') or, in the most complete version, by considering also the correlation between the assigned weights and scores (weight-score-adjusted satisfaction index, Γ''). In Appendix A these three versions of the student satisfaction index are formulated for the interested reader. Moreover, in the Plan phase the student feedback collection and the most suitable data analysis must be planned.

After the Plan phase, the course is carried out and the teacher collects the student feedback (Do phase). This is followed by a quantitative analysis of the collected data by using the student satisfaction index (Study phase). In the Act phase, the teacher should reflect upon the results and take decisions concerning subsequent courses.

The feedback tool adopted here is based on the quality dimensions of the well-known SERVQUAL model (Parasuraman *et al*., 1988). According to this model, the perceived quality of a generic service has five dimensions: responsiveness, tangibles, assurance, empathy and reliability. For a definition of these dimensions adapted to the service of a university course see Appendix D.

The feedback tool aims to measure the degree of satisfaction perceived by students attending a course, based on the instructional delivery modalities of that course. The tool is made up of four documents. The Weighting Grid (Appendix B) and the Evaluation Form (Appendix C) are to be filled in by students. The Course Quality Dimensions (Appendix D) and the Evaluation Matrix (Appendix E) help students fill in the Weighting Grid and the Evaluation Form. The Weighting Grid allows students to assign an importance weighting to the Course Quality Dimensions. It allows student expectations to be recorded and it should be submitted at the beginning of the course. The Evaluation Form, a questionnaire consisting of 21 items that interpret the quality dimensions, is designed to gather the perceptions of students and should be submitted at the end of the course.

The purpose and structure of the feedback tool, as well as the meaning of the five Course Quality Dimensions, are explained to students at the beginning of the course. When students provide their feedback they will refer to the Course Quality Dimensions definitions for a shared interpretation of the quality dimensions and will check the Evaluation Matrix to see how the Likert scale scores correspond with their related meanings.

3.2 TESF in presence of teaching experiments

Conversely, when it is decided to perform teaching experiments, the design of experiments methodology is called for. The responses to be analyzed are still in terms of student satisfaction and collected through the feedback tool. The feedback tool and the procedures adopted in the presence of experiments are slightly different from the case where no experiments are carried out. The teaching experiment factors are decided by the teacher as the variables potentially

affecting the course quality dimensions, by focusing attention on what can be perceived by a student. The selection of variables to adopt as control factors is done on the basis of the ability of the teacher to act on them. The number and levels of control factors must be chosen on the basis of the experimental plan to be adopted (Plan phase).

In the Do phase the teacher carries out the experiments and collects the student feedback throughout the course. A quantitative analysis of the collected data is made in the study phase followed by improvement actions (Act phase). An application of TESF in the presence of experiments is illustrated in detail in Barone & Lo Franco (2009).

4. Application of the methodology, an illustrative case

The case refers to the statistics course included in the degree program in Environmental Engineering at the University of Palermo. The adoption of the TESF methodology started at the first course edition in the academic year (A.Y.) 2004/2005. At that time the methodology was still at an early stage including only the possibility to make teaching experiments. Four control factors were chosen: modality of practices/laboratory work; board type; case studies; teacher-students interaction. All details of these experiments are reported in Barone & Lo Franco, 2009. The same experimental plan was replicated in the following course edition (A.Y. 2005/2006), due to the presence of a latent factor (availability of handouts) which was believed to affect the responses of the previous year experiments. The main indications coming from the analysis of these two years experimentation were the following: a) the board type (e.g. traditional blackboard/overhead projector) does not affect student satisfaction, provided that all handouts are given to the students early in advance; b) students appeared not to like the interaction solicited by the teacher.

In the Plan phase of the following course edition (A.Y. 2006/2007) it was decided to collect student feedback and calculate and analyze the student satisfaction index. That allowed checking the benefit of the previously run experiments and having a new baseline to define a student satisfaction target for subsequent course editions. The feedback tool was submitted to 69 students who were attending the course (Do phase). Two cases were excluded from the analysis because they were highly incomplete (students 20 and 21, see Table 2). Therefore, the analysis concerns the responses to 67 evaluation forms and weighting grids (Study phase).

Figure 2 shows the radar plot of the weight-score-adjusted satisfaction index $\Gamma"$ vs. each item of the Evaluation Form. The index value is the radial coordinate of the intersection points of the shaded area borderline and the lines corresponding to the items. It can be read with the help of the approximately circular lines equivalent to $\Gamma" = -0.72$ at the center of the plot and $\Gamma" = 0.88$ at its border.

Item II.1 “Classroom comfort” got the lowest index value ($\Gamma" = 0.15$). Conversely, Items III.1 “Teacher’s mastery of the course topics” ($\Gamma" = 0.48$), III.4 “Students’ ability to ask questions and make comments during class” ($\Gamma" = 0.52$), and III.5 “Teacher’s warmth and availability” ($\Gamma" = 0.49$) obtained the highest scores, indicating that “Assurance” was the most satisfactory dimension.

Table 2. Scores assigned to the Evaluation Form items by 69 students.

