A national student survey for the Italian higher education system

Serafina Pastore, Amelia Manuti & Anna Fausta Scardigno

Published online: 17 Oct 2021.

Submit your article to this journal ☑

View related articles ☑

View Crossmark data ☑

Routledge
Taylor & Francis Group

A national student survey for the Italian higher education system

Serafina Pastore ⓘD, Amelia Manuti ⓘD and Anna Fausta Scardigno ⓘD

Department of Education, Psychology, Communication, University of Bari Aldo Moro, Bari, Italy

ABSTRACT

Despite the strong criticisms, mass student surveys play an important role in the quality assurance process of national higher education systems. While some national higher education systems have a long history of using student surveys to inform teaching and course quality improvement (e.g. the UK, Australia), in Italy, where the quality assurance system has been put in place at both state and institutional levels, a national student survey has never been implemented. As a consequence, higher education institutions have used, and still use, only student evaluations of teaching. If, on the one hand, these local surveys tend to overlap with the rationale and structure of a national survey, on the other hand, they prevent universities from comparing quality indicators. Given the strong drive to use student surveys as a quality assurance mechanism, the present article reports a study aimed to design and probe a new questionnaire to be used at the national level. A total of 572 final year students enrolled in a public university participated in the study. Data were examined using a principal component analysis. Study results could set the groundwork for a critical debate on changes and improvements in the quality assurance process.

Introduction

Over the past decades, mass student surveys have received considerable interest in the higher education context (Harvey 2003; Kane, Williams, and Cappuccini-Ansfield 2008; Gibbs 2010; Buckley 2012). Some higher education systems have a long history of using student surveys (Marsh, Overall, and Kesler 1979; Sadler 2017), and have regularly used them to inform teaching and course quality improvements (e.g. in Australia or the UK). Historically, student surveys 'were introduced for various reasons relating to accountability, such as justifying the investment of public money and assuring the quality of provision to a range of stakeholders as well as contributing to informing students regarding their choice of institution' (Yorke 2013, 6). Generally identified as a proxy for teaching and learning quality, mass student surveys have dominated educational policy (and educational research) for years.

Addressing different aspects of student experience regarded as important (e.g. facilities, student service, learning resources), student surveys are crucial to gather valid and reliable evidence on the strengths and weaknesses of higher education institutions. Student surveys have been linked to quality assurance mechanisms and accreditation systems in order to enhance

the comparability and compatibility of higher education structures and processes across the world. Within the quality assurance framework, and through the harmonization inspired by the Bologna Process (Jungblut, Vukasovic, and Stensaker 2015; Maassen and Stensaker 2019), several efforts have been made to connect, in a coherent and balanced way, student feedback to course/program improvement (Harvey 2003; Maki 2015; Sadler 2017).

Despite the intrinsic limitations and strong criticisms addressed by different research studies (Yorke 2009; Richardson 2005, 2013; Tight 2013; Canning 2015; Arthur 2020; Borch, Sandvoll, and Torsten 2020), student satisfaction surveys still provide a gauge for universities to ascertain students' satisfaction with their university experience; consequently, the debate on student surveys is still in great turmoil (Merry et al. 2013; Bell and Brooks 2018).

Entering this debate, the present paper reports on a validation study of a student survey in the Italian higher education context, where the lack of a national student survey represents a significant gap in the quality assurance system (especially in terms of comparable levels of quality across the country).

Quality assurance and student surveys

At the top of the policy agenda in many countries across Europe and beyond, quality in higher education has inspired a lively debate. Moving from the original attempts to define the multi-faceted concept, the major efforts, over the last 20 years, have been directed towards the harmonization of different conceptualizations of quality, as well as towards the development of instruments to measure and monitor quality in the higher education contexts (Stensaker 2008; Manatos, Sarrico, and Rosa 2017; De Vincenzi, Garau, and Guaglianone 2018; Borch 2020).

The impact of neo-liberalism and the emphasis on different aspects like efficiency, productivity, auditing and accounting have led to quality assurance being considered as a fundamental tool for higher education institutions (Amaral et al. 2009). At the same time, the recognition that quality assurance would/should lead higher education enhancement has raised a variety of perspectives about the relationship between quality assurance and quality enhancement (e.g. as distinct, even opposite, activities, in terms of progression, or as integral parts of the same process: Williams 2016; Tight 2020). If, on the one hand, academic staff (and students) made notable efforts to improve the quality of teaching and learning, on the other hand, it has to be recognized that higher education institutions have a considerable interest in the development and use of a wide range of measures of quality assurance. Concepts such as learning outcomes, labour outcomes and placement, student retention, and student satisfaction generally, have been recognized as evidence to be provided in institutional accountability (at national and international level), as well as for the improvement of higher education institutions at the micro and the meso levels.

