



Evaluation of university teaching through students' questionnaire

La valutazione della didattica universitaria attraverso il questionario studenti

Stefania Nirchi

University of Cassino and Southern Latium, Cassino, Italy
s.nirchi@unicas.it

ABSTRACT

In recent years, the evaluation of the University has become the focal point of an intense debate the essential element of which is based on the ability of the teaching faculty to carrying out its activities; on the one hand the possibility of improvement of the Italian education system in general and, on the other, the learning outcomes of the students. One of the methods now adopted by almost all universities in the measurement of the quality of teaching, is the evaluation of teachers by students. This is completed by means of an anonymous questionnaire designed to measure the opinions of students on the teachings and courses taken, in order to highlight the main positive aspects and also critical aspects, and propose any improvements. The main objective of this article is to underline any eventual problems related to the different use that universities can apply to the outcomes of the student's questionnaires, analyzing the main points of the survey and highlighting docimological characteristics of the instrument utilized.

Negli ultimi anni la valutazione delle Università è al centro di un intenso dibattito che vede come elemento imprescindibile dal quale partire, il ruolo chiave giocato dall'abilità del corpo docente nello svolgere la propria attività; un ruolo importantissimo sul quale viene misurata, da una parte la possibilità di miglioramento del sistema di istruzione italiano in generale e, dall'altra, gli esiti di apprendimento degli studenti. Uno degli step seguiti ormai da quasi tutte le Università per misurare la qualità della didattica è la valutazione dei docenti da parte degli studenti. Si tratta di una valutazione per la quale viene impiegato come strumento di rilevazione un questionario anonimo volto a misurare le opinioni dei ragazzi sugli insegnamenti, al fine di metterne in evidenza punti forti ed aspetti critici e proporre eventuali azioni di miglioramento. L'articolo, partendo dall'analisi dei punti salienti delle indagini fatte intorno al tema della valutazione dei docenti ed evidenziando le caratteristiche docimologiche dello strumento di rilevazione impiegato, ha lo scopo di sottolineare gli eventuali problemi collegati al diverso uso che le Università possono fare degli esiti dei questionari-studenti.

KEYWORDS

Student Evaluation of Teaching, Higher Education, Student Questionnaire, Evaluation Purpose, Use of Ratings.

Valutazione della didattica, Insegnamento universitario, Questionario degli studenti, Scopo della valutazione, Uso dei dati.

Introduction

In recent years, the evaluation of the educational activity of university professors has become, more than in the past, a milestone in the measurement of teaching skills and of the performance of professors and, in general, in monitoring the effectiveness and the efficiency of the system of higher education. To accomplish this, an anonymous questionnaire is administered to all students generally at the end of the courses. The questionnaire is characterized by structured questions (*rating scales*) and a series of open-ended questions, the aim of which is to gather opinions on a range of dimensions and aspects related to individual courses and their lecturers, in order to assess quality, effectiveness and efficiency. Due to the existing terminological confusion, the debate around the topic is very active. The importance of this type of evaluation is emphasized also by the numerous surveys that have completed over the years, and the role played by students – not only as a contribution to the teaching evaluation process – has become an essential element of the process itself. With regards the terminological confusion, it must be clarified that one often tends to confuse “*rating*” with “*evaluation*”, “*teaching*” with “*teacher*”, which require very different approaches and evaluation scenarios together. In fact, the terms “*rating*” and opinions refer to data which needs more evaluation interpretations to be correctly read, as opposed to the very concept of “*evaluation*” which already expresses a value judgment (eg. Benton & Cashin, 2012; Zambelli, 2006). The same reasoning is applied for the terms “*teaching*” (or *instruction*) related to the teaching-learning process, while “*teacher*” and teaching to people. Similar considerations can be taken if the Italian panorama in which one moves from expressions like “teaching evaluation by students”, to “evaluations of university teachers by students,” up to “survey of students’ opinions on the teaching”, is taken into account (operative proposal ANVUR - AVA, 2013)¹.

