

Tuning Journal for Higher Education

Volume 10, Issue No. 2, May 2023

DOI: <https://doi.org/10.18543/tjhe1022023>

Student and Teacher perceptions and experiences:
How do they align?

ARTICLES

Major increases in teachers' performance evaluations: Evidence from student evaluation of teaching surveys

Jaime Prieto, Rocío Guede-Cid, Ana I. Cid-Cid, and Santiago Leguey

doi: <https://doi.org/10.18543/tjhe.2299>

Received: 22 December 2021

Accepted: 26 February 2023

E-published: May 2023

Copyright

Copyright for this article is retained by the Publisher. It is an Open Access material that is free for full online access, download, storage, distribution, and or reuse in any medium only for non-commercial purposes and in compliance with any applicable copyright legislation, without prior permission from the Publisher or the author(s). In any case, proper acknowledgement of the original publication source must be made and any changes to the original work must be indicated clearly and in a manner that does not suggest the author's and or Publisher's endorsement whatsoever. Any other use of its content in any medium or format, now known or developed in the future, requires prior written permission of the copyright holder.

Major increases in teachers' performance evaluations: Evidence from student evaluation of teaching surveys

Jaime Prieto, Rocío Guede-Cid, Ana I. Cid-Cid, and Santiago Leguey*

doi: <https://doi.org/10.18543/tjhe.2299>

Received: 22 December 2021

Accepted: 26 February 2023

E-published: May 2023

Abstract:

Purpose: This exploratory study examined major increases in teachers' performance evaluations and their immediate impact on next year's score for those instructors that taught the same subject for at least two years in a row. The purpose was twofold. Firstly, to identify those Student Evaluation of Teaching (SET) survey items associated with major increases in teacher evaluations. Secondly, to examine if there is evidence of the use of these SET results by instructors to improve their teaching.

Design: The sample comprised SET survey ratings from one university over a five consecutive year period, for a total sample of 13,052 teacher evaluations and 3,893 teachers-subject observations under analysis. Frequency tables and Student's t-test were used for analysis.

Findings: The results highlighted the three SET survey items captured by the dimension of teaching methodology as those most closely related to major increases in teacher evaluations. Regarding the second objective, the results show no generalised response from teachers who experience major increases in SET ratings. This suggests

* **Jaime Prieto** (jaime.prieto@urjc.es), PhD, is a lecturer on information and communications technology at the Rey Juan Carlos University (Spain).

Rocío Guede-Cid (rocio.guede@urjc.es), PhD, is a lecturer on Mathematics and Statistics at the Rey Juan Carlos University (Spain).

Ana I. Cid-Cid (ana.cid@urjc.es), PhD, is a senior lecturer on Mathematics and Statistics at the Rey Juan Carlos University (Spain).

Santiago Leguey (corresponding author, santiago.leguey@urjc.es), PhD, is a senior lecturer on Mathematics and Statistics at the Rey Juan Carlos University (Spain).

More information about the authors is available at the end of this article.

Acknowledgements: None.

Funding: None.

Conflict of interest: None.

that the use of SET results is either limited or does not have an immediate measurable effect on student satisfaction.

Originality/Value: To the best of our knowledge, this was the first study to specifically examine major increases in teachers' performance evaluations and their immediate impact on next year's score based on evidence from SET surveys.

Keywords: Teacher evaluation; student evaluation; student evaluation of teaching; SET; teaching evaluation; higher education; university teaching; teacher performance evaluation; teaching excellence; SET surveys.

I. Introduction

Obtaining feedback from students through Student Evaluation of Teaching (SET) surveys is a widely extended practice in universities internationally that provides diagnostic feedback to instructors on the quality of their teaching.^{1,2} However, from the point of view of the teaching staff, there is no consistent evidence of the teachers' direction of opinion towards the use, validity and consequences of SET results.³ Published research examining instructors' attitudes to student ratings shows a broad range of responses, with teachers showing both positive and negative attitudes towards using SET surveys.^{4,5} This lack of consensus is amplified in many cases due to opinions at the extremes, finding from strongly supportive teachers to the use of SET surveys to extremely critical ones.⁶ It has been argued that the concerns of the instructors with the use of SET surveys is due to its dual usage for formative purposes (i.e., students' diagnostic feedback for improving teaching) as well as for summative purposes (i.e., administrative policymaking about faculty personnel and

¹ Rachel Johnson, "The authority of the student evaluation questionnaire," *Teaching in Higher Education* 5, no. 4 (2000): 419-434, <https://doi.org/10.1080/713699176>.

² Francisco Zabaleta, "The use and misuse of student evaluation of teaching," *Teaching in Higher Education* 12, no. 1 (2007): 55-76, <https://doi.org/10.1080/13562510601102131>.

³ Cecilia K. Y. Chan, Lillian Y. Y. Luk, and Min Zeng, "Teachers' perceptions of student evaluations of teaching," *Educational Research and Evaluation* 20, no. 4 (2014): 275-289, <https://doi.org/10.1080/13803611.2014.932698>.