In this framework, student surveys of their course experience in higher education, as pointed out by Howson and Buckley (2017), have dominated the policy sphere through league tables and key performance indicators (especially in some nationals like the UK). The universities' need for insights into socio-economic, political and cultural impacts on students' experience from their perspective has driven the development of student surveys. Thus, quality indicators for measurable performance outcomes, such as the student experience, perceptions or satisfaction, have been overlooked in the arguments about quality assurance. Indeed, over the years, student surveys have been widely criticized, problematized in terms of validity, reliability and gender bias (Yorke 2009; Richardson 2005, 2013), and sometimes radically reviewed (e.g. in the UK in 2007 and then in 2018).

While recent research suggests new approaches to assess teaching quality (Kane et al. 2013), some theoretical and practical issues still remain unresolved (e.g. the overlapping focus on teacher efficacy, student perceptions or student satisfaction). Furthermore, the intrinsic limitations

of the measurement of satisfaction via student feedback survey have to be acknowledged: students' satisfaction with their educational experience, similarly to customer satisfaction, is the result of a complex set of factors (personal perceptions, conceptions, assumptions). Recent studies point to how feedback, on the backdrop of the quality assurance process, still represents one of the most problematic aspects of students' university experience. At the same time, feedback is a tricky task for teachers because, even though they recognize its utility, they often do not use feedback to review their instructional practice (Bell and Brooks 2018).

The recent shift in the conceptualization of feedback has led to radical review of assessment and feedback practices focusing on teacher and student feedback literacy (Carless and Boud 2018; Dawson, Carless, and Lee 2021). However, there are some other issues to be thoroughly considered with the feedback gathered through student surveys (Richardson 2005): it is important to clearly define how feedback will be interpreted and used; if the feedback will be made public or not; and to what extent students are involved in the feedback process. Some studies maintain that students do not understand the rationale of this process and therefore have troubles using and contextualizing feedback information (Alderman, Towers, and Bannah 2012; Porter 2011). Another important issue is related to the administration time of student surveys: while research evidence clearly shows how the immediacy of feedback positively impacts on teachers' and students' responsiveness, as well as on their perceptions of feedback usefulness (Merry et al. 2013; Dawson, Carless, and Lee 2021), the use of student evaluations as summative measures for quality assurance is persistent.

Despite the problems, the concerns and the criticalities highlighted by the literature, and despite the lack of consensus among scholars and practitioners, student surveys remain the tool most utilized in quality assurance systems in higher education. Policymakers, university managers and academics, and sometimes public opinion, consider student surveys, and therefore student evaluations, as the most important indicators for educational quality. Not surprisingly, demands for students to provide feedback about the different aspects of the academic experience have increased and led to more evaluation activities, as well as to more emphasis on expectations related to evaluations and to their impact on educational policies (Saunders 2011).

While using student evaluations or student feedback as summative measures to determine retention, promotion and teaching competence, or as a formative tool for the improvement of teaching and learning (Alzafari and Kratzer 2019; Darwin 2021), is a matter of concern, in Italy the debate on this topic is almost silent and student evaluations appear to be less used than expected. If, on the one hand, student evaluations have formally allowed each higher education institution to collect evidence on their students' experience regarding teaching and learning and act accordingly in order to improve their offer, on the other hand, the lack of a national student survey has hindered the comparison of institutions and programs and led to an unbalanced quality assurance system in Italy. To address this gap, the current study reports the development of a national student survey.

The Italian higher education system and the quality assurance process

Student evaluations of teaching and programmes were introduced in the Italian higher education system in the early 1990s. However, they became a central part of the quality assurance system in 2010 when the University Reform Law introduced a quality assurance system clearly aligned with the Bologna Process. This evaluation system is aimed to ensure a continuous improvement of educational quality through systematic documentation and monitoring of teaching, research and management of higher education institutions across the country (Figure 1).

In this framework, student ratings of teaching have been progressively recognized as one of the most effective mechanisms to improve and shape the quality of teaching and courses. With the implementation of the quality assurance process, since 2012–2013 the National Agency for
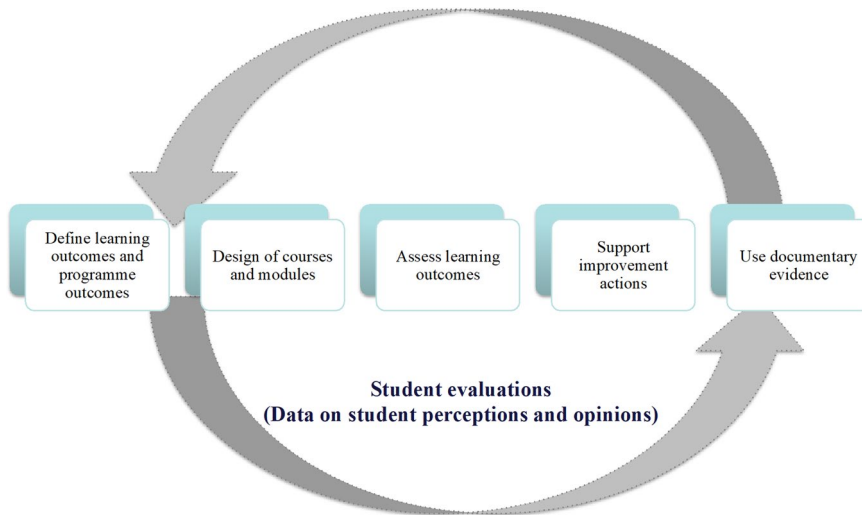
Figure 1. The quality assurance process in the Italian higher education system.