	student	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
item	I.1	0.75	0.50	0.75	0.75	1.00	0.50	0.25	1.00	0.75	0.75	1.00	0.75	0.50	0.75	0.50	0.75	0.25	0.75	1.00	0.75	0.50	0.75	1.00
	I.2	0.75	0.50	0.75	0.50	1.00	0.25	0.75	0.50	0.50	0.75	1.00	1.00	0.25	0.75	0.50	0.75	0.00	0.25	0.75	0.75	0.25	0.25	0.75
	I.3	0.25	0.75	1.00	0.75	0.25	0.75	0.50	0.75	0.75	0.50	1.00	0.75	0.25	0.50	0.50	0.50	0.50	0.50	1.00	0.25		0.50	1.00
	II.1	0.50	0.25	0.25	0.50	0.25	0.00	0.00	0.25	0.50	0.25	0.25	0.25	0.00	0.25	0.75	0.25	0.00	0.00	0.00	0.50	0.50	0.25	0.75
	II.2	0.75	0.50	0.50	0.25	0.50	0.50	0.00	0.50	0.75	0.50	0.75	0.25	0.25	0.50	0.50	0.75	0.25	0.25	0.75	0.25	0.50	0.50	0.75
	II.3	0.75	0.50	0.75	0.50	0.75	0.50	0.50	1.00	0.75	0.75	1.00	0.25	0.50	0.75	0.75	0.50	0.25	0.75	1.00	0.75		0.25	0.75
	III.1	1.00	0.75	0.75	0.75	0.75	0.75	0.75	1.00	0.75	0.75	1.00	1.00	0.75	0.75	0.75	1.00	0.75	0.50	1.00	0.75	0.50	0.50	1.00
	III.2	0.50	1.00	0.50	0.75	0.75	0.50	0.75	0.75	0.50	0.50	1.00	0.50	0.50	0.75	0.50	0.50	1.00	0.25	0.50	0.50		0.50	0.75
	III.3	0.25	0.75	0.50	0.50	0.50	0.75	0.50	0.75	0.50	0.25	1.00	0.75	0.25	0.50	0.75	0.75	0.25	0.25	0.50	0.50		0.75	0.75
	III.4	0.75	0.75	1.00	0.75	1.00	1.00	0.50	1.00	1.00	0.75	1.00	1.00	1.00	0.75	0.50	1.00	1.00	0.75	1.00	1.00		1.00	0.75
	III.5	0.75	0.75	1.00	0.75	0.75	1.00	0.75	0.75	1.00	1.00	1.00	1.00	0.75	1.00	0.75	1.00	0.75	0.75	1.00	0.50		1.00	0.75
	IV.1	0.75	0.75	1.00	0.75	0.75	0.75	0.75	1.00	1.00	0.50	1.00	0.75	0.75	0.75	0.50	0.50	0.25	0.50	0.75		0.50	0.50	0.75
	IV.2	1.00	0.75	1.00	1.00	0.50	1.00	0.50	1.00	0.75	0.50	1.00	1.00	1.00	0.75	0.75	1.00	0.50	0.50	1.00			0.50	0.75
	IV.3	1.00	0.50	0.75	0.50	0.75	0.50	0.50	0.50	0.75	0.50	1.00	0.75	0.50	0.50	0.75	0.50	0.00	0.00	1.00			0.25	0.75
	IV.4	0.75	0.50	0.50	0.50	0.75	1.00	0.50	0.25	0.50	0.75	0.75	0.50	0.75	0.50	0.50	0.50	0.25	0.00	1.00			0.25	0.50
	IV.5	0.50	0.25	0.75	0.50	0.25	1.00	0.50	0.75	0.25	0.75	0.75	0.75	0.50	0.50	0.50	0.50	0.00	0.00	1.00			0.25	0.50
	V.1	0.75	0.50	0.50	0.75	0.75	0.50	0.75	0.75	0.50	0.50	1.00	0.75	0.50	0.50	0.25	0.75	0.25	0.25	1.00			0.50	0.75
	V.2	0.75	0.50	0.75	0.75	0.75	0.50	1.00	0.75	0.75	0.50	1.00	0.75	0.50	0.75	0.50	0.50	0.00	0.00	1.00		0.50	0.50	0.75
	V.3	0.75	0.25	1.00	0.50	1.00	0.25	0.75	1.00	0.50	0.75	1.00	0.50	0.50	0.75	0.75	0.75	0.50	0.00	0.75			0.25	0.75
	V.4	0.75	0.25	1.00	0.75	1.00	0.75	1.00	0.50	0.75	0.75	1.00	0.75	0.50	1.00	0.50	0.50	0.25	0.50	0.75			1.00	1.00
	V.5	1.00	0.50	1.00	0.75	0.75	0.75	1.00	1.00	0.75	0.50	1.00	0.50	0.50	0.75	0.75	0.25	0.75	0.25	1.00			0.75	1.00

	student	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46
item	I.1	0.75	1.00	0.50	0.50	1.00	0.50	0.75	0.75	1.00	0.75	0.75	1.00	0.75	1.00	0.75	1.00	0.75	1.00	0.75	1.00	1.00	0.25	0.50
	I.2	0.50	0.50	0.75	0.25	0.75	0.25	0.50	1.00	0.25	0.75	0.50	1.00	0.50	1.00	0.50	1.00	0.50	0.75	0.50	0.75	1.00	0.50	0.75
	I.3	0.50	0.25	0.50	0.00	1.00	0.75	0.75	0.75	0.75	0.75	0.25	1.00	0.75	0.75	0.50	0.50	0.75	0.50	0.75	0.25	0.75	0.75	0.50
	II.1	0.25	0.50	0.25	0.75	0.50	0.00	0.25	0.50	1.00	0.50	0.25	0.25	0.75	0.50	0.75	0.50	0.25	0.50	0.75	0.50	0.50	0.00	0.25
	II.2	0.75	0.50	0.50	1.00	0.25	0.50	0.75	0.50	0.50	0.25	0.50	1.00	0.50	1.00	0.75	0.25	0.50	0.25	0.50	0.50	0.50	0.50	0.75
	II.3	1.00	0.50	0.75	0.75	0.25	0.50	0.75	0.50	0.50	0.75	0.75	1.00	0.50	0.25	1.00	0.75	0.50	0.50	0.50	0.50	1.00	0.25	0.75
	III.1	1.00	1.00	0.75	1.00	1.00	0.75	1.00	0.75	1.00	1.00	1.00	1.00	0.75	0.50	0.75	1.00	0.75	0.75	1.00	1.00	1.00	0.50	0.75
	III.2	0.75	0.75	0.50	0.50	1.00	0.25	0.75	0.75	1.00	0.75	0.75	1.00	0.75	0.25	0.75	1.00	0.75	0.50	1.00	0.50	0.75	0.50	0.75
	III.3	0.75	0.50	0.50	1.00	0.75	0.50	0.50	0.75	0.75	1.00	0.75	1.00	0.75	0.75	0.75	0.75	0.50	0.50	1.00	0.50	1.00	0.75	0.50
	III.4	0.75	1.00	0.75	1.00	1.00	0.75	1.00	0.75	1.00	1.00	1.00	1.00	0.75	1.00	1.00	0.75	0.75	0.75	1.00	1.00	1.00	1.00	1.00
	III.5	0.75	0.50	0.75	0.25	1.00	0.75	1.00	1.00	1.00	1.00	0.75	1.00	0.75	1.00	0.75	0.75	0.75	0.75	1.00	0.75	1.00	0.75	0.75
	IV.1	0.75	1.00	0.50	0.50	0.75	0.75	0.75	0.75	0.75	1.00	0.50	1.00	0.75	0.75	0.75	0.75	0.50	0.50	0.75	0.75	0.50	0.75	0.50
	IV.2	0.75	1.00	0.75	0.25	0.75	0.75	0.75	0.75	0.75	1.00	0.75	1.00	0.75	0.75	0.75	0.75	0.50	0.25	1.00	0.75	1.00	0.75	0.75
	IV.3	0.50	0.75	0.50	0.75	0.50	0.25	0.50	0.50	0.75	0.75	0.75	1.00	0.75	0.50	0.50	0.75	0.50	0.25	1.00	0.50	0.75	0.75	0.25
	IV.4	0.50	0.75	0.50	0.75	0.50	0.50	0.75	0.75		0.50	0.25	1.00	0.50	0.75	0.75	0.75	0.25	0.00	1.00	0.50	0.50	0.75	1.00
	IV.5	0.50	0.75	0.50	0.50	0.75	0.25	0.75	0.50	0.50	0.50	0.25	1.00	0.50	0.50	0.50	0.50	0.50	0.50	0.75		0.50	0.50	0.75
	V.1	0.75	0.50	0.25	0.75	0.50	0.50	0.50	0.75	0.75	1.00	0.50	1.00	0.50	0.75	1.00	0.75	0.75	0.50	1.00	0.50	0.75	0.50	0.75
	V.2	0.75	0.50	0.50	0.75	0.75	0.75	0.50	0.75	0.75	1.00	0.75	1.00	0.50	0.50	1.00	0.75	0.50	0.75	0.75	0.75	0.75	0.50	0.75
	V.3	0.75	0.50	0.50	0.75	0.75	0.50	0.50	0.50	0.75	1.00	0.25	1.00	0.50	0.25	0.75	0.75	0.25	0.50	0.75	0.50	0.75	0.25	0.50
	V.4	0.75	0.50	0.75	1.00	0.75	0.25	0.75	0.75	0.75	1.00	1.00	1.00	0.50	0.25	0.75	1.00	0.75	0.50	0.75	0.75	0.75	0.50	0.75
	V.5	0.75	0.50	0.75	0.50	1.00	0.25	0.75	0.75	1.00	0.75	1.00	1.00	0.50	0.50	0.75	0.50	0.50	0.75	1.00	1.00	0.75	0.25	0.75