⁴ Tanya N. Beran and Jeniffer L. Rokosh, "Instructor's perspectives on the utility of student ratings of instruction," *Instructional Science* 37, (2009): 171-184, <https://doi.org/10.1007/s11251-007-9045-2>.

⁵ Paul W. G. Surgenor, "Obstacles and opportunities: Addressing the growing pains of summative student evaluation of teaching," *Assessment & Evaluation in Higher Education* 38, (2013): 363-376, <https://doi.org/10.1080/02602938.2011.635247>.

⁶ Fadia Nasser and Barbara Fresko, "Faculty views of student evaluation of college teaching," *Assessment & Evaluation in Higher Education* 27, no. 2 (2002): 187-198, <https://doi.org/10.1080/02602930220128751>.

key factor within institutional audits), with many teachers being highly suspicious and often hostile towards the use of SET results as a critical factor for administrative decision-making,^{7,8,9} but agreeing that SET ratings provide instructors with valuable information on how to refine their teaching based on how their students have perceived their teaching practice during the course.^{10,11,12,13}

SET surveys try to assess instructors' teaching effectiveness or teaching quality of a particular course surveying students' opinion, usually, through Likert-scale questionnaires (in this regard and despite not being the focus of the present study, see Bedggood and Donovan,¹⁴ for an overview of published research criticising whether students' ratings constitute a measure of students' satisfaction as consumers rather than a measure of teaching quality). The standardised instruments that are most widespread in the literature on teacher evaluation in higher education, and therefore suggest greater prominence in the international university environment, are the Student Evaluation of Educational Quality (SEEQ), the questionnaire for student evaluation of teaching SET37 and the Students' Evaluation of Teaching Effectiveness Rating Scale (SETERS) (see Moreno-Murcia, Silveira and Belando,¹⁵ for a brief overview on SET instruments used by

⁷ Senior Bolivar, "Student teaching evaluations: Options and concerns," *Journal of Construction Education* 5, no. 1 (2000): 20-29.

⁸ Cecilia. K. Y. Chan, Lillian Y. Y. Luk, and Min Zeng, "Teachers' perceptions of student evaluations of teaching," *Educational Research and Evaluation* 20, no. 4 (2014): 275-289, <https://doi.org/10.1080/13803611.2014.932698>.

⁹ John C. Ory, "Faculty thoughts and concerns about student ratings," *New Directions for Teaching and Learning* 87, (2001): 3-15, <http://dx.doi.org/10.1002/tl.23>.

¹⁰ John C. Ory, "Faculty thoughts and concerns about student ratings," *New Directions for Teaching and Learning* 87, (2001): 3-15, <http://dx.doi.org/10.1002/tl.23>.

¹¹ Magdalena Cladera, "Let's ask our students what really matters to them," *Journal of Applied Research in Higher Education* 13, no. 1 (2021): 112-125, <https://doi.org/10.1108/JARHE-07-2019-0195>.

¹² James A. Kulik, "Student ratings: Validity, utility and controversy," *New Directions for Institutional Research* 27, (2002): 9-25, <http://dx.doi.org/10.1002/ir.1>.

¹³ Fadia Nasser and Barbara Fresko, "Faculty views of student evaluation of college teaching," *Assessment & Evaluation in Higher Education* 27, no. 2 (2002): 187-198, <https://doi.org/10.1080/02602930220128751>.

¹⁴ Rowan E. Bedggood and Jerome D. Donovan, "University performance evaluations: what are we really measuring?" *Studies in Higher Education* 37, no. 7 (2012): 825-842, <https://doi.org/10.1080/03075079.2010.549221>.

¹⁵ Juan Antonio Moreno-Murcia, Yolanda Silveira, and Noelia Belando, "Questionnaire Evaluating Teaching Competencies in the University Environment. Evaluation of Teaching Competencies in the University," *Journal of New Approaches in Educational Research* 4, no. 1, (2015): 54-61, <https://doi.org/10.7821/naer.2015.1.106>.

universities worldwide). Other institutions do not employ standardised questionnaires but develop their SET instruments. Therefore, the current practice of SET surveys is littered with instruments that differ on the items they incorporate and in the particular dimensions they capture to try to adequately operationalise the teaching effectiveness construct.¹⁶ In this sense, there is general agreement that SET instruments must capture the multidimensional structure of the teaching process and, therefore, reflect this multidimensionality incorporating several dimensions related to effective teaching.¹⁷ However, the SET literature reflects a wide variety both in the nature and number of the dimensions that are measured in SET instruments (see Spooren, Brockx and Mortelmans,¹⁸ for an overview of the dimensions that are captured in recently published literature on SET instruments). In this regard, a relatively recent study by Bedggood and Donovan¹⁹ (p. 831) identified “quality of instruction” (i.e., referring to “both teachers skills and ability, and also to their friendliness, enthusiasm and approachability”), “task difficulty” (i.e., “in terms of demands and effort required by students to achieve their desired result”), and “academic development and stimulation” (i.e., regarding to “how stimulated and motivated a student feels, and whether they believe they are growing and developing their academic skills”) as the three most commonly identified dimensions in SET surveys.