Table 1. Core dimensions of the Italian student evaluation survey.

Teaching quality	Teacher quality	Student interest
1. Adequacy of rudiments and prior knowledge for new learning 2. Study load in terms of credits 3. Adequacy of teaching material 4. Clarity in the description of assessment of learning methods	5. Observance of lesson times 6. Students' motivation 7. Clarity in the communication/explanation 8. Usefulness of supplementary educational activities 9. Correctness of the information reported on the web site of the degree course 10. Availability to provide to students further clarifications and explanations	11. Students' interest for topics/ subject matter

the Assurance of Higher Education and Research (ANVUR) has created a student questionnaire (with 11 questions) focused on three main dimensions: teaching quality, teacher quality and student interest (Table 1).

This survey was purposed to:

- gather information to improve teaching practices;
- evaluate teaching quality;
- inform stakeholders for university evaluation.

The original aim of the survey was to allow a comparison of institutions and programmes at national level in terms of external mechanism of the quality assurance process (Harvey 2006). However, higher education institutions have had the chance to freely change the original 11 questions. Thus, each higher education institution progressively introduced new and different questions focused on aspects such as academic support, organization, management and learning facilities. The common body of questions (and therefore of comparable indicators of quality across the country) disappeared. Even though the structure of these surveys is more similar to a national student survey than to a local student evaluation of teaching, the data have been used only at local level.

While in other higher education systems student evaluations of teaching are differentiated (e.g. in the UK the National Student Survey gathers final year undergraduates opinions about their university experience, and student evaluation of teaching through the Teaching Excellence

Framework assesses the quality of teaching in higher education institutions), in Italy there has been a conflation of two different questionnaires (the national student survey and the student evaluation of teaching) in one survey. The local surveys, developed on the model of the national student survey, have a variable structure across higher education institutions and generate data useful only for internal evaluation. The lack of an empirical basis for assessing the impact of quality changes in higher education institutions impedes a reasonable comparison across the country and time.

These surveys are administrated not only to final year students, but to all students enrolled in graduate and post-graduate courses. Therefore, students had (and still have) to fill in the same questionnaire for each teaching module/course they attended: the survey is mandatory for all students before an examination. The over-production of student evaluations, as well as deficiencies in the data collected, has resulted in the data gathered being not really informative and useful for quality assurance at local and national levels. While these surveys were focused on teaching, the information gathered on other dimensions were not comparable between different institutions across the country. The lack of accurate and robust evidence deeply impacted the quality assurance process (Beecham 2009; Kelly 2012; Hornstein 2017).

The lack of teaching standards in Italy, as well as the lack of a framework for teaching quality (like the Teaching Excellence Framework in England) makes it very difficult to interpret gathered data and to take reasoned, defensible choices about improvement actions at course, programme and institution levels. Over the last years, ANVUR has tried to review its student survey, reducing the 11 questions, differentiating between students who have or have not attended courses, and changing the rating scale for student satisfaction. These changes, however, have not reconsidered the rationale and the structure of the survey; therefore, a national student survey for the Italian higher education system is still missing. The present study aims to develop a national student survey for the Italian higher education system.

The validation study

A comparative analysis of the *National Student Survey* (UK) and the *Course Experience Questionnaire* (Australia) was carried out. The National Student Survey (NSS), launched in 2005 in the UK, is carried out in the final year of undergraduate study. It is based on the Course Experience Questionnaire (CEQ), the annual government-mandated survey sent to the previous year's graduating students. The Australian higher education system has a long history of collecting student satisfaction surveys: student evaluations, indeed, have been regularly used to inform teaching and course quality improvements since the 1970s (Marsh, Overall, and Kesler 1979).

Despite the methodological criticisms (e.g. what is actually measured, the use of only positive questions), the NSS and the CEQ are used to inform student choice, quality assurance and quality enhancement. Therefore, they were used in the current study to guide the survey development.

The NSS questionnaire was changed in 2017: the number of questions has been reduced and new sections (i.e. learning community and student voice) have been added. The CEQ, since 2016, was included in the Graduate Outcomes Survey (GOS). This national survey of the Australian Government Department of Education and Training is purposed to collect information on graduates' labour market outcomes and further study activities.

NSS and CEQ target different thematic areas (Table 2): while the NSS has currently 28 core questions, the CEQ, which originally comprised 25 items, has now 19. Both surveys include a global measure of satisfaction with the course and ask to students to express their agreement on a five-point scale.

This analysis highlights the need for a survey that could balance organizational and managerial elements, as well as those aspects related to the most innovative dimensions of the

Table 2. Student evaluations surveys: a comparison.