Since the theme of the role of students in the university teaching evaluation process is rather complex and articulated, a docimological analysis of the major assets and of the critical points of the students’ questionnaire, will be examined in this paper, the theoretical framework of which is the results of studies conducted over the last few years and that have encouraged the debate.

1. The national and international context

The evaluation of the educational activity completed by students (which finds its roots in the twenties, Guthrie, 1954), is mainly characterised as being completely voluntary. In fact, as from 2001 there are traces of a different type of evaluation: it becomes obligatory and is applied to a large number of Universities (Theall, Abrami & Mets, 2001). And in this same period a different use of the answers given by the students in the questionnaires, starts to take shape, not only applied to educational evaluations “in itinere” but also to a summative evaluation the aim of which was to examine (per es. Centra, 1993; Kulik, 2001; Marsh, 2007; Marsh & Roche, 1993; Richardson, 2005):

1 Cfr. http://www.anvur.org/attachments/article/26/5.%20RilevazioneOpinioneStud-Def_06_11_13.pdf consulted on the 21/03/2016.

- *The accountability* of the teaching activities carried out by the lecturers;
- The decisions taken regarding the recruitment and career development of lecturers;
- The possibility for students to choose to follow subjects on the basis of point given to lectures in the previous years.

Internationalisation, immigration, multiculturalism, ongoing education, have all contributed in the production of different methods of learning, different expectations, different orientations and different starting points. This heterogeneity of students has recalled the need of a more articulated and renewed educational system, capable of responding efficiently to the radical changes which have also effected higher education (Giovannini, 2010). Proof of this, at a European level is the “*European Space for Higher Education and Quality Assurance*”, particularly the paper “*Standards and Guidelines for Quality Assurance in European Higher Education Area*”, of the European Association for Quality Assurance in Higher Education (ENQA, 2009).

The participation of students in the process of Quality Assurance has increased in the years (as from the Prague Communiqué of 2001 to the document entitled “*Bologna with student yes*” published in 2007, and “*The National Unions of Students in Europe*” (ESIB, 2007), using different modalities from nation to nation, but also in universities of one same country. The most recent normative reference in Italy with regards the evaluation of University educational activities is represented by DPR 76/2010 which assigns to ANVUR – the Italian National Agency for the Evaluation of the University System and Research – the assignment of defining criteria and methodologies for the evaluation of university study programmes based on objective and certified criteria, and also of enhancing a positive attitude of students regarding teaching and learning (art.3, comma 1, letter b).

As already stated, even though the questionnaires are organised and constructed differently from one university to another, they are the main instrument used by students to assess teaching (eg. Centra, 1979; Martinez-Gomez, Carot Sierra, Jabaloyes & Zarzo, 2011; Pounder, 2007; Powell, Hunt & Irving, 1997; Seldin, 1999; Wright, 2006) and are nearly universally accepted.

Many studies have been conducted on the use of the questionnaires, in the course of years. In particular in 1993 Centra identifies four studies on the evaluation of university courses:

- *Purdue Rating Scale for Instructors* (Remmers and collaborators at Purdue University, 1927-1960);
- Voluntary questionnaires (Sixties) encouraged by individual lecturers to better teaching (Ory, 2000);
- Research on the validity and utility of the questionnaires (Seventies);
- First analysis of the answers given by the students (1980-1993).

As from the Nineties up to now, many studies on student evaluation have been conducted: these studies have been characterised in time to examine different aspects linked to this argument: the effectiveness of a module; the perception of those involved (thus students and lecturers); the type of competencies students must already possess in order to evaluate lecturers and courses; the influence of certain character traits of the lecturer, such as charisma, legal status; the number of students following a particular study programme, their age; if a course is compulsory or not, etc.