SET research has been a hot topic for international academic researchers for a long time. It indeed still is, primarily because of the number of concerns involved in the use of formal instruments for obtaining students' feedback in higher education and their consequences.²⁰ A large body of literature has

¹⁶ Lawrence A. Roche and Herbert W. Marsh, “Multiple dimensions of university teacher self-concept,” *Instructional Science* 28, (2000): 439-468, <http://dx.doi.org/10.1023/A:1026576404113>.

¹⁷ Herbert W. Marsh, “Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness,” in *The scholarship of teaching and learning in higher education: An evidence-based perspective*, ed. R. P. Perry and J. C. Smart (New York: Springer, 2007a), 319-383.

¹⁸ Pieter Spooren, Bert Brockx, and Dimitri Mortelmans, “On the validity of student evaluation of teaching: The state of the art,” *Review of Educational Research* 83, no. 4 (2013): 598-642, <https://doi.org/10.3102/0034654313496870>.

¹⁹ Rowan E. Bedggood and Jerome D. Donovan, “University performance evaluations: what are we really measuring?” *Studies in Higher Education* 37, no. 7 (2012): 825-842, <https://doi.org/10.1080/03075079.2010.549221>.

²⁰ Dennis E. Clayton, “Student evaluations of teaching: are they related to what students learn? A meta-analysis and review of the literature,” *Journal of Marketing Education* 31, no. 1 (2009): 16-30, <https://doi.org/10.1177/0273475308324086>.

addressed the reliability, stability, and validity of the questionnaires in search of more valid and reliable SET instruments that will, therefore, help increase the trust in SET results.²¹ Likewise, published SET literature has largely examined the possible influence of potential biasing factors in student evaluations (e.g., gender, race, ethnicity, culture) and how they may affect SET results interpretation.^{22,23} Other studies compared SET results when evaluations were collected in-class with those gathered using online methods,^{24,25} others have used student ratings to benchmark universities,²⁶ while others have attempted to identify “motivators, barriers, and strategies to improve response rate to student evaluation of teaching”.²⁷

Despite this richness of teaching evaluation literature, the majority of studies either rely on case studies or small cross-sectional data sample sizes obtained from one academic year.²⁸ In this regard, considering a longitudinal framework is crucial to investigate teachers' performance across years.²⁹ Despite the existence of several SET studies that have used longitudinal data

²¹ Jing Zhao and Dorinda J. Gallant, “Student evaluation of instruction in higher education: Exploring issues of validity and reliability,” *Assessment & Evaluation in Higher Education* 37, no. 2 (2012): 227-235, <https://doi.org/10.1080/02602938.2010.523819>.

²² Yanan Fan, L. J. Shepherd, E. Slavich, D. Waters, M. Stone, M., R. Abel, and E. L. Johnston, “Gender and cultural bias in student evaluations: Why representation matters,” *PLoS ONE* 14, no. 2 (2019): e0209749, <https://doi.org/10.1371/journal.pone.0209749>.

²³ Thi Thu Trang Tran and Truong Xuan Do, “Student evaluation of teaching: do teacher age, seniority, gender, and qualification matter?” *Educational Studies*, (2020). <https://doi.org/10.1080/03055698.2020.1771545>.

²⁴ Rosemary J. Avery, Keith Bryant, Alan Mathios, Hyojin Kang, and Duncan Bell, “Electronic course evaluations: does an online delivery system influence student evaluations?” *The Journal of Economic Education* 36, no. 1 (2006): 21-37, <https://doi.org/10.3200/JECE.37.1.21-37>.

²⁵ Yesim Capa-Aydin, “Student evaluation of instruction: comparison between in-class and online methods,” *Assessment & Evaluation in Higher Education* 41, no. 1 (2016): 112-126, <https://doi.org/10.1080/02602938.2014.987106>.

²⁶ Cheng and Marsh, “UK National Student Survey: Are differences between universities and courses reliable and meaningful,” *Oxford Review of Education* 36, no. 6 (2010): 693-712, <<https://doi.org/10.1080/03054985.2010.491179>>.

²⁷ Catherine Cone, Velliyur Viswesh, Vasudha Gupta, and Elizabeth Unni, “Motivators, barriers, and strategies to improve response rate to student evaluation of teaching,” *Currents in Pharmacy Teaching and Learning* 10, no. 12 (2018): 1543-1549, <https://doi.org/10.1016/j.cptl.2018.08.020>.

²⁸ Yanan Fan, L. J. Shepherd, E. Slavich, D. Waters, M. Stone, M., R. Abel, and E. L. Johnston, “Gender and cultural bias in student evaluations: Why representation matters,” *PLoS ONE* 14, no. 2 (2019): e0209749, <https://doi.org/10.1371/journal.pone.0209749>.