National Student Survey	Course Experience Questionnaire
• The teaching on my course (4 items) • Learning Opportunities (3 items) • Assessment and Feedback (4 items) • Academic Support (3 items) • Organisation and Management (3 items) • Learning Resources (3 items) • Learning Community (2 items) • Student Voice (6 items)	• Overall Satisfaction (1 item) • Good Teaching Scale (6 items) • Generic Skills Scale (6 items) • Graduate Qualities Scale (6 items)

Table 3. Italian national survey dimensions.

Course organization	Assessment of learning and quality assurance	Learning support	Learning resources	Dublin descriptors
8 items: Teaching organization and management aspects (program, syllabus, information on the web site) in terms of coherence between what has been communicated and what has been realized during the course/module	8 items: Assessment criteria, modalities, strategies, feedback practice for the quality assurance process and for the improvement of student learning	5 items: Student workload; hands-outs, materials for students; teacher availability; tutor and other services for support student learning	4 items: spaces, libraries, learning tools, ICT, etc.	4 items: Student perception of learning activities meaningfulness, usefulness of contents, knowledge, and learning outcomes

teaching-learning aspects, such as the Dublin Descriptors. In the Bologna Process the Dublin Descriptors are generic statements on typical expectations of achievements and abilities associated with awards: these statements are extremely relevant in the instructional design phase (i.e. program, course or module) and in the assessment of student learning outcomes.

The new survey is made up of 29 items related to five different dimensions (Table 3):

- course organization (8 items);
- assessment and quality assurance (8 items);
- learning support (5 items);
- learning resources (4 items);
- Dublin descriptors (4 items).

The validation study was conducted in one university. The survey was preliminary subject to review by five experts in the quality assurance field. The content validity ratios (CVRs, Lawshe 1975) were computed. The results showed satisfactory substantive agreement and positive CVRs: then the questionnaire, with the 29 statements listed in a random order, was sent to 572 final year students enrolled in different bachelor's and master's degree courses.

The survey was administrated through a Moodle e-learning platform. Students were invited to fill in personal information (e.g. age, gender, year of enrolment) where requested, and to express their agreement with a range of statements about their learning experience during the course attended using a 4-point Likert scale (from 1 = completely disagree to 4 = completely agree). Students were informed about the aims of the study and were assured of anonymity, they voluntarily consented to participate according to the ethical guidelines required by law. All forms were valid for data processing.

A principal component analysis (PCA) was performed on students' scores on the five scales to assess the structure of measures developed and to reduce its dimensions. Before running the analysis, assumptions were checked by looking at the Kaiser Meyer Olkin test (>.09) and at the Bartlett test ($\chi = 2714.4$; $p = .000$) and both indices were good. Commonalities were also assessed.

Table 4. Italian student survey: items frequency (percentage distribution of responses).

A. Course organization	1	2	3	4
1. The organization of this course was sustainable for me	41.3	27.3	10.1	9.5
2. The teaching schedule was respected	2.3	13.6	46.7	25.6
3. Likely changes in the teaching activities were effectively communicated	4.8	12.6	33.7	20.7
4. The teacher arose my interest for his/her subject matter	6.6	25	28.4	18
5. The teacher clearly communicated the topic of his/her lesson	6	18.2	34.6	20.9
6. The teacher gave further explanations, if required	6	18.2	43	20.9
7. The course was aligned with the syllabus	4.9	16.1	42.2	24.4
8. Other teaching activities (e.g. seminars, case-studies, work-groups, etc.) were useful for the preparation of the exam	5	15.4	29.4	17
B. Assessment of learning and quality assurance				
9. The teacher clearly defined his/her assessment modalities	3.7	16.1	42.8	74.6
10. The teacher explained what he/she expected as learning outcomes	2.9	13.4	57.8	28.9
11. The teacher explained the criteria he/she would have used for grading	7	21.8	26	26.6
12. The assessment criteria were reported in the syllabus	6	14.1	35.3	25.3
13. During the course, there were interim assessments	12.2	25.6	28.1	14.5
14. During the course, I was allowed to give a feedback to my teacher	11.2	22.3	32.8	13.2
15. The teacher considered my feedback after the course to revise his/her practice	12.8	14.3	9.7	15.5
16. My feedback was used to improve the course	9.7	18	10.1	25.8
C. Learning support				
17. The study workload of this course was proportional to the university credits	11.4	13.8	16.3	9.5
18. The teaching material (e.g. hands-out and supplementary readings) was adequate to prepare the exam	17.8	20.9	26.7	6.8
19. The teacher was available during the reception hours	16.5	21.5	25	8.5
20. The teacher was available to communicate through email	19.4	20.7	22.9	8.3
21. During the course, I relied upon some student services (e.g. tutorship, mentoring, peer to peer learning)	8.7	19	31	13.6
D. Learning resources				
22. The learning space in the classroom was adequate	16.5	21.5	25	8.5
23. Other learning spaces (e.g. library, laboratories, study rooms, pc stations) were adequate	16.7	23.6	22.7	9.3
24. ICT supported my learning	13.6	23.3	17.8	11.2
25. My previous knowledge was sufficient to understand the contents of this course	5	14.7	29.3	21.7
E. Dublin descriptors				
26. This course helped me to learn useful theoretical contents for my future profession	5	14.7	29.3	21.7
27. This course helped me to gain precious knowledge for my future profession	7.6	18.4	26.2	18.6
28. This course helped me to learn how to make critical judgements for my future profession	7.6	20	28.7	14.5
29. This course helped me to learn communication and life-long learning skills useful for my future profession	15.1	24.6	24	6.8

Factors with eigenvalues equal or greater than one and loadings equal or greater than 0.5 were considered and retained. Intercorrelations between the items showed values greater than .25, therefore Oblimin rotation with Kaiser normalisation was chosen assuming consistent relations between the dimensions investigated. Table 4 shows the matrix with factor loadings.