	student	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69
item	I.1	0.50	0.50	0.75	1.00	1.00	0.75	1.00	0.75	0.50	0.75	0.50	1.00	0.75	1.00	0.25	1.00	0.75	1.00	0.75	1.00	0.75	0.75	0.00
	I.2	0.50	0.25	0.75	1.00	0.75	0.75	0.50	0.50	0.75	0.50	0.25	1.00	0.50	0.75	0.75	0.75	0.50	1.00	0.50		0.75	0.75	0.25
	I.3	0.75	0.50	0.50	0.75	0.75	0.75	1.00	0.75	0.50	1.00	0.50	1.00	1.00	0.50	0.50	0.75	0.25	0.75	0.75	0.50	0.50	0.50	0.00
	II.1	0.00	0.75	0.25	0.50	0.75	0.50	0.00	0.00	0.25	0.25	0.00	0.75	0.00	0.50	0.25	1.00	0.75	0.25	0.50	0.75	0.25	0.50	0.25
	II.2	0.75	0.50	0.75	0.50	1.00	0.50	1.00	0.50	0.25	0.75	0.25	0.75	0.75	0.75	0.75	1.00	1.00	0.50	0.50	1.00	0.75	0.25	0.75
	II.3	0.25	0.75	0.50	0.50	0.75	0.50	0.50	0.75	0.50	0.75	0.50	0.75	0.75	0.75	0.50	1.00	0.75	0.75	0.50	0.75	0.75	0.50	0.75
	III.1	0.75	0.75	0.75	1.00	1.00	0.75	0.75	1.00	0.75	1.00	0.75	1.00	0.75	1.00	0.00	1.00	1.00	1.00	0.75	1.00	1.00	0.50	1.00
	III.2	0.75	0.25	0.50	0.75	1.00	0.75	0.75	0.75	0.50	0.25	0.50	1.00	0.50	1.00	0.25	0.75	0.75	0.75	0.50	0.75	0.50	0.50	0.75
	III.3	0.25	0.50	0.75	0.75	0.75	0.75	0.75	0.75	0.50	0.50	0.50	1.00	0.50	0.50	0.25	1.00	0.75	0.50	0.50	0.75	0.50	0.50	0.75
	III.4	1.00	0.75	0.50	1.00	1.00	0.75	1.00	0.75	0.75	0.75	0.50	1.00	1.00	1.00	0.50	1.00	1.00	1.00	0.75	1.00	1.00	0.75	1.00
	III.5	1.00	1.00	0.75	1.00	1.00	0.75	1.00	1.00	0.75	0.75	0.50	1.00	1.00	1.00	0.00	1.00	0.50	1.00	0.50	1.00	1.00	0.75	0.75
	IV.1	0.75	0.75	0.50	1.00	1.00	0.75	1.00	0.75	0.50	0.50	0.50	1.00	0.75	0.75	0.00	1.00	0.75	0.75	0.50	0.75	0.75	0.50	0.75
	IV.2	0.75	0.75	0.75	1.00	0.75	0.75	0.75	0.75	0.75	0.75	0.50	1.00	0.75	0.75	0.75	0.75	1.00	0.75	0.50	1.00	1.00	0.50	0.75
	IV.3	0.75	0.50	0.50	1.00	0.75	0.75	0.75	0.50	0.50	0.25	0.50	1.00	0.50	0.50	0.50	1.00	0.75	0.75	0.50	0.75	0.75	0.50	0.75
	IV.4	0.00	0.50	0.75	1.00	0.50	0.75	0.75	0.25	0.50	0.00	0.25	1.00	0.50	0.75	0.50	0.75	0.75	0.75	0.50	0.75	0.75	0.25	0.75
	IV.5	0.00	0.50	0.75	1.00	0.75	0.75	0.50	0.50	0.50	0.25	0.25	1.00	0.50	0.50	0.75	1.00	0.75	0.75	0.50	0.75	0.75	0.25	0.50
	V.1	0.25	0.50	0.50	1.00	1.00	0.50	1.00	0.75	0.50	0.50	0.50	1.00	0.75	1.00	0.25	1.00	0.75	0.75	0.50	0.75	0.75	0.50	0.00
	V.2	0.25	0.50	0.75	1.00	1.00	0.75	1.00	0.75	0.75	0.75	0.50	1.00	0.75	0.75	0.50	1.00	0.75	1.00	0.75	0.75	1.00	0.25	0.25
	V.3	0.25	0.50	0.50	0.75	0.75	0.50	1.00	0.75	0.50	0.50	0.50	1.00	0.75	0.75	0.50	0.75	0.50	0.75	0.50	0.75	1.00	0.25	0.25
	V.4	0.25	1.00	0.75	1.00	1.00	0.75	1.00	0.75	0.50	0.75	0.75	1.00	0.75	1.00	0.25	1.00	0.25	1.00	0.50	1.00	0.75	0.75	0.25
	V.5	0.25	1.00	0.75	1.00	1.00	0.50	1.00	0.75	0.75	0.75	0.75	0.50	0.75	1.00	0.75	1.00	0.75	1.00	0.75	1.00	1.00	0.75	0.50

On the basis of these findings the teacher informed the degree program coordinator that the comfort of the classroom chosen for the course was considered very unsatisfactory from the students. The coordinator ordered a technical inspection of the room, then he took some actions: maintenance of the air conditioning, new lighting and furniture (especially chairs) were ordered (Act phase).