²⁹ Silvia Bacci, “Longitudinal data: Different approaches in the context of item-response theory models,” *Journal of Applied Statistics* 39, no. 9 (2012): 2047-2065, <https://doi.org/10.1080/02664763.2012.700451>.

to assess teachers' ratings over the years (e.g.,^{30,31,32}), there is still a need for more long-term longitudinal studies to track and analyse the ratings of the same cohort of teachers over extended periods.³³ Specifically, the year-by-year analysis of the ratings obtained by a teacher in a particular subject might provide teachers with useful insights on how their teaching performance is perceived by their students with the close experience of having taught the subject for another year, helping them to identify their strengths or weaknesses in their way of teaching in the short term,^{34,35} thus allowing lecturers to prepare the subject in a better way.³⁶ In particular, focusing more specifically on those teaching components in which instructors' ratings significantly increase from year to year might contribute to a better understanding of the path to teaching excellence (see Jones,³⁷ for an overview of how to measure the quality of higher education when linked to teaching quality measures).

However, and to our best knowledge, no studies have specifically analysed the year-by-year behaviour of those SET items and dimensions in

³⁰ Herbert W. Marsh and Lawrence A. Roche, "Effects of grading leniency and low workload on students' evaluations of teaching: Popular myth, bias, validity, or innocent bystanders?" *Journal of Educational Psychology* 92, no. 1 (2000): 202-228, <https://doi.org/10.1037/0022-0663.92.1.202>.

³¹ Herbert W. Marsh, "Do university teachers become more effective with experience? A multilevel growth model of students' evaluations of teaching over 13 years," *Journal of Educational Psychology* 99, no. 4 (2007b): 775-790, <https://doi.org/10.1037/0022-0663.99.4.775>.

³² Isabella Sulis, Mariano Porcu, and Vincenza Capursi, "On the Use of Student Evaluation of Teaching: A Longitudinal Analysis Combining Measurement Issues and Implications of the Exercise," *Social Indicators Research* 142, (2019): 1305-1331, <https://doi.org/10.1007/s11205-018-1946-8>.

³³ Herbert W. Marsh, "Do university teachers become more effective with experience? A multilevel growth model of students' evaluations of teaching over 13 years," *Journal of Educational Psychology* 99, no. 4 (2007b): 775-790, <https://doi.org/10.1037/0022-0663.99.4.775>.

³⁴ Sarah. J. Stein, Dorothy Spiller, Stuart Terry, Trudy Harris, Lynley Deaker, and Jo Kennedy, *Unlocking the impact of tertiary teachers' perceptions of student evaluation of teaching Wellington*, New Zealand: Ako Aotearoa National Centre for Tertiary Teaching Excellence, 2012.

³⁵ Paul W. G. Surgeon, "Obstacles and opportunities: Addressing the growing pains of summative student evaluation of teaching," *Assessment & Evaluation in Higher Education* 38, (2013): 363-376, <https://doi.org/10.1080/02602938.2011.635247>.

³⁶ Magdalena Cladera, "Let's ask our students what really matters to them," *Journal of Applied Research in Higher Education* 13, no. 1 (2021): 112-125, <https://doi.org/10.1108/JARHE-07-2019-0195>.

³⁷ Sandra Jones, "Measuring the quality of higher education: Linking teaching quality measures at the delivery level to administrative measures at the university level," *Quality in Higher Education* 9, no. 3 (2003): 223-229, <https://doi.org/10.1080/1353832032000151094>.

which teachers' ratings increase significantly. Therefore, this exploratory study aimed to increase knowledge on the SET topic by examining major increases in teachers' performance evaluations and its immediate impact on next year's score for those instructors that taught the same course or subject for at least two years in a row. Specifically, this paper had a twofold aim. Firstly, to identify those SET survey items associated with major increases in teacher evaluations of a particular subject. Secondly, to examine if there is evidence of the use of these SET results by instructors to improve their teaching when analysing the SET ratings behaviour in the years before and after the major increase occurred.

II. Method

II.1. Sample

The SET surveys of a public university (Madrid, Spain) over five consecutive years were analysed. Teachers were evaluated each year in all subjects and groups taught. For the study, the evaluations of all the groups corresponding to the same teacher, subject and year were grouped. A total number of 13,052 teacher evaluations was obtained from the 21 departments of all the different faculties. The average number of questionnaires collected per subject evaluated was 40.8. The pairs of evaluations corresponding to the same teacher and subject in two consecutive years were then selected for a total of 3,893 pairs of evaluations under analysis.

II.2. Instrument

The instrument for collecting students' evaluations was the standard used by the university. The researchers did not participate in any way in its development. The questionnaire consisted of ten questions grouped into three dimensions defined a priori: planning and organisation of the subject, teacher obligations and teaching methodology. In addition to the last question in which the students' overall satisfaction with the teacher is collected. All questions were formulated in terms of the degree of agreement of the students with different statements about various aspects of teaching. Responses were given on a 5-point Likert scale that ranged from "Strongly disagree" (1 on the scale) to "Strongly agree" (5 on the scale). Table 1 presents the dimensions captured by the instrument and the items included. It also includes a series of identification codes for each dimension and for each item that are used to more easily display the results of the analysis.