Descriptive analysis, means, standard deviations and Cronbach alpha for all dimensions are shown in Table 5. Internal consistency of the whole scale was assessed with the Cronbach alpha showing a positive value ($a = .93$).

Results

Table 6 shows the new factorial structure deriving from PCA and retaining 23 out of 29 items. Five factors were found accounting for the 63.6% of the total variance.

Factor one was labelled 'course organization'. It originally contained eight items but only five of them originally loaded under this factor. This factor accounts for the basic aspects that are most important when considering the organization of the course from the student perspective (e.g. the teaching schedule, the possibility to receive from the teacher information about the course). It is interesting to note that two items that were initially coded as part of the learning support ('*The teacher was available during the reception hours*' and '*The teacher was available to communicate through email*') were significantly associated with the organization of teaching activities in the course.

Table 5. Means, standard deviation, Cronbach alpha, correlations ($N=175$).

Variables	Mean (SD)	1	2	3	4	5
1. Course organization	3.31 (0.47)	($\alpha=0.85$)				
2. Assessment of learning and quality assurance	3.14 (0.53)	0.473**	($\alpha=0.82$)			
3. Feedback learning and support	3.23 (0.51)	0.531**	0.511**	($\alpha=0.67$)		
4. Learning resources	2.86 (0.55)	0.600**	0.562**	0.660**	($\alpha=0.61$)	
5. Learning experience	3.20 (0.54)	0.604**	0.384**	0.466**	0.549**	($\alpha=0.85$)

*$p<0.001$ (2-tailed)
**$p<0.005$ (2-tailed).

Table 6. Principal factor analysis on student survey.

	Course organization	Assessment of learning and quality assurance	Student feedback and teaching practice	Learning resources	Learning experience
The teacher was available during the reception hours	.812				
The teacher was available to communicate through email	.759				
The teacher gave further explanations, if required	.703				
The teaching schedule was respected	.671				
The course was aligned with the syllabus	.666				
Likely changes in the teaching activities were effectively communicated	.621				
The teacher clearly communicated the topic of his/her lesson	.615				
The teacher explained what he/she expected as learning outcomes		.881			
The teacher explained the criteria he/she would have used for grading		.837			
The teacher clearly defined his/her assessment modalities		.792			
The assessment criteria were reported in the syllabus		.738			
The teacher considered my feedback after the course to revise his/her practice			.862		
My feedback was used to improve the course			.814		
During the course, I was allowed to give a feedback to my teacher			.798		
The study workload of this course was proportional to the university credits				.771	
During the course, I relied upon some student services (e.g. tutorship, mentoring, peer to peer learning)				.742	
The organization of this course was sustainable for me				.726	
My previous knowledge was sufficient to understand the contents of this course				.673	
This course helped me to gain precious knowledge for my future profession					.905
This course helped me to learn useful theoretical contents for my future profession					.831
This course helped me to learn how to make critical judgements for my future profession					.826
Other teaching activities (e.g. seminars, case-studies, work-groups, etc.) were useful for the preparation of the exam					.608
This course helped me to learn communication and life-long learning skills useful for my future profession					.607

Oblimin rotated solution (Kaiser normalization).

Factor two was labelled 'assessment of learning and quality assurance'. It also initially encompassed eight items but the PCA reduced them to four. In this dimension we found items related to students' perception of what is expected from them in terms of quality of the learning outcomes and criteria for the assessment of learning (e.g. '*The teacher explained what in terms of your learning he/she will assess*'; '*The teacher explained which criteria he/she will use for grading*'; '*The teacher clearly defined his/her assessment modalities*'; '*The assessment criteria were reported in the syllabus*'). These items are clearly aligned with the idea that assessment is a quality assurance mechanism and that formal requirements are relevant in terms of transparency, rigor, and course quality.

This structure shows also that probably the concept of assessment is much more focused on the need for students to understand what teachers want from them rather than giving feedback to their mentors. The item '*During the course, there were interim assessments*' was discarded. The last three items ('*The teacher considered my feedback*'; '*My feedback was used to improve the course*', and '*During the course I was allowed to give a feedback to my teacher*') were split out from this factor. In the new structure, these items refer to the possibility to give feedback to the teacher to improve the course and teaching practices. The third factor was therefore re-labelled 'Student feedback and teaching practice'.