Respect to the results of the experimentation carried out during the two preceding years, the values of Γ'' show that:

- The "functionality of the tools used for lessons" (Item II.2), including also the adopted board-type did not represent a weakness ($\Gamma'' = 0.30$), which confirms the foregoing evidence.
- The behavior and style of the teacher during the 2006/2007 course, especially concerning interaction with the students, was in accordance with the foregoing experimental results. This meant that the teacher did not solicited interaction when not explicitly requested by students. This behavior was better appreciated by the students (see the values of Γ'' for the items III.4, III.5, and mainly IV.2, $\Gamma'' = 0.45$).

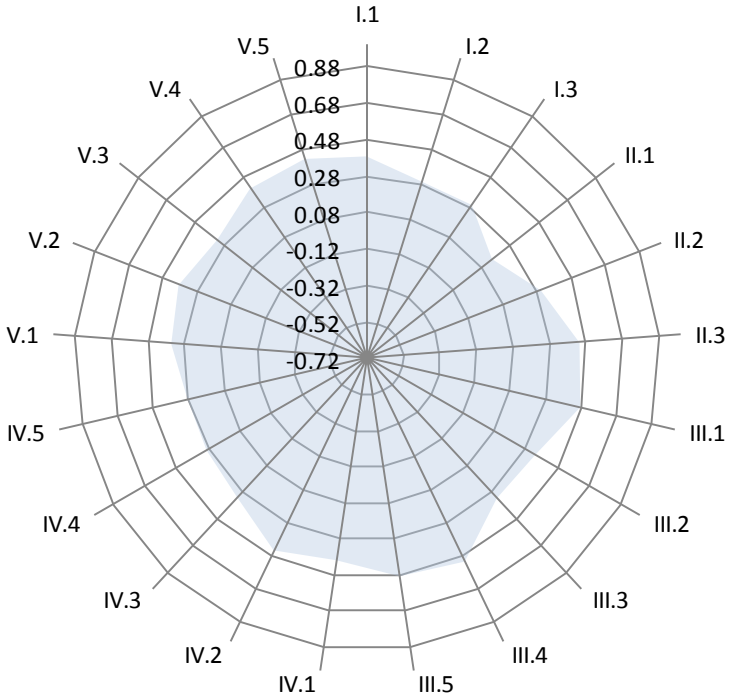

Figure 2. Weight-score-adjusted satisfaction index Γ'' vs. item.

Some additional results were:

- The weights-scores correlation coefficients for all students were centered around zero and lying in the range [–0.8, 0.62], see Figure 3. Note that for each student the correlation coefficient is calculated by considering all assigned scores and weights (see equation 10, Appendix A). This is important evidence confirming the fundamental assumption leading to the formulation of the index Γ'', according to which it is important to consider the weights-scores correlation.
- The average values of the importance weights attributed to the five course quality dimensions (Figure 4) range between 0.17 ('Reliability') and 0.23 ('Responsiveness'); the standard deviations range between 0.07 and 0.09. The dispersion for 'Assurance' was the lowest, but 11 outliers were found.

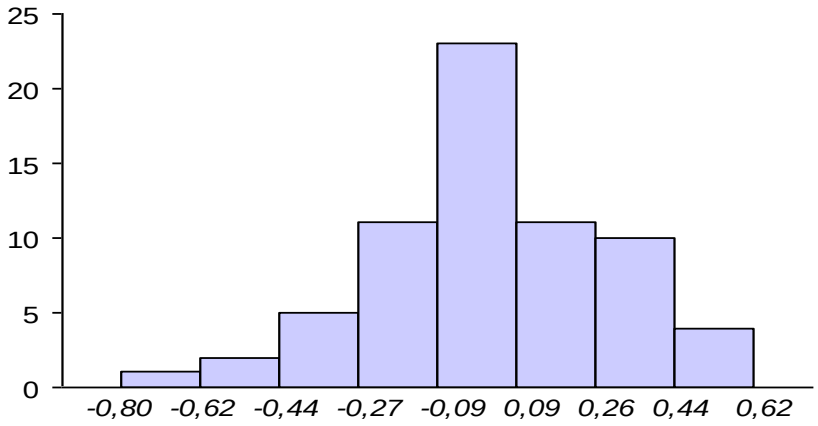

Figure 3. Histogram of the weights-scores correlation coefficients.

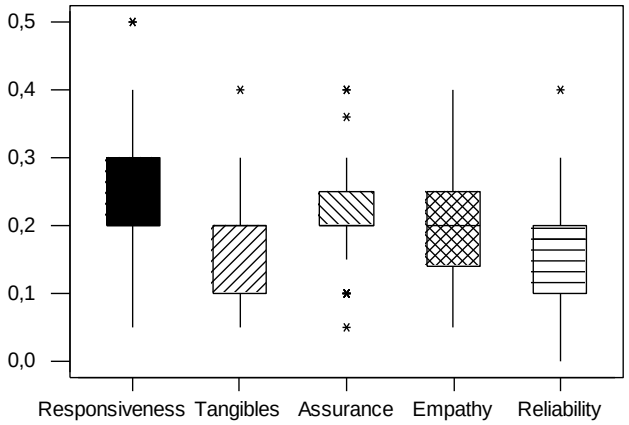

Figure 4. Box-whiskers plot of the importance weights vs. Quality Dimensions.

5. Conclusion

This paper has reviewed and updated the TESF methodology and its feedback tool. A new student satisfaction index has also been presented. The methodology allows university teachers, especially statistics educators, to acknowledge the "voice of the student", a valuable resource that a teacher can use to improve his or her course by carefully looking at its delivery modalities. Indeed, the analysis of student feedback allows the perceptions of students attending a course to be summarized, taking into account the relative importance accorded by them to the course quality dimensions. Furthermore, the weights assigned by the students to the course quality

dimensions allow the correlation between such weights (expression of the recognized importance and, therefore, of the expectations for the course) and the assigned scores (expression of perceptions about the completed lessons) to be taken into account.

The student satisfaction index formulated in the paper is intended as a way to summarize data collected through any feedback tool based on a Likert scale. This type of data is usually analyzed in terms of the average score per questionnaire item/section or the percentage of respondents for each scale level and per item (see, for example, Tsao, 2006 and Dutton & Dutton, 2005, respectively). The case study described in the paper showed how the methodology has been satisfactorily applied to a statistics course.

TESF responds to a self-assessment principle and predisposes teachers to accept students' opinions. It is complementary to other methods aimed at providing teachers with input to improve their performance; for example, those aimed at verifying the knowledge and skills that students acquire through a course. The authors recognize the importance of measuring the improvement of a course not only in student attitudes but also in student learning outcomes. However, a relevant assumption of this work is that student satisfaction affects learning, which is the final and most important result of a course.

Finally, the proposed methodology encourages students to be conscious of the active part they play in the courses they attend.