Table 1
Survey items and dimensions captured
by the instrument for collecting students' feedback

Dimension	Survey items	Id code
<u>Planning and organisation (PO)</u>		
	The teacher explains in detail to the students the teaching guide of the subject at the beginning of the course	PO1
	The teacher has informed clearly about the assessment criteria of the subject	PO2
	The teacher, in addition to the face-to-face classes, has planned complementary activities (e.g., problem-solving, readings, practical exercises) that facilitate the learning of the subject	PO3
<u>Teacher obligations (TO)</u>		
	The teacher respects the class schedules	TO1
	The teacher is available to attend to the students	TO2
	Teaching activities to meet the objectives, contents and methodology specified in the teaching guide of the subject	TO3
<u>Teaching methodology (TM)</u>		
	The teacher adequately clarifies the doubts of the different activities proposed in the subject	TM1
	The teacher explains clearly	TM2
	The development of the subject allows me adequate monitoring and learning	TM3
<u>Overall satisfaction (OS)</u>		
	Taking into account all the aspects mentioned, I am satisfied with the work carried out by the teacher	OS

II.3. Procedure

The procedure for obtaining student feedback was in a face-to-face classroom setting. The evaluations of the subjects of a single term duration were carried out at the end of each term immediately prior to the start of the examination period. The evaluations of the subjects of annual duration were

collected at the end of the second term. A survey team provided a questionnaire to each student containing the instructions and the survey questions along with optical reading sheets in which the students fill in the answers. No data were requested to allow the identification of the students. An optical mark reader software was used to scan the answers automatically (Dara Optical Mark Reader, Dara Group, Spain). The research complied both with the ethical principles of research of the university where the research was conducted and with the Ethical Guidelines for Educational Research published by the British Educational Research Association (BERA, fourth edition, 2018).³⁸ Collected data did not allow the identification of the teacher or the students nor did it require consent because data had already been collected for administrative purposes by the University and no intervention was conducted, thus guaranteeing compliance with internationally recognized scientific legislation and protocols advocating for the generation, dissemination, and application of research results for the scientific, technical and cultural development of society.³⁹

II.4. Statistical analysis

The study aimed to analyse major increases in teachers' performance evaluations corresponding to the same teacher and subject in two consecutive years. In the first phase of the analysis, the pairs of evaluations susceptible to comparison were selected by calculating the difference in scores for each item and the average score thereof. In a second phase, the pairs of evaluations with an increase in the score were classified, both in average terms and for each item. The 95% percentile of the differences between the ratings obtained by the teachers in the subjects in two consecutive years was the cut-off point for the consideration of major increases.

To analyse the association of the survey items with major increases, these were classified into two types: average and isolated. Those evaluations that experienced a major average increase in the overall rating (i.e. when considering all the survey items) were classified as major average increases. Those evaluations in which there was a major increase in at least one of the items but did not produce a major average increase were identified as isolated major increases. To determine the relationships between major increases

³⁸ British Educational Research Association (BERA), *Ethical Guidelines for Educational Research* (London: BERA, 2018).

³⁹ UK Data Service, "Consent for data sharing," 2022, <https://ukdataservice.ac.uk/learning-hub/research-data-management/ethical-issues/consent-for-data-sharing/>.

(average and isolated) and the items in the questionnaire, the frequencies of occurrence were obtained. For examining the evidence of the use of SET results by instructors, a single mean value for each assessment in the years before and after the major increase was considered. The Student's t-test of the difference between average scores of paired data was computed. Special attention was paid to the behaviour of the scores in the year after the major increase in relation to the two previous years. That case-by-case comparison was impossible when the teacher did not teach the same subject the year after the major increase. Therefore, in these cases, the unpaired t-test was used. The statistical package SPSS (v21.0, IBM Corporation, USA) was used for analysis.

III. Results

III.1. Association of the survey items with major increases

The values from which 50%, 75%, 90%, 95% and 99% (percentiles) of the differences in the evaluations of the teachers and subject ratings in consecutive years are presented in Table 2. The values are presented for each questionnaire item, for the set of ten items and the mean score of all the items. The cut-off points for selecting the major increases were 0.855 for the set of all questions and 0.975 for the mean values. A first approximation shows TO1, TM2, TM3 and OS as the items in which the greatest differences occur. The 2,107 questions that exceeded these limits referred to 567 different pairs of teacher and subject ratings, of which 194 were significant average increases.