Factor four was re-labelled 'learning resources'. This dimension, in the new structure, is composed by four items related to the affordances that students have in the academic context (e.g. tutorship, mentorship), as well as to the proportionality of study workload to the university credits. It is interesting that this dimension encompassed an item originally in the 'course organization' dimension ('*The organization of this course was sustainable for me*'), and another one that referred to a basic personal learning resource namely previous knowledge ('*My previous knowledge was sufficient to understand the contents of this course*') that contributed to reassure students about the possibility to overcome difficulties.

The last dimension that was originally linked to the 'Dublin Descriptors' was renamed 'learning experience' and evoked the significance of university experience (through different teaching activities) as a privileged context to gain knowledge and to develop skills that could be potentially useful for the professional future (e.g. '*This course helped me to gain precious knowledge for my future profession*'; '*This course helped me to learn useful theoretical contents for my future profession*'; '*This course helped me to learn how to make critical judgements for my future profession*').

In the new structure of the student survey, it is interesting to note the reduction of the items originally under the scale 'assessment of learning and quality assurance'. The more formal aspects of quality assurance are clearly shifted from the role that assessment and feedback have for student learning.

Another interesting aspect is related to the new dimension in this version of the student survey: the section dedicated to a re-elaboration of the Dublin descriptors in terms of learning experience. This dimension was designed and included in the survey to understand if students, after almost a decade of the Dublin descriptors introduction to the Italian higher education system, have become more familiar with a different way of conceiving and practicing the teaching and learning process. Students demonstrated their understanding of the rationale of this educational innovation and considered the aspects related to the quality of higher education in terms of what they expect to learn and to acquire for their future professional life. This should be an aspect to be considered in the design of more responsive national student surveys.

Limitations

Although the findings of the PCA showed sound and reliable measures, some limitations must be discussed.

Firstly, due to the reduced sample and to the contextualized nature and the cross-sectional design of the study, no formal generalization is possible. However, it would be of interest to

know whether similar results would be found in other universities in the Italian higher education context. Future research is needed to investigate if, and to what extent, the results of this study could be extended. Moreover, further item analyses are also needed to improve the quality and the accuracy of the measures, and their external and concurrent validity: thus, a replication study is highly encouraged. This is a required step to enhance not only the quality of higher education but also to ensure a more coherent and cohesive system in Italy.

Despite the numerous criticisms related to national student surveys, without a survey designed to compare higher education institutions the risk to use different measures for different objects is high, as well as the risk to use existing student course evaluations for ranking institutions instead of improving course quality (internal evaluation).

Secondly, qualitative research would be helpful to investigate if, and to what extent, aspects and factors defined in the proposed survey might impact student perception of quality in the higher education field.

Conclusions

In the accountability era, a wide range of practices and strategies has been developed to assess, measure, rank and improve higher education institutions (Bamber and Anderson 2012; Wright and Jenkins-Guarnieri 2012; Yorke 2013). In this framework, 'informal discussion or conversations; formal qualitative sessions, such as focus groups, facilitate discussions or suggestion boxes; representative or consultative committees; questionnaires' (Harvey 2003, 3) are some of the strategies most utilized and embedded in higher education. However, the primary data sources for evaluation, national benchmarking and decision-making in higher education remain student evaluations. The aim of these evaluations is to measure lecturers' performance and document student experiences, allowing the comparison of standards across a university (Arthur 2020). Over the years, extensive research has been undertaken under this mechanism of quality assurance based on student opinions.

In the Italian higher education system, the lack of educational research on this field and the lack of a national student survey led to the development and use of multiple versions of student evaluations of teaching. While the lack of professional standards for teachers and academic staff impedes the use of student evaluations for appraisal purposes (e.g. tenure/promotion decisions), data gathered through these evaluations are not a robust and comparable measure of higher education institutions quality.

The present study has tried to fill this gap, shedding light on the use of students' surveys for quality assurance in Italy: these findings can be helpful to better understand incongruences and criticalities in the quality assurance process. Moreover, these study results can inform quality assurance practice encouraging ANVUR and the Italian universities in a deep revision of quality assurance mechanisms including the student survey.

Finally, although the context of this paper is the Italian higher education system, it has relevance to the international debate on teaching, learning and assessment practices in the higher education context. Embedding quality into higher education institutions is a challenging matter given the complexity of the higher education environment. The changing context of European higher education (especially after the Covid-19 pandemic) raises a host of important considerations for quality assurance.

The responses to questions related to the accuracy, consistency and applicability of the national student survey could instigate a reflection on convergences and discrepancies that appear across the policy documents, the practice of quality assurance and the current needs of higher education institutions. Cross-national studies are necessary to understand the challenges to quality in the European higher education area and the ways to overcome them.

This study provides a basis for more research that may support educational policy-makers and educational practitioners for quality enhancement in the European higher education context.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Serafina Pastore is an Associate Professor in the Department of Education, Psychology, Communication, University of Bari (Italy). Her research focuses the intersections of assessment practice, teacher education and educational policy as they operating in the context of school and university innovation.