2. Problems related to students and questionnaires in pertinent dimensions

Keeping in mind the shared principle that “a good educational service is more complex than any list of qualities or characteristics can suggest” (Centra, 1003, p.4) it is understandable, when one tries to analyse the characteristics of the students’ questionnaire and the accuracy of the results regarding the effectiveness of the lecturer and the quality of the study programme, that one can find difficulties.

Every lecturer tend to define effectiveness based on one’s own personal opinion and by choosing a reference paradigm between behavioural, cognitive and/or social and constructive.

The students’ questionnaires are usually characterised by a series of items which are specifically measured by a Likert scale (from three to seven points) and in some cases, by a series of open-ended questions which are considered as important feedback for the lecturer. Even though a *multidimensional* agreement exists between lecturers and students, the criticality is related to the choice of the dimensions which have to be evaluated, to the “weighting” to be attributed to them, and to the utility of a general question on the effectiveness of the teaching/learning process. To embark the problematic nature of these effects, over the years other proposals were presented to accompany the results of the questionnaires with other types of evaluation (self-assessment, peer to peer evaluation, external observation), keeping in mind other potential *bias*, such as the number of students in a class, the students’ interests, workload, and so on (Marsh, 2007).

In particular:

- *The peer review*: commonly used in the United States and in Great Britain as a means of evaluation, it represents one of the instruments used in the process of Accreditation of the Universities. It consists of a visit by an expert of a group of experts, who also compiles a final evaluation report;
- *The teaching portfolio or teaching dossier*: this consists of a series of documents which present the professional career of the lecturer (the curriculum, work experience, evaluation obtained from students over the years, evaluation obtained from peer or other observers, type and quality of the didactic offer, area of research and publications, etc);
- *The focus group*, the aim of which is the discussion, with students, of the main and highlighted aspects which characterise the module;
- A series of questionnaires or interviews to students who are about to graduate, or to ex-students, who can evaluate the important contribution of a particular university course in their working career;
- *Self-assessment reports*: reports which are completed by the University with the aim of producing a report on its Quality and other particular aspects.

Students have always been recognised, in various studies, not only as a great source of information regarding teaching (for the length of the period involved in the teaching-learning process), but also because they are the best evaluators of the professional and ethical behaviour of the lecturer, and the ability of the latter to establish a proper communicative relationship, the level of attention and aroused involvement, the workload and of course materials, etc. (Eg. Braskamp & Ory, 1994). However, some researchers emphasize the inability of students to judge aspects such as the competency of lecturers in disciplinary matters and the use of teaching aids (Seldin, 1993), especially in the case of freshmen (Crumbley, Henry & Kratchman , 2001; Fish, 2005).

3. Possible bias which influence students' answers

As already stated, there are certain factors which, although not directly linked to the teaching-learning process, can affect the answers given by the students in the questionnaire (cfr. Lalla, Frederic e Ferrari, 2011, Zambelli, 2006):

- Factors related to the lecturers (sex, age, teaching experience, role, behavioural and character traits, difficulty of the subject taught, tendencies to give anticipations re the final exam, high grades);
- Factors related to the same students (sex, level, motivation, a sense of revenge on professors);
- Factors related to the module (course content and level of difficulty, options, level, number of students following the course, timetable).

In this context the analysis of the “Dr. Fox effect”² is also very interesting as a potential *bias*, on which many authors have carried out research but achieving often discordant results (Spooren et al., 2013). This term refers to the influence of the personality of the lecturers on the assessments from students. Through a meta-analysis carried out on 12 studies conducted between 1975 and 1982 in the Canadian context, Abrami et al. (1982) studied this issue by analyzing the relationship between the lecturers' expressiveness and opinions of students and the conclusion was that the expression of the lecturer has a significant impact on students' answers in the questionnaires, but has little effect on their performance.

In several studies factors related to gender (eg. Feldman, 1993) also emerged, and to the so-called *leniency hypothesis* (eg. Gump, 2007), namely the idea that students give more positive evaluations to professors who are less demanding in terms of workload and are more soft in exams and grades.