Table 2
Percentiles of differences in the teacher ratings in consecutive years

Survey item	P50%	P75%	P90%	P95%	P99%
PO1	0.06	0.38	0.73	0.96	1.5
PO2	0.06	0.37	0.72	0.97	1.46
PO3	0.04	0.39	0.74	0.99	1.5
TO1	0.03	0.36	0.71	1.01	1.69
TO2	0.02	0.32	0.65	0.88	1.46
TO3	-0.01	0.28	0.6	0.80	1.40
TM1	0.03	0.35	0.71	0.95	1.53

Survey item	P50%	P75%	P90%	P95%	P99%
TM2	0.03	0.37	0.75	1.00	1.54
TM3	0.07	0.40	0.75	1.03	1.53
OS	0.02	0.37	0.75	1.06	1.65
Set of ten items	0.03	0.32	0.63	0.86	1.4
Mean of all items	0.03	0.36	0.71	0.97	1.53

The percentage of items in which there was a major increase is presented in Table 3. Regarding isolated major increases, the items that appeared most frequently were PO3 (i.e., the teacher plans complementary activities that facilitate the learning of the subject) and TO1 (i.e., compliance with class schedules), with 25% and 27% respectively. Conversely, the items that showed the lowest frequencies were TO3 (i.e., teaching activities to meet the specifications of the teaching guide of the subject), TO2 (i.e., the teacher is available to attend to the students) and TM1 (i.e., the teacher clarifies the doubts properly), with 4%, 9% and 10% respectively. A more in-depth analysis showed that isolated major increases typically occurred in one (61.7%) or two (22.3%) of the questionnaire items, with items belonging to the dimensions of planning and organisation (PO) and teacher obligations (TO) showing the highest frequency of occurrence. In particular, items PO3 and TO1 accumulated the highest frequency when major increases occurred specifically in one item (reaching figures of 37.0% and 35.2% respectively). Interestingly and conversely, items PO3 and TO1 were among the three items that appeared less frequently among major average increases (57% and 56% respectively). Items relating to teaching methodology (TM dimension) showed the lower frequencies among isolated major increases.

With respect to the major average increases, all items showed frequencies of occurrence above 50%. When a major average increase occurred, major increases occurred in at least four survey items (6.8 items on average). Specifically, in 19% of cases, major increases occurred in all the items of the questionnaire. The three items that appeared most frequently among major average increases were those related to teaching methodology (i.e., TM1, TM2, TM3), all above 70%. Major average increases were strongly associated with the item that retrieved the students' overall satisfaction (OS item). In 87% of the cases in which there was a major average increase, there was also a major increase in this item. Interestingly, the item related to the teacher's compliance with the activity specifications as stipulated in the teaching guide

(TO3 item) showed the lowest frequency of occurrence both for isolated major increases (4%) and major average increases (55%).

Table 3

Percentage of items in which there was a major increase by type

Survey item	Isolated major increases	Major average increases	All increases
PO1	17%	64%	33%
PO2	16%	68%	34%
PO3	25%	57%	36%
TO1	27%	56%	37%
TO2	9%	60%	26%
TO3	4%	55%	21%
TM1	10%	75%	32%
TM2	18%	74%	37%
TM3	16%	83%	39%
OS	19%	87%	42%

III.2. Evidence of the use of SET results by teachers

The case-by-case comparison in the year after the major increase yielded the following results. In 36% of the cases the score obtained in the year after the major increase continued to rise, with teachers obtaining even better ratings (mean score differences greater than zero; mean difference = 0,46, standard deviation SD = 0,40, paired t-test p-value < 0.0001). In the rest of the cases (64%), the scores obtained the year after the major increase went down, showing mean score differences significantly lower than zero (mean difference = -0.50, SD = 0.37, paired t-test p-value < 0.0001) but significantly above the scores obtained in the course prior to the major increase (mean difference = 0.82, SD = 0.46, paired t-test p-value < 0.0001). Overall, the year after the major increase almost all teachers (95.7%) obtained an average score higher than that of the year prior to it.

Regarding the comparison between the year after the major increase and the set of all scores, the results showed that the average of the scores was significantly lower than that of the set of all scores (mean difference = -0.23, SD = 0.67, independent samples t-test p-value < 0.0001). That is, the highest

scores are obtained the year of the major increase. Besides, the scores obtained the year before showed lower values than those obtained in the subsequent year. These scores are also lower than those of the set of evaluations in which no major increases were detected.

IV. Discussion

To the best of our knowledge, this was the first study to specifically examine major increases in teachers' performance evaluations and their immediate impact on next year's score based on evidence from SET surveys.

The results of the analysed SET instrument highlighted the three items captured by the dimension of teaching methodology as those most closely tied to major average increases. Even though there is no single definition of its scope, teaching methodology is understood as the "set of strategies, procedures and actions consciously and thoughtfully organised and planned by teachers to guarantee student learning and the attainment of the stated objectives".⁴⁰ It is precisely this conscious action by teachers when implementing the most appropriate strategies, procedures and actions that the literature has highlighted about the importance of SET scores to promote teachers' self-reflection on their teaching quality and ability for teaching improvement purposes.^{41,42,43}

These findings suggest the possible existence of a strong relationship between the extent to which students rate those SET survey aspects regarding teaching methodology and their degree of satisfaction with the teacher. Accordingly, focusing specifically on those teaching components in which instructors' ratings significantly increase from year to year, beyond the specific scores obtained, could contribute to deciphering the pathway towards excellence in teaching and learning, bringing with it the consequent benefit for students, teachers, institutions and society as a whole. Very

⁴⁰ Eurydice, "Teaching and learning in Primary Education. European Commission, Education Information Network in Europe," 2020, https://eacea.ec.europa.eu/national-policies/eurydice/content/teaching-and-learning-primary-education-42_en.