Amelia Manuti is an Associate Professor in Organizational Psychology in the Department of Education, Psychology, Communication, University of Bari (Italy). Her research interests refer to the meaning of working in school-to-university and university-to-work transitions.

Anna Fausta Scardigno is an Associate Professor in Sociology of Education in the Department of Education, Psychology, Communication, University of Bari (Italy). Her recent research interests focus on university credentialism and migrants' and refugees' educational policies.

ORCID

Serafina Pastore http://orcid.org/0000-0001-8598-021X
Amelia Manuti http://orcid.org/0000-0001-5308-7494
Anna Fausta Scardigno http://orcid.org/0000-0002-0020-8966

References

Alderman, L., S. Towers, and S. Bannah. 2012. "Student Feedback Systems in Higher Education: A Focused Literature Review and Environmental Scan." *Quality in Higher Education* 18 (3): 261–280. doi:10.1080/13538322.2012.730714.

Alzafari, K., and J. Kratzer. 2019. "Challenges of Implementing Quality in European Higher Education: An Expert Perspective." *Quality in Higher Education* 25 (3): 261–288. doi:10.1080/13538322.2019.1676963.

Amaral, A., G. Neave, C. Musselin, and P. Maassen, eds. 2009. *European Integration and the Governance of Higher Education and Research*. Dordrecht: Springer.

Arthur, L. 2020. "Evaluating Student Satisfaction – Restricting Lecturer Professionalism: Outcomes of Using the UK National Student Survey Questionnaire for Internal Student Evaluation of Teaching." *Assessment & Evaluation in Higher Education* 45 (3): 331–344. doi:10.1080/02602938.2019.1640863.

Bamber, V., and S. Anderson. 2012. "Evaluating Learning and Teaching: Institutional Needs and Individual Practices." *International Journal for Academic Development* 17 (1): 5–18. doi:10.1080/1360144X.2011.586459.

Beecham, R. 2009. "Teaching Quality and Student Satisfaction: Nexus or Simulacrum?" *London Review of Education* 7: 135–146. doi:10.1080/14748460902990336.

Bell, A. R., and C. Brooks. 2018. "What Makes Student Satisfied? A Discussion and Analysis of the UK's National Student Survey." *Journal of Further and Higher Education* 42 (8): 1118–1142. doi:10.1080/0309877X.2017.1349886.

Borch, I. H. 2020. "Lost in Translation: From the University's Quality Assurance System to Student Evaluation Practice." *Nordic Journal of Studies in Educational Policy* 6 (3): 231–244.

Borch, I. H., R. Sandvoll, and R. Torsten. 2020. "Discrepancies in Purposes of Student Course Evaluations: What Does It Mean to Be 'Satisfied'?" *Educational Assessment, Evaluation and Accountability* 32 (1): 83–102. doi:10.1007/s11092-020-09315-x.

Buckley, A. 2012. *Making It Count: Reflecting on the National Student Survey in the Process of Enhancement*. York, UK: Higher Education Academy.

Canning, J. 2015. "A New Measurement and Ranking System for the UK National Student Survey." *Perspectives: Policy and Practice in Higher Education* 19 (2): 56–65. doi:10.1080/13603108.2015.1021401.

Carless, D., and D. Boud. 2018. "The Development of Student Feedback Literacy: Enabling Uptake of Feedback." *Assessment & Evaluation in Higher Education* 43 (8): 1315–1325. doi:10.1080/02602938.2018.1463354.

Darwin, S. 2021. "The Changing Topography of Student Evaluation in Higher Education: Mapping the Contemporary Terrain." *Higher Education Research & Development* 40 (2): 220–233. doi:10.1080/07294360.2020.1740183.

Dawson, P., D. Carless, and P. P. W. Lee. 2021. "Authentic Feedback Supporting Learners to Engage in Disciplinary Feedback Practices." *Assessment & Evaluation in Higher Education* 46 (2): 286–296. doi:10.1080/02602938.2020.1769022.

De Vincenzi, A., A. Garau, and A. Guaglianone. 2018. "Has the Quality of Teaching and Learning Processes Improved as a Result of the Implementation of Quality Assurance Coordinated by the State?" *Quality in Higher Education* 24 (1): 55–65. doi:10.1080/13538322.2018.1426382.

Gibbs, G. 2010. *Dimensions of Quality*. York, UK: Higher Education Academy.

Harvey, L. 2003. "Student Feedback." *Quality in Higher Education* 9 (1): 3–20. doi:10.1080/13538320308164.

Harvey, L. 2006. "Impact of Quality Assurance: Overview of a Discussion between Representatives of External Quality Assurance Agencies." *Quality in Higher Education* 12 (3): 287–290. doi:10.1080/13538320601051010.

Hornstein, H. A. 2017. "Student Evaluations of Teaching Are an Inadequate Assessment Tool for Evaluating Faculty Performance." *Cogent Education* 4 (1): 1304016. doi:10.1080/2331186X.2017.1304016.

Howson, C. K., and A. Buckley. 2017. "Development of the UK Engagement Survey." *Assessment & Evaluation in Higher Education* 42 (7): 1132–1144. doi:10.1080/02602938.2016.1235134.