The existence of a *gender bias* in evaluations is also confirmed by the diversity of the results of evaluations completed by female and male professors. In fact, Sinclair's and Kunda's (2000) experiment shows that, when the standards are high and the teaching module is difficult, the hostility towards female professors is stronger than towards male professors. The two researchers conclude that students, when faced with a critical situation, tend to unconsciously adopt negative stereotypes to reduce the validity of the difficulty aroused, and to increase their personal self esteem (Kaschak, 1978).

Another important factor that can occur is an “*asymmetric information*”, in the sense that students are in the position of having to judge their lecturers followed, in a situation of a lack of appropriate knowledge (Olshavsky & Spreng, 1995). For example, they are not able to understand whether the professor has provided them with the relevant material for their study and /or if this has been correctly done. Therefore in such a situation of *asymmetric information*, it can result that different characteristics other than the quality of teaching can influ-

2 The term originated in an experiment in which Dr. Myron Fox had to give a one-hour lecture to a group of 55 people, including psychiatrists, educators and social workers. The theme of the lecture was the “game theory”. Dr Fox was a skilled orator with a clear voice and a pleasant and convincing attitude. Moreover, his lecturer proved to be brilliant, rich in amusing anecdotes, but lacked completely of any meaning. The reaction of the audience found Dr. Fox's speech interesting, clear, stimulating and nobody objected it had no real sense.

ence the judgement expressed by students in regards of their lecturers, such as for example, workload and/or grades (Marsh, 1987). In addition, students are led to judge positively the communication skills of lecturers regardless of the effect on their learning, of the teaching aids and material, of how lectures were carried out, whether in a traditional manner and /or using digital media, *case studies and working groups* (Yurkiewics, Hoefler and Byrne, 2006), or it can happen that they reward the system for attributing grades or even the appearance (Wright, 2000).

With regards grades given, many empirical studies on evaluation carried out in a number of American universities, have shown that students tend to give higher votes to professors whose courses they think they will obtain a high grade (Krautmann e Sander, 1997, Isely e Singh, 2005, Stapleton e Murkinson, 2001, Whitworth, Price e Randall, 2002). At this point it is easily deductable that high grades at examinations is proof of the lecturer's ability to teach. In actual facts empirical evidence shows that the same lecturer who delivers the same course in different classes will obtain the best evaluations in the class where the distribution of grades was higher. A higher correlation exists between the grade given at the exam and the vote given to the lecturer if we consider the relative grade of the student (rice, 1988; Greenwald and Gillmore, 1997). This evidence suggest that we are faced with a distorted behaviour, a situation which is defined as "*leniency bias hypothesis*", the inclination to consider and positively evaluate those who approve us rather than those who criticize us. This phenomenon is particularly important when the lecturers evaluated are established or not: the risk is that the second type (as the case of adjunct professors) are more liable to satisfy students by giving them high grades so as to obtain better evaluations which would help them in their career in the University; or the same professors could feel undermined in their own academic liberty so much as to avoid contrast with students and thus giving more importance to one's own behaviour rather than to the core of what is lectured.

Another risk factor which could create problems is characterised by the anonymity of the questionnaires, which, although this represents an element which must be guaranteed, can however create criticism since it is not clear if the student answering the questionnaire is a student who attends lectures on a regular basis or is a student who attends once in a while, and also if the student complains generally of all courses or of the same course in a continuous manner.

4. Objectives of the students' questionnaire

Usually questionnaires are distributed and are traceable for different motivations:

- To help professors to identify the weak points in their teaching method;
- To decide if a contract is to be renewed and reward the best professors;
- To monitor the efficiency of teaching methodologies and to activate adequate procedures to ensure the quality of the results;
- To help students in the choice of courses and lecturers.

Amongst the different elements listed above, the one which is a source of major critical issues, is represented by the decision taken by the governing authorities on the basis of the data gathered in the questionnaire (career advancement, accreditation, confirmation of an engagement); in respect of this aim a specific guardianship is asked for.