⁴¹ Cecilia. K. Y. Chan, Lillian Y. Y. Luk, and Min Zeng, "Teachers' perceptions of student evaluations of teaching," *Educational Research and Evaluation* 20, no. 4 (2014): 275-289, <https://doi.org/10.1080/13803611.2014.932698>.

⁴² Beatrice Tucker, Sue Jones, Leon Straker, and Joan Cole, "Course evaluation on the web: Facilitating student and teacher reflection to improve learning," *New Directions for Teaching and Learning* 96, (2003): 81-94, <http://dx.doi.org/10.1002/tl.125>.

⁴³ Yuankun Yao and Marilyn Grady, "How do faculty make formative use of student evaluation feedback? A multiple case study," *Journal of Personnel Evaluation in Education* 18, (2005): 107-126, <https://doi.org/10.1007/s11092-006-9000-9>.

interestingly, Beran and Rokosh⁴⁴ found that SET scores were used to a lesser extent to make choices about course textbooks, exams, and student assignments. Three particular aspects that are more related to the design of the subject or course rather than to how the subject is taught. These findings would be in line with the results of the present study and would suggest that the aspects related to the dimension of planification and organisation of the subject are less important in the general perception of student satisfaction than the specific actions carried out by the teacher (i.e. teaching methodology).

Regarding the isolated increases and together with the above, the results showed that it was less frequent to find isolated major increases in the items relative to the dimension of teaching methodology. This ratifies the idea that when there is a major increase in the items related to teaching methodology, the overall satisfaction of the student with the teacher increases, causing a drag effect on the rest of the items. Isolated major increases were mainly associated with items related to the dimensions of teacher obligations (TO) and planning and organisation (PO). In particular, items TO3 (i.e., teaching activities to meet the specifications of the teaching guide of the subject), TO2 (i.e., the teacher is available to attend to the students) and TM1 (i.e., the teacher clarifies the doubts properly) rarely appeared alone in isolated increases, nor did they show a strong association with major average increases. This could be an indicator that the scores of some of the different aspects of teaching are being differentiated and might refer to aspects of teaching that are less important for the students' general satisfaction.

Regarding the analysis of the scores when examining the previous year and the year following a major increase, the results showed that major increases generally started from scores below the set of evaluations in which no major increases were detected in the previous year. However, these increases were not consolidated in the year following the major increase. Most of the teachers lowered their scores the year after a major increase occurred, both for the teachers who improved or lowered their ratings. Although this is not incompatible with a possible reaction of the teaching staff to the feedback of their students for improving the teaching of a particular subject in the following academic year, this could also be due to natural variations in SET scores (that is, major increases would result from the combination of a low score in one year and a high score in the following

⁴⁴ Tanya N. Beran and Jeniffer L. Rokosh, "Instructor's perspectives on the utility of student ratings of instruction," *Instructional Science* 37, (2009): 171-184, <https://doi.org/10.1007/s11251-007-9045-2>.

year, without any of these scores being exceptional). Previous research analysing teachers' attitudes and reactions toward SET results has found conflicting results, covering the entire range from total acceptance to the strongest opposition.^{45,46}

On the one hand, research has found instructors who recognise the importance of SETs and who consider that the systematic feedback they receive year on year from their students constitutes a very valuable and useful tool for the improvement of teaching (e.g.,^{47,48}) and thus better learning for students.^{49,50} Specifically, in a study investigating teachers' attitudes about SET ratings from a sample of 357 teachers conducted at a Canadian university by Beran and Rokosh,⁵¹ the authors found that the teachers considered SET results to be most useful "for improving general teaching quality, for refining overall instruction, and for improving lectures". In this particular sense and very interestingly, a longitudinal study on the impact of lecturers reflective practices as an essential aspect of professional development found that SET scores increased for all reflective teachers year after year and, more significantly for instructors who showed higher levels of reflection.⁵² In light of the results of the present research, this would suggest that the premise on the required level of reflection was

⁴⁵ Nasser, Fadia, and Barbara Fresko. "Faculty views of student evaluation of college teaching." *Assessment & Evaluation in Higher Education*, 27, no. 2 (2002): 187-198. <https://doi.org/10.1080/02602930220128751>

⁴⁶ Arthur, Linet. "From performativity to professionalism: Lecturer's responses to student feedback." *Teaching in Higher Education*, 14, (2009): 441-454. <https://doi.org/10.1080/13562510903050228>