Appendix A. The student satisfaction index

Basic version of the satisfaction index

Let x_{ij} be the *j*-th order statistic ($j = 1,...,n$) of the scores assigned to the *i*-th item ($i = 1,...,m$) of the Evaluation Form by *n* students. It is a quantitative variable measured on the Likert scale.

Let $x_{\min}$ be the smallest score of the adopted Likert scale and $x_{\max}$ be the largest one. It is assumed that $x_{\min} = 0$ and $x_{\max} = 1$. If this is not the case, the scores must be normalized to the range [0,1].

Let $r_{ij} = j/n$ ($i = 1,...,m; \quad j = 1,...,n$) be defined as the relative rank, based on the score assigned to the *i*-th item.

Let ξ_{ij} be defined as the ratio between the cumulated score of the first *j* ranked scores and the maximum possible value, obtainable if all students assign $x_{\max}$:

$$\xi_{ij} = \frac{1}{n x_{\max}} \sum_{k=1}^{j} x_{ik} \qquad i = 1,...,m \quad j = 1,...,n \tag{1}$$

For each item the points with coordinates (r_{ij}, ξ_{ij}) are plotted on a Cartesian graph and a connecting line is drawn, starting from the origin (Figure 5). The area between the line and the horizontal axis is defined as the "Area of Satisfaction" (A_S) and is computed as:

$$A_{S_i} = \frac{1}{2} \sum_{j=0}^{n-1} \left(\xi_{ij} + \xi_{i(j+1)} \right) \left(r_{i(j+1)} - r_{ij} \right) \qquad i = 1,...,m \tag{2}$$

by having posed $r_{i0} = 0, \; \xi_{i0} = 0 \; \forall i.$ If all students assign the score $x_{\max}$, the line is the bisector line (Full satisfaction line). It bounds the so called "Area of Full Satisfaction" A_{FS} whose value is 0.5.

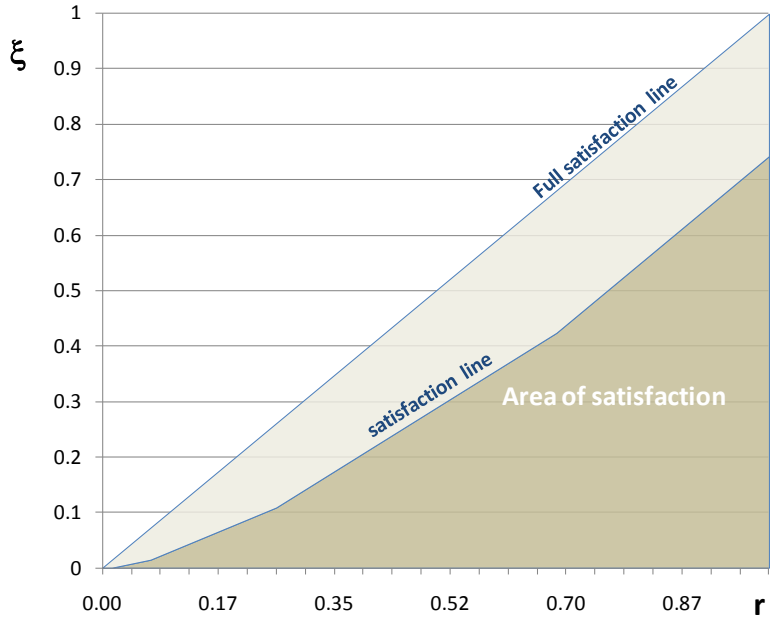

Figure 5. Areas of satisfaction and full satisfaction.

Figure 5 shows the satisfaction line for a subset of the data of Table 2 (Item I.1). It is possible to note that the line is made of five connected segments corresponding to the five score values of the adopted Likert scale.

The *Satisfaction Index* is defined as the ratio between the Area of Satisfaction and the Area of Full Satisfaction:

$$\Gamma_i = \frac{A_{S_i}}{A_{FS}} = \sum_{j=0}^{n-1} \left(\xi_{ij} + \xi_{i(j+1)}\right)\left(r_{i(j+1)} - r_{ij}\right) \qquad i = 1, \ldots, m \tag{3}$$

It is a summary measure of the satisfaction expressed by students and can be calculated for each item of the Evaluation Form, always ranging between 0 and 1. With the data of Figure 5, for the Item I.1 the Satisfaction Index is $\Gamma = 0.62$.

Weight-adjusted satisfaction index

According to the SERVQUAL model, a service's quality is the result of a comparison between the expectations and the perceptions related to that service. The application of this model often consists of measuring expectations and perceptions through two feedback tools submitted at distinct times (Sahney, Banwet & Karunes, 2006). The methodology presented in this article takes student expectations into account by considering the degree of importance that each student assigns to each of the course quality dimensions. In fact, the degree of importance attributed to each quality dimension is an expression of *desiderata* and, therefore, it contributes to the expectations in relation

to a service. Furthermore, measuring the perception of the 'reliability' dimension implicitly means taking expectations into account since they are influenced by the received promises.

Let w_{ij} $(i=1,...,m; \quad j=1,...,n)$ be the importance weight assigned by the *j*-th student to the quality dimension including the *i*-th item. Let $w_{\min}$ and $w_{\max}$ indicate the minimum and the maximum attributable weights, respectively. Assume that $w_{\min}=0; \quad w_{\max}=1$. If the adopted weight range differs from [0,1], it is necessary to adopt a normalization as for the scores.

To formally introduce the weights, the following variable is defined:

$$x'_{ij} = x_{ij} - (w_{ij} - w^*) \qquad \begin{matrix} i=1,...,m \\ j=1,...,n \end{matrix} \tag{4}$$

where w^* is the "neutral" weight meaning indifference between the *D* dimensions:

$$w^* = \left(w_{\max} - w_{\min}\right)/D \tag{5}$$

In case $D=5$, as assumed in this paper, $w^*=0.2$.

Basically, equation 4 implies a reduction of the assigned score if the assigned weight is higher than w^*, and vice-versa. This assumption will be better explained below. Based on equation 4:

$$x'_{\max} = x_{\max} - (w_{\min} - w^*) = 1-(0-0.2) = 1.2$$

$$x'_{\min} = x_{\min} - (w_{\max} - w^*) = 0-(1-0.2) = -0.8.$$

By introducing equation 4 into equation 1, one obtains:

$$\xi'_{ij} = \frac{1}{n x'_{\max}} \sum_{k=1}^{j} x'_{ik} \qquad \begin{matrix} i=1,...,m \\ j=1,...,n \end{matrix} \tag{6}$$

Therefore:

$$A'_{S_i} = \frac{1}{2} \sum_{j=0}^{n-1} \left(\xi'_{ij} + \xi'_{i(j+1)}\right)\left(r'_{i(j+1)} - r'_{ij}\right) \tag{7}$$

being r'_{ij} the relative rank of the students based on the weight-adjusted scores and by having posed $r'_{i0}=0, \; \xi'_{i0}=0 \quad \forall i$.