Previous studies have highlighted the risk that a similar use of the data can influence lecturers who tend to lighten the workload for the students, with the aim of obtaining higher votes and thus affecting the quality of the output data (eg. Beran & Rokosj, 2009; Centra, 1993; Seldin, 1984; Theall & Franklin, 2001). For this reason it is necessary to control these variables and to keep in mind that these distortions can occur and therefore take them in consideration in the design of the questionnaire, and at the same time students should be informed regarding the use of their judgement, motivating them and involving them as much as possible. What emerges is that the scenario in which the debate of the use of the data gathered from the questionnaires, is very lively, even though researchers believe that it is important to use different instruments to better teaching methods and also for the evaluation of the study programmes accomplished by the academic governing bodies (eg. Benton & Cashin, 2012). This criteria has been implemented by some universities: for example in a university in Ontario the only use of the results of the questionnaires is to consign them to the lecturer with the aim of bettering his didactic abilities (Kelly, 2012).

5. The use of the questionnaire's outcome as a support to teaching: what direction?

If the questionnaires on the quality of the teaching process are to provide useful information for the lecturer four conditions are necessary:

- The answers must give an added value;
- This added value must be important;
- What one must do to better himself must be clear;
- One must be motivated to better himself or herself.

At least studies show that there is a correlation between the effectiveness of teaching perceptions of students and those of teachers (eg. Nasser & Fresko, 2006).

Among the several items, those that undoubtedly provide valuable information are the open-ended questions, because they are more accurate and more functional to the improvement of teaching (Ory, 2000).

Perhaps what seems more problematic is represented by the moment when the lecturer receives the results of the questionnaire. Since the latter is typically administered towards the end of the course, the results are visible to the lecturer too late to operate any changes in the course, because the same is in the closing phase. In this regard, several studies show that a survey carried out in the middle of the course has a more positive outcome, as well as online assessment can allow more descriptive opinions (Venette, Sellnow & McIntyre, 2011). What could also be helpful is, apart from conducting the survey in another moment during the course, the opportunity to create a confrontation with other colleagues (eg. providing refunds / group discussions rather than providing the data to the individual teacher).

In the light of what has previously been exposed the growing importance of creating an integrated vision of the teaching and learning processes, is the basis of university teaching evaluation, because if it's true that learning is not an inducted and passive process, the weight and responsibility of a qualified education cannot fall completely on the study programmes. As we know each student plays an active role in the development of their learning processes and knowledge and for this reason it contributes to the process of continuous improvement. (Zambelli 2006, p. 56).

In conclusion

To improve university teaching, a prominent role for students must be recognised, to give them the possibility to express their own judgments regarding professors and courses, but the process of evaluation and self-evaluation should be undoubtedly enhanced by themselves, motivating them and involving them in the improvement even through reflection on their own way of learning (*what* and *how* they are learning). This allows for the activation of a confrontation and a discussion between professors and students that appears necessary to promote an effective process of continuous improvement of teaching and consequently of the learning outcomes achieved by students.