Jungblut, J,., M. Vukasovic, and B. Stensaker. 2015. "Student Perspective on Quality in Higher Education." *European Journal of Higher Education* 5 (2): 157–180. doi:10.1080/21568235.2014.998693.

Kane, D., J. Williams, and G. Cappuccini-Ansfield. 2008. "Student Satisfaction Surveys: The Value in Taking an Historical Perspective." *Quality in Higher Education* 14 (2): 135–155. doi:10.1080/13538320802278347.

Kane, T. J., D. F. McCaffrey, T. Miller, and D. O. Staiger. 2013. *Have We Identified Effective Teachers? Validating Measures of Effective Teaching Using Random Assignment*. Seattle, WA: Bill & Melinda Gates Foundation.

Kelly, M. 2012. *Student Evaluations of Teaching Effectiveness: Considerations for Ontario Universities* (COU #866). Toronto: Council of Ontario Universities.

Lawshe, C. H. 1975. "A Quantitative Approach to Content Validity." *Personnel Psychology* 28 (4): 563–575. doi:10.1111/j.1744-6570.1975.tb01393.x.

Maassen, P., and B. Stensaker. 2019. "From Organised Anarchy to de-Coupled Bureaucracy: The Transformation of University Organisation." *Higher Education Quarterly* 73 (4): 456–468. doi:10.1111/hequ.12229.

Manatos, M,., C. Sarrico, and M. Rosa. 2017. "The Integration of Quality Management in Higher Education Institutions: A Systematic Literature Review." *Total Quality Management & Business Excellence* 28 (1–2): 159–175. doi:10.1080/14783363.2015.1050180.

Maki, P. L. 2015. *Assessment That Works: A National Call, a Twenty-First-Century Response*. Washington, DC: Association of American Colleges & Universities.

Marsh, H. W., J. U. Overall, and S. P. Kesler. 1979. "Class Size, Students' Evaluations, and Instructional Effectiveness." *American Educational Research Journal* 16 (1): 57–70. doi:10.3102/00028312016001057.

Merry, S., M. Price, D. Carless, and M. Taras. 2013. *Reconceptualising Feedback in Higher Education*. Abingdon, UK: Routledge.

Porter, S. 2011. "Do College Student Surveys Have Any Validity?" *The Review of Higher Education* 35 (1): 45–76. doi:10.1353/rhe.2011.0034.

Richardson, J. 2005. "Instruments for Obtaining Student Feedback: A Review of the Literature." *Assessment & Evaluation in Higher Education* 30 (4): 387–415. doi:10.1080/02602930500099193.

Richardson, J. 2013. "The National Student Survey and Its Impact on UK Higher Education." In *Enhancing Student Feedback and Improvement in Tertiary Education*, edited by M. Shah and C. S. Nair, 76–84. Abu Dhabi: Commission for Academic Accreditation Quality Series No. 5.

Sadler, D. R. 2017. "Academic Achievement Standards and Quality Assurance." *Quality in Higher Education* 23 (2): 81–99. doi:10.1080/13538322.2017.1356614.

Saunders, M. 2011. "Insights into Programmatice Evaluative Practice in HE: A Commentary." In *Reconceptualising Evaluation in Higher Education: The Practice Turn*, edited by M. Saunders, P. Trowler, and V. Bamber, 113–124. Maidenhead: Open University Press.

Stensaker, B. 2008. "Outcomes of Quality Assurance: A Discussion of Knowledge, Methodology and Validity." *Quality in Higher Education* 14 (1): 3–13. doi:10.1080/13538320802011532.

Tight, M. 2013. "Students: Customers, Clients or Pawns?" *Higher Education Policy* 26 (3): 291–307. doi:10.1057/hep.2013.2.

Tight, M. 2020. "Research into Quality Assurance and Quality Management in Higher Education." In *Theory and Method in Higher Education Research*. Vol. 6 of *Theory and Method in Higher Education Research*, edited by J. Huisman and M. Tight, 185–202. Bingley: Emerald Publishing Limited.

Williams, J. 2016. "Quality Assurance and Quality Enhancement: Is There a Relationship?" *Quality in Higher Education* 22 (2): 97–102. doi:10.1080/13538322.2016.1227207.

Wright, S., and S. Jenkins-Guarnieri. 2012. "Student Evaluations of Teaching: Combining the Meta-Analyses and Demonstrating Further Evidence for Affective Use." *Assessment & Evaluation in Higher Education* 37 (6): 683–699. doi:10.1080/02602938.2011.563279.

Yorke, M. 2009. "Student Experience' Surveys: Some Methodological Considerations and an Empirical Investigation." *Assessment & Evaluation in Higher Education* 34 (6): 721–739. doi:10.1080/0260293080 2474219.

Yorke, M. 2013. "Surveys of the Student Experience and the Politics of Feedback." In *Reconceptualising Feedback in Higher Education*, edited by P. Merry, M. Price, D. Carless, and M. Taras, 6-18. Abingdon, UK: Routledge.