To calculate a weight-adjusted satisfaction index Γ', i.e. the satisfaction index with weight-adjusted scores, A'_{S_i} is divided by A'_{FS_i} which is the area of satisfaction obtained if for all students $x'_{ij} = x'_{\max}$.

$$\Gamma'_i = \frac{A'_{S_i}}{A'_{FS_i}} = \sum_{j=0}^{n-1} \left(\xi'_{ij} + \xi'_{i(j+1)}\right)\left(r'_{i(j+1)} - r'_{ij}\right) \tag{8}$$

It is easy to find that $A'_{FS_i} = 0.5$ and that $-\frac{2}{3} \le \Gamma'_i \le 1$.

To clarify the ideas behind the formulation of Γ'_i, imagine a situation in which one student assigns the same score, say 0.5, to two items *a* and *b* of the Evaluation Form. Furthermore assume that this student assigns an importance weight 0.4 to the item *a* and 0.1 to the item *b*. According to this situation it is possible to deduce that the student is equally satisfied with *a* and *b*, but he considers *a* four times more important than *b*. Therefore, the teacher should primarily act to improve *a*.

By applying the equation 4 to this illustrative case, x'_a is equal to 0.3 and x'_b is equal to 0.6. Consequently Γ'_a will be lower than Γ'_b, inducing the teacher to act on the item *a* more urgently than item *b*.

Weight-score-adjusted satisfaction index

Further reflections on data analysis reveal the possibility of considering the correlation between the scores and the weights assigned by each student. In practice, the scores can be further modified:

$$x''_{ij} = x'_{ij} - \rho_j \cdot \tau = x_{ij} - (w_{ij} - w^*) - \rho_j \cdot \tau \tag{9}$$

where ρ_j is the weights-scores linear correlation coefficient for the *j*-th student, calculated over the set of *m* items of the Evaluation Form, i.e.:

$$\rho_j = \frac{\frac{1}{m}\sum_{i=1}^{m}\left(x_{ij} - \bar{x}_{\cdot j}\right)\left(w_{ij} - \bar{w}_{\cdot j}\right)}{\sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(x_{ij} - \bar{x}_{\cdot j}\right)^2}\sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(w_{ij} - \bar{w}_{\cdot j}\right)^2}} \qquad j = 1,\ldots,n \tag{10}$$

Note that the scores may differ from item to item for the same student, while the weights may differ only from dimension to dimension; that is, they are equal for all items pertaining to the same dimension.

The constant τ in equation 9 is the unit step of the Likert scale used for the satisfaction judgment ($\tau = 0.25$ in the Evaluation Form, Appendix C). Equation 9 implies that a reduction of the score will result in presence of a positive correlation, while an amplification of the score will result in case of negative correlation. Accordingly, if a student tends to assign high scores to the items to which he or she assigned high importance weightings and low scores to the items corresponding to low weightings (positive correlation), or vice-versa (negative correlation) the application of equation 9 will help neutralize the bias effect of this tendency.

The updated limits of the scores are:

$$x''_{\max} = x'_{\max} + \tau = x_{\max} - \left(w_{\min} - w^*\right) + \tau = 1.45$$
$$x''_{\min} = x'_{\min} - \tau = x_{\min} - \left(w_{\max} - w^*\right) - \tau = -1.05$$

As in the transition from *x* to *x'*, the aim when moving from x' to x'' is to correct the scores to account for explicit or latent variables that could potentially affect them, to make them more homogeneous, and to drive teachers to the most appropriate improvement actions.

Consequently, by adopting the equation 9, the satisfaction index becomes:

$$\Gamma''_i = \frac{A''_{S_i}}{A''_{FS_i}} = \sum_{j=0}^{n-1}\left(\xi''_{ij} + \xi''_{i(j+1)}\right)\left(r''_{i(j+1)} - r''_{ij}\right) \tag{11}$$

where $A''_{FS_i} = 0.5$, based on the assumption that for all students $x''_{ij} = x''_{\max}$.

Figure 6 shows the weight-score-adjusted satisfaction line (the curved one) relative to the same data as in Figure 5. The satisfaction area is delimited by the satisfaction line and the horizontal axis. The two extreme cases for the satisfaction line are also depicted in the figure (lower and upper bounds). It easy to verify that $\Gamma''_{\min} = -0.724$, $\Gamma''_{\max} = 1$. For the case of Figure , $\Gamma'' = 0.39$.

In summary, Γ'' takes into account factors that could potentially influence the assessment, by considering the relative importance of different course aspects for students and the correlation between importance and perception.

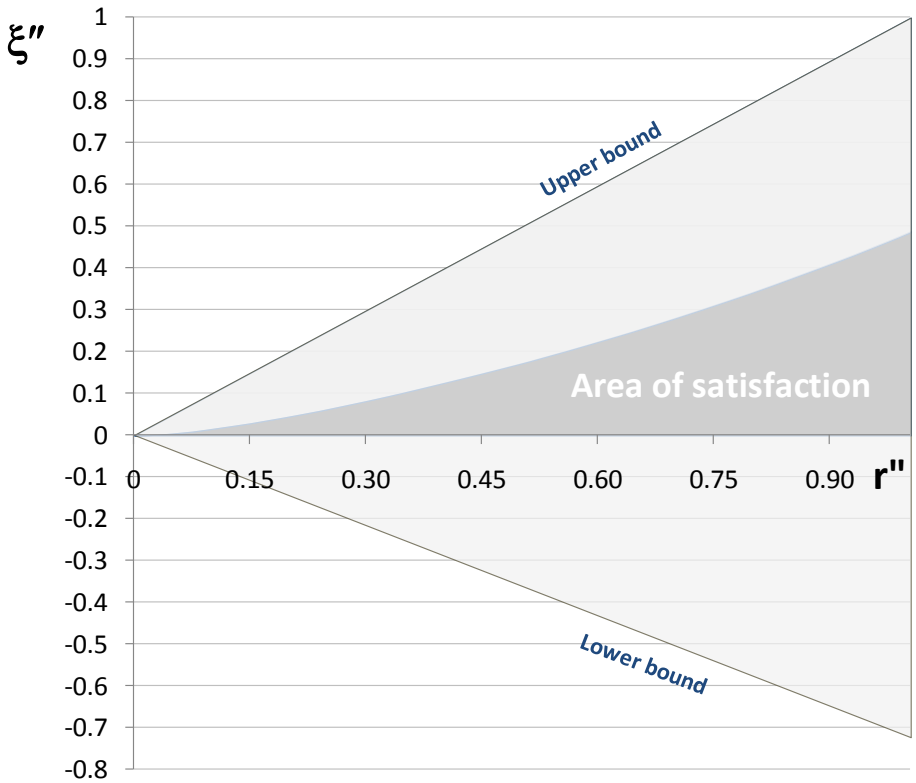

Figure 6. Areas of satisfaction for the second modified satisfaction index.