References

- Abrami, P.C. (2001). Improving judgements about teaching effectiveness using teacher rating forms. In M. Theall P. C. Abrami & L. A. Mets (Eds.). *The student ratings debate: Are they valid? How can we best use them?* [Special issue]. *New Directions for Institutional Research*, 109, 59-87.
- Abrami, P. C., d'Apollonia, S., & Rosenfield, S. (2007). The dimensionality of student ratings of instruction: What we know and what we do not. In R.P. Perry & J.C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 385-446). Dordrecht, The Netherlands: Springer.
- Abrami, P.C., Leventhal, L., & Perry, R.P. (1982). Educational Seduction. *Review of Educational Research*, 52(3), 446-464.
- ANVUR – AVA (2013). *Proposta operativa per l'avvio delle procedure di rilevamento dell'opinione degli studenti per l'A.A. 2013-2014*. Testo aggiornato al 6/11/2013. http://www.anvur.org/attachments/article/26/RilevazioneOpinioneStudDef_06_11_13.pdf.
- Aquario, D. (2007). *Università in cambiamento e valutazione. Il punto di vista degli studenti*. Padova: CLEUP.
- Benton, S.L., & Cashin, W. E. (2012). *Student Ratings of Teaching: A Summary of Research and Literature, IDEA Paper no.50*. Manhattan, Kansas State University: The IDEA Center.
- Beran, T. N., & Rokosh, J. L. (2009). Instructors' perspectives on the utility of student ratings of instruction. *Instructional Science. An International Journal of the Learning Sciences*, 37(2), 171-184.
- Braskamp, L., & Ory, J.C. (1994). *Assessing faculty work: Enhancing individual and institutional performances*. San Francisco, CA: Jossey-Bass.
- Brinko, K.T. (1990). Instructional consultation with feedback in higher education. *Journal of Higher Education*, 61, 65-83.
- Brockx, B., Spooren, P., & Mortelmans, D. (2011). Taking the grading leniency story to the edge. The influence of student, teacher, and course characteristics on student evaluation of teaching in higher education. *Education Assessment, Evaluation and Accountability*, 23, 289-306.
- Cashin, W.E. (1999). Student Rating of Teaching: Uses and Misuses. In P. Seldin (Ed.), *Changing Practices in Evaluating Teaching. A practical guide to improved faculty performance and promotion/tenure decisions* (pp.25-44). Bolton, MA: Anker.
- Centra, J.A. (1979). *Determining faculty effectiveness: Assessing teaching, research, and service for personnel decisions and improvement*. San Francisco: Jossey-Bass.
- Centra, J.A. (1993). *Reflective faculty evaluation: Enhancing teaching and determining faculty effectiveness*. San Francisco: Jossey-Bass.
- Cohen, P.A. (1983). Comment on A Selective Review of the Validity of Student Ratings of Teaching. *The Journal of Higher Education*, 54(4), 448-458.
- Crumbley, L., Henry, B., & Kratchman, S. (2001). Students' perceptions of the teaching of college teaching. *Quality Assurance in Education*, 9(4), 197-207.