Appendix B. Weighting Grid

Regarding a generic university course, please distribute one unit to the following course quality dimensions, based on the importance level that you give to each of them.	
I. Responsiveness	
II. Tangibles	
III. Assurance	
IV. Empathy	
V. Reliability	
Total	1

Appendix C: Evaluation Form

What is your satisfaction level regarding the following items?		*Tick only one cell for each item*				
I – RESPONSIVENESS		**0**	**0.25**	**0.5**	**0.75**	**1**
I.1	Teacher's support to students during class					
I.2	Teacher's support to students out of class					
I.3	Management of class time					
II – TANGIBLES		**0**	**0.25**	**0.5**	**0.75**	**1**
II.1	Classroom comfort					
II.2	Functionality of the tools used for lessons					
II.3	Suitability of the recommended study material					
III – ASSURANCE		**0**	**0.25**	**0.5**	**0.75**	**1**
III.1	Teacher's mastery of the course topics					
III.2	Usefulness of the course topics					
III.3	Continuity of the course topics					
III.4	Students' ability to ask questions and make comments during class					
III.5	Teacher's warmth and availability					
IV – EMPATHY		**0**	**0.25**	**0.5**	**0.75**	**1**
IV.1	Teacher's brightness in explaining course topics					
IV.2	Teacher's ability to interact with students					
IV.3	Teacher's ability to attract students' attention					
IV.4	Teacher's tendency to pay individualized attention					
IV.5	Teacher's ability to retain students' attention					
V - RELIABILITY Concerning the following course aspects, have the promises made by the teacher, administrative staff, etc. been respected?		**0**	**0.25**	**0.5**	**0.75**	**1**
V.1	Aims					
V.2	Contents					
V.3	Study support materials					
V.4	Practices/exercises/laboratory work					
V.5	Course scheduling					

Appendix D. Course Quality Dimensions

Responsiveness

Willingness and readiness of the teacher and his or her staff to respond to student expectations during the course. Moreover, their capacity to face unforeseen events of an organizational nature without disappointing students; willingness to support students in their learning process during the course; willingness to transfer all necessary information about the course.

Tangibles

Includes the physical aspects of the course, such as the textbooks suggested by the teacher for studying the subject, the material tools that help the teacher explain the lesson (traditional blackboard, computer, slides, etc.), the room conditions in which lessons take place.

Assurance

Teacher's competence (i.e., the abilities and knowledge necessary to teach the course contents) and credibility (i.e., the ability to inspire loyalty and honesty). It also means courtesy (kindness and respect towards students) and safety (freedom from doubt and uncertainty).

Empathy

A teacher's ability to transfer his/her own knowledge to students clearly, and the ability to retain interest during a lesson. The teacher is committed to understanding student needs and pays attention to them individually. It also means access to lessons and contact with the teacher (timetable and place where lessons and student reception are delivered).

Reliability

The reliability implies an agreement between the perceived service and the trust that the student places on the course as a result of the information found (e.g., at the department, on websites, etc.) and received (e.g., by the teacher, by other students who attended the same course, etc.). It means the ability of the course to keep its promises, in the sense that everything that was initially announced regarding the course (its contents, class schedule, examinations modalities, recommended study materials, etc.) occurred without any unwanted change.

Appendix E. Evaluation Matrix

I – RESPONSIVENESS		**0**	**0.25**	**0.5**	**0.75**	**1**
I.1	Teacher's support to students during class	Totally absent	Scarce	Sufficient	Good	Excellent
I.2	Teacher's support to students out of class	Totally absent	Scarce	Sufficient	Good	Excellent
I.3	Management of class time	Totally absent	Scarce	Sufficient	Good	Excellent
II - TANGIBLES		**0**	**0.25**	**0.5**	**0.75**	**1**
II.1	Classroom comfort	Totally absent	Scarce	Sufficient	Good	Excellent
II.2	Functionality of the tools used for lessons	Totally absent	Scarce	Sufficient	Good	Excellent
II.3	Suitability of the recommended study material	Unsuitable	Scarcely suitable	Sufficiently suitable	Very suitable	Excellently suitable
III – ASSURANCE		**0**	**0.25**	**0.5**	**0.75**	**1**
III.1	Teacher's mastery of the course topics	Absent	Scarce	Sufficient	Good	Excellent
III.2	Usefulness of the course topics	Not at all useful	Scarce	Sufficient	Good	Excellent
III.3	Continuity of the course topics	Absent	Scarce	Sufficient	Good	Excellent
III.4	Students' ability to ask questions and make comments during class	No possibility	Scarce	Sufficient	Good	Excellent
III.5	Teacher's warmth and availability	Absent	Scarce	Sufficient	Good	Excellent
IV - EMPATHY		**0**	**0.25**	**0.5**	**0.75**	**1**
IV.1	Teacher's brightness when explaining course topics	Absent	Scarce	Sufficient	Good	Excellent
IV.2	Teacher's ability to interact with students	Absent	Scarce	Sufficient	Good	Excellent
IV.3	Teacher's ability to attract students' attention	Absent	Scarce	Sufficient	Good	Excellent
IV.4	Teacher's tendency to pay individualised attention	Absent	Scarce	Sufficient	Good	Excellent
IV.5	Teacher's ability to retain students' attention	Absent	Scarce	Sufficient	Good	Excellent
V - RELIABILITY		**0**	**0.25**	**0.5**	**0.75**	**1**
V.1	Aims	Not at all	Scarcely	Sufficiently	Good	Perfectly
V.2	Contents	Not at all	Scarcely	Sufficiently	Good	Perfectly
V.3	Study support materials	Not at all	Scarcely	Sufficiently	Good	Perfectly
V.4	Practices/exercises/laboratory work	Not at all	Scarcely	Sufficiently	Good	Perfectly
V.5	Course scheduling	Not at all	Scarcely	Sufficiently	Good	Perfectly

Acknowledgements

The authors, who equally contributed to this article, wish to sincerely thank the journal editor and the anonymous referees for their valuable comments that led to a substantial improvement of the article.

The work was supported by the project PRIN2008 "Innovation in service quality management: statistical approach and application in some fields of national interest" funded by the Italian Ministry of University and Research.

References

Alves, H., Raposo, M. (2007). "Student Satisfaction Index in Portuguese Public Higher Education". *The Service Industries Journal*, 27(6): 795–808.

Barone, S., Lo Franco, E. (2008). "Design and analysis of teaching experiments for course quality in the academic setting". Invited seminar at the University of Georgia, Athens (USA), August 21, 2008.