- ENQA (2005). *Standards and Guidelines for Quality Assurance in the European Higher Education Area*. Helsinki:ENQA. La terza ed. del 2009 è scaricabile da http://www.enqa.eu/wp-content/uploads/2013/06/ESG_3edition-2.pdf.
- ESIB - The national unions of students in Europe (2007). *Bologna through student eyes 2007 edition*. London, UK: Bologna Process Committee 2005 - 2007.
- European University Association (2006). Annual Report. <http://www.eua.be/publications/eua-activity-reports.aspx>
- Feldman, K.A. (1993). College students' views of male and female college teachers: part II. Evidence from student's evaluations of their classroom teachers. *Research in Higher Education, 34*(2), 151-211.
- Feldman, K.A. (2007). Identifying Exemplary Teachers and Teaching: Evidence from Student Rating. In R.P. Perry & J.C. Smart (Eds.). *The Scholarship of Teaching and Learning in Higher Education: An Evidence-Based Perspective* (pp. 93-143). Dordrecht: Springer.
- Fish, S. (2005). Who is charge here? *The Chronicle of Higher Education, 51*(22), p. 2.
- Franklin, J., & Theall, M. (1989). *Who reads ratings: knowledge, attitudes, and practices of users of student ratings of instruction*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.
- Galliani, L. (Ed.). (2011). *Il docente universitario*. Tomo I e II. Lecce: Pensa MultiMedia editore.
- Giovannini, M.L. (2010). Imparare il "mestiere" di insegnare all'università: orientamenti e sfide. Learning to teach in Higher Education: Directions and challenges. In M.L. Giovannini (Ed.), *Insegnare all'Università. Modelli di formazione in Europa. Learning to Teach in Higher Education. Approaches and Case Studies in Europe* (pp. 1-56). Bologna: CLUEB.
- Giovannini, M.L. (2014). Il dibattito su valutazione ed evidenze: per un processo valutativo credibile e trasparente. *ECPS Journal, 9*, 101-126.
- Giovannini, M.L., & Rosa A. (2012). La valutazione di impatto dei progetti di formazione all'insegnamento dei docenti universitari: quali indicazioni dalle rassegne delle ricerche empiriche? *Giornale Italiano della Ricerca Educativa, V*(8), 93-104.
- Greenwald, A.G. (Ed.). (1997). Student rating of professors [Current Issues]. *American Psychologist, 52*(11), 1182-1225.
- Gump, S.E. (2007). Student Evaluations of Teaching Effectiveness and Leniency Hypotesis: A Literature Review. *Educational Research Quarterly, 30*(3), 56-69.
- Guthrie, E.R. (1954). *The evaluation of teaching: A progress report*. Seattle: University of Washington.
- Kelly, M. (2012). *Student Evaluations of Teaching Effectiveness: Considerations for Ontario Universities*. Vilfrid Laurier University. <http://www.cou.on.ca/publications/publications-archives/academiccolleague-papers/pdfs/ac-paper---student-evaluations-may-23-12>.
- Kulik, J.A. (2001). Student ratings: validity, utility and controversy. *New Directions for Institutional Research, 109*, 9-25.
- Lalla, M., Frederic P., & Ferrari, D. (2011). Students' Evaluation of Teaching Effectiveness: Satisfaction and Related Factors In M. Attanasio & V. Capursi (Eds.), *Statistical Methods for the Evaluation of University Systems* (pp. 113-129). Berlin Heidelberg DEU: Springer.
- Marsh, H.W. (1987). Students' evaluations of university teaching: research findings, methodological issues and directions for future research. *Journal of Educational Research, 11*, 253-388.
- Marsh, H.W. (2005). La valutazione della didattica da parte degli studenti. In R. Semeraro (Ed.), *La valutazione della didattica universitaria in Italia, in Europa, nel mondo. Atti di un convegno internazionale* (pp.121-157). Milano: FrancoAngeli.
- Marsh, H. W. (2007). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness. In R. P. Perry & J. C. Smart (Eds.), *The Scholarship of teaching and learning in higher education: An evidence-based perspective* (pp.319-383). Dordrecht, The Netherlands: Springer.
- Marsh, H.W., & Roche, .A. (1993). The use of students' evaluations and an individually structured intervention to enhance university teaching effectiveness. *American Educational Research Journal, 30*, 217-251.

- Martinez-Gomez, M., Carot Sierra, J.M., Jbaloyes, J., & Zarzo, M. (2011). A multivariate method for analysing and improving the use of student evaluation of teaching questionnaires: a case study. *Quality & Quantity*, 45, 1415-1427.
- Menges, R.J., & Brinko, K.T. (1986). Effects of Student Evaluation Feedback: A Meta-analysis of Higher Education Research. *Paper presented at the Annual Meeting of the American Educational Research Association*, San Francisco, CA, April 16-20, 1986.
- Ministero dell'Università e della Ricerca. Comitato nazionale per la valutazione del sistema universitario (2010). *Progettazione, implementazione e validazione di un questionario per la valutazione della didattica erogata a studenti universitari*, RdR 2/10. http://www.cnvsu.it/_library/downloadfile.asp?id=11775.
- Ministero dell'Università e della Ricerca Scientifica e Tecnologica. Osservatorio per la valutazione del sistema universitario (1998). *Valutazione della didattica da parte degli studenti*. RdR 1/98. <http://osservatorio.murst.it/didattic.htm>
- Nasser, F., & Fresko, B. (2002). Faculty views of student evaluation of college teaching. *Assessment & Evaluation in Higher Education*, 27(2), 187-198.
- Nasser, F., & Fresko, B. (2006). Predicting student ratings: The relationship between actual student ratings and instructors' predictions. *Assessment & Evaluation in Higher Education*, 31(1), 1-18.
- Ory, J.C. (2000). Teaching evaluation: past, present and future. *New Directions for Teaching and Learning*, 83, 13-18.
- Perry, R.P., & Smart, J.C. (Eds.). (2007). *The scholarship of teaching and learning in higher education: An evidence-based perspective*. Dordrecht, The Netherlands: Springer.
- Pounder, J.S. (2007). Is student evaluation of teaching worthwhile? An analytical framework for answering the question. *Quality Assurance in Education*, 15(2), 178-191.
- Powell, A.M., Hunt, A., & Irvin, A. (1997). Evaluating of courses by whole student cohorts: A case study. *Assessment & Evaluation in Higher Education*, 22(4), 397- 404. *North American Colleges & Teachers of Agriculture (NACTA) Journal*, 51(3), 10-14.
- Remmers, H.H. (1927). The Purdue Rating Scale for Instructors. *Educational Administration and Supervision*, 6, 399-406.
- Remmers, H.H. (1928). The relationship between students' marks and student attitudes towards instructors. *School and Society*, 28, 759-760.
- Remmers, H.H. (1930). To what extent do grades influence student ratings of instructors? *Journal of Educational Research*, 21, 314-316.
- Remmers, H.H., & Brandenburg, G. C. (1927). Experimental data on the Purdue Rating Scale for instructors. *Educational Administration and Supervision*, 13, 519- 527.
- Richardson, J.T.E. (2005). Instruments for obtaining student feedback: a review of the literature. *Assessment & Evaluation in Higher Education*, 30(4), 387-415.
- Seldin, P. (1984). *Changing practices in faculty evaluation*. San Francisco: Jossey-Bass.
- Seldin, P. (1993). The use and abuse of student ratings of professors. *The Chronicle of Higher Education*, 39(3), 40.
- Seldin, P. (1999). *Changing practices in evaluating reaching. A practical guide to improved faculty performance for promotion/tenure decisions*. Boston, MA: Ancher.
- Semeraro, R., & Aquario, D. (2011). L'autovalutazione della didattica da parte dei docenti. Presentazione dei risultati di una ricerca esplorativa condotta presso l'Università di Padova, *ECPS Journal*, 3, 25-51.
- Spooren, P., Brocx, B., & Mortelmans, D. (2013). On the Validity of Student Evaluation of Teaching: The State of the Art. *Review of Educational Research*, 83(4), 598-642.
- Theall, M., Abrami, P.C., & Mets, L.A. (Eds.). (2001). The Student Ratings Debate: Are They Valid? How Can We Best Use them? *New Directions for Institutional Research*, 109, 1-6.
- Theall, M., & Franklin, J. (2001). Looking for Bias in All the Wrong Places: A Search for Truth or a Witch Hunt in Student Ratings of Instruction? *New Direction for Institutional Research*, 109, 45-56.
- Venette, S., Sellnow, D., & McIntyre, K. (2010). Charting new territory: Assessing the online frontier of student ratings of instruction. *Assessment & Evaluation in Higher Education*, 35, 101-115.
- Wright, R.E. (2006). Student evaluations of faculty: concerns raised in the literature, and possible solutions. *College Student Journal*, 40(2), 417-422.

- Wright, S.L., & Jenkins-Guarnieri, M.A. (2012). Student evaluations of teaching combining the meta-analyses and demonstrating further evidence for effective use. *Assessment & Evaluation in Higher Education*, 37(6), 683-699.
- Zambelli, F. (2006). Strumenti e metodi di valutazione: una analisi critica. In R. Semeraro (Ed.). *La valutazione della didattica universitaria. Docenti e studenti protagonisti in un percorso di ricerca* (pp. 26-60). Milano: FrancoAngeli.