Barone, S., Lo Franco, E. (2009). "Design of a university course quality by Teaching Experiments and Student Feedback (TESF)". *Total Quality Management*. 20(7): 687-703.

Babbar S. (1995). Applying total quality management to educational instruction. A case study from a US public university. *International Journal of Public Sector Management*, 8(7), 35-55.

Bailey D., Bennett J.V. (1996). The realist model of higher education. *Quality Progress*, 29(11), 77-79.

Bier I.D., Cornesky R. (2001). Using QFD to construct a higher education curriculum. *Quality Progress*, 34(4), 64-68.

Brady, J.E., Allen, T.T. (2002), "Case Study Based Instruction of DOE and SPC". *The American Statistician*, 56(4): 312–315.

Braskamp, L.A., Wise, S.L., Hengstler, D.D. (1979). "Student Satisfaction as a measure of Departmental Quality". *Journal of Educational Psychology*, 71(4): 494–498.

Canic, M.J., McCarthy, P.M. (2000). Service quality and higher education do mix. *Quality Progress*, 33(9), 41-46.

Chinn P.W.U., Hilgers T.L. (2000). A conceptual model of service quality and its implication for future research. *Journal of Research in Science Teaching*, 37(1), 3-25.

Clewes D. (2003). A student-centred conceptual model of service quality in higher education. *Quality in Higher Education*, 9(1), 69-85.

Dutton, J., Dutton M. (2005). "Characteristics and Performance of Students in an Online Section of Business Statistics". *Journal of Statistics Education*, 13(3).

Emery C., Kramer T., Tian R. (2001). Customers vs. products: adopting an effective approach to business students. *Quality Assurance in Education*, 9(2), 110-115.

Ensby M., Mahmoodi F. (1997). Using the Baldrige award criteria in college classroom? *Quality Progress*, 30(4), 85-91.

Felder R.M., Brent R. (1999). How to improve teaching quality. *The Quality Management Journal*, 6(2), 9-21.

Fram E.H., Camp R.C. (1995). Finding and implementing best practices in higher education. *Quality Progress*, 28(2), 69-73.

Harvey L. (2003). Student feedback [1]. *Quality in Higher Education*, 9(1), 3-20.

Helms S., Key C.H. (1994). Are students more than customers in the classroom? *Quality Progress*, 27(9), 97-99.

Hogg, R.V., Newton, H.J., Cobb, G.W., Bryce, G.R. *et al*. (1999). "Let's use CQI in our Statistics Programs". *The American Statistician*, 53(1): 7–28.

ISO (2005). "ISO 9000:2005 Quality management system – Fundamentals and vocabulary standard".

Kanji, G.K., Bin Al Tambi A.M. (1999). Total Quality Management in UK higher education institutions. *Total Quality Management*, 10(1), 129-153.

Karapetrovic S., Rajamani D., Willborn W.W. (1999). University, Inc. To regain competitiveness, higher education must adopt industrial techniques. *Quality Progress*, 32(5), 87-95.

Kember, D., Leung, D.Y.P., Kwan, K.P. (2002). "Does the Use of Student Feedback Questionnaires Improve the Overall Quality of Teaching?" *Assessment & Evaluation in Higher Education*, 27(5): 411–425.

Leckey, J., Neill, N. (2001). "Quantifying Quality: the importance of student feedback". *Quality in Higher Education*, 7(1): 19–32.

Maguad, B.A. (2007). “Identifying the needs of customers in higher education”. *Education*, 127(3): 332–343.

Marozzi M. (2009). “A composite indicator dimension reduction procedure with application to university student satisfaction”, *Statistica Neerlandica*, 63(3): 258–268.

Masjuan, J. M., Troiano, H. (2009). “University students’ success: a psycho-sociological approach”. *Higher Education*, 58: 15–28.

Matthews W.E. (1993). Total quality management in academia. The missing element in higher education. *The Journal for Quality and Participation*, 16(1), 102-108.

Owlia, M.S., Aspinwall, E.M. (1996). “Quality in higher education: a survey”. *Total Quality Management*, 7(2): 161–171.

Panasuk, R.M., LeBaron, J. (1999), “Student Feedback: a tool for improving instruction in graduate education”, *Education*, 120(2): 356–368.

Parasuraman, A., Zeithaml, V.A. Berry, L.L. (1988). SERVQUAL: A multiply – item scale for measuring customer perceptions of service quality. *Journal of Retailing*, 64, Spring: 12–40.

Polese F., Moletta G. (2006). Value creation and related measurement in universities. An empirical application. *Total Quality Management*, 17(2), 243-263.

Politis, Y., Sisko, Y. (2004). Multi criteria methodology for the evaluation of a Greek engineering department. *European Journal of Operational Research*, 156(1), 223-240.

Quinn A., Lemay G., Larsen P., Johnson D.M. (2009). “Service quality in higher education”. *Total Quality Management*, 20(2): 139–152.

Redding, P. (2005). “The evolving interpretations of customers in higher education: empowering the elusive”. *International Journal of Consumer Studies*, 29(5): 409–417.

Richardson, J.T.E. (2005). “Instruments for obtaining student feedback: a review of the literature”. *Assessment & Evaluation in Higher Education*, 30(4): 387–415.

Rowley, J. (2003). “Designing student feedback questionnaires”. *Quality Assurance in Education*, 11(3): 142–149.

Sahney, S., Banwet, D.K., Karunes, S. (2006). “An Integrated Framework for Quality in Education: Application of Quality Function Deployment, Interpretive Structural Modelling and Path Analysis”. *Total Quality Management*, 17(2): 265–285.

Sirvanci M. (1996). Are students the true customers of higher education? *Quality Progress*, 29(10), 99-102.

Snee, R. D. (1990). "Statistical Thinking and Its Contribution to Total Quality". *The American Statistician*, 44(2): 116–121.

Toland, M.D., De Ayala R.J. (2005). "A multilevel factor analysis of students' evaluations of Teaching". *Educational and Psychological Measurement*, 65(2): 272–296.

Tsao. Y. (2006). "Teaching Statistics with Constructivist-Based Learning Method to Describe Student Attitudes Towards Statistics". *Journal of College Teaching & Learning*, 3(4): 59–63.

Tudor, G. E. (2006). "Teaching Introductory Statistics Online – Satisfying the Students". *Journal of Statistics Education*, 14(3).

Wallace J.B. (1999). The case for student as customer. *Quality Progress*, 32(2) 47-51.

Stefano Barone
Chalmers University of Technology
Department of Technology Management and Economics
mailto:stefano.barone@chalmers.se
University of Palermo
Department of Manufacturing Technology and Managerial Engineering
Viale delle Scienze, Edificio 8. 90128 Palermo, ITALY
mailto:stefano.barone@unipa.it

Eva Lo Franco
University of Palermo
Department of National Accounting and Analysis of Social Processes
mailto:eva.lofranco@unipa.it