⁴⁷ Nasser, Fadia, and Barbara Fresko. "Faculty views of student evaluation of college teaching." *Assessment & Evaluation in Higher Education*, 27, no. 2 (2002): 187-198. <https://doi.org/10.1080/02602930220128751>

⁴⁸ Surgenor, P. W. G. "Obstacles and opportunities: Addressing the growing pains of summative student evaluation of teaching." *Assessment & Evaluation in Higher Education*, 38, (2013): 363-376. <https://doi.org/10.1080/02602938.2011.635247>

⁴⁹ Ballantyne, Roy, Jill Borthwick, and Jan Packer. "Beyond student evaluation of teaching: Identifying and addressing academic staff development needs." *Assessment & Evaluation in Higher Education*, 25, no. 3 (2000): 221-236. <https://doi.org/10.1080/713611430>

⁵⁰ Zhao, Jing, and Dorinda J. Gallant. "Student evaluation of instruction in higher education: Exploring issues of validity and reliability." *Assessment & Evaluation in Higher Education*, 37, no. 2 (2012): 227-235. <https://doi.org/10.1080/02602938.2010.523819>

⁵¹ Tanya N. Beran and Jeniffer L. Rokosh, "Instructor's perspectives on the utility of student ratings of instruction." *Instructional Science* 37, (2009): 171-184, <https://doi.org/10.1007/s11251-007-9045-2>.

⁵² Tiffany M. Winchester and Maxwell Winchester, "A longitudinal investigation of the impact of faculty reflective practices on students' evaluations of teaching." *British Journal of Educational Technology* 45, no. 1 (2014): 11-124, <https://doi.org/10.1111/bjet.12019>.

not fulfilled. However, more research would be needed on this regard (e.g., qualitative research to investigate the possible relations between instructors' beliefs, thoughts or beliefs about their experiences with SET major increases).

On the other hand, teachers that show negative attitudes towards the use of SETs for the improvement of their teaching mainly argue that the aspects covered in the evaluations do not reflect their perceptions of good teaching (e.g.,^{53,54,55}), therefore making them consider SET results of low or null utility for refining instruction and thus making little or no use of student feedback.⁵⁶ Nevertheless, and to a certain extent paradoxically, some of these same studies^{57,58,59} have observed a general recognition by teachers of the suitability and usefulness of SET surveys for other purposes such as administrative decision-making or institutional integrity assessment. In this regard, previous research argued that teachers' response to students' feedback is a complex process involving multiple factors (e.g., teachers' background and experience, teacher's personality, students' characteristics, teaching strategies used) that is influenced by instructors' perceptions, beliefs and feelings⁶⁰ and that, ultimately, is more related to the teachers' desire and ability for change than to the

⁵³ Peter Burden, "Does the end of semester evaluation forms represent teacher's views of teaching in a tertiary education context in Japan?" *Teaching and Teacher Education* 24, no. 6 (2008): 1463-1475, <https://doi.org/10.1016/j.tate.2007.11.012>.

⁵⁴ Peter Burden, "Creating confusion or creative evaluation? The use of student evaluation of teaching surveys in Japanese tertiary education," *Educational Assessment, Evaluation and Accountability* 22, (2010): 97-117, <https://doi.org/10.1007/s11092-010-9093-z>.

⁵⁵ Paul W. G. Surgenor, "Obstacles and opportunities: Addressing the growing pains of summative student evaluation of teaching," *Assessment & Evaluation in Higher Education* 38, (2013): 363-376, <https://doi.org/10.1080/02602938.2011.635247>.

⁵⁶ Fadia Nasser and Barbara Fresko, "Faculty views of student evaluation of college teaching," *Assessment & Evaluation in Higher Education* 27, no. 2 (2002): 187-198, <https://doi.org/10.1080/02602930220128751>.

⁵⁷ Peter Burden, "Does the end of semester evaluation forms represent teacher's views of teaching in a tertiary education context in Japan?" *Teaching and Teacher Education* 24, no. 6 (2008): 1463-1475, <https://doi.org/10.1016/j.tate.2007.11.012>.

⁵⁸ Peter Burden, "Creating confusion or creative evaluation? The use of student evaluation of teaching surveys in Japanese tertiary education," *Educational Assessment, Evaluation and Accountability* 22, (2010): 97-117, <https://doi.org/10.1007/s11092-010-9093-z>.

⁵⁹ Fadia Nasser and Barbara Fresko, "Faculty views of student evaluation of college teaching," *Assessment & Evaluation in Higher Education* 27, no. 2 (2002): 187-198, <https://doi.org/10.1080/02602930220128751>.

⁶⁰ Arthur Linet, "From performativity to professionalism: Lecturer's responses to student feedback," *Teaching in Higher Education* 14, (2009): 441-454, <https://doi.org/10.1080/13562510903050228>.

belief in the usefulness of SETs.⁶¹ In this sense and very interestingly, Hendry, Lyon and Henderson-Smart⁶² observed from a survey of 121 lecturers covering student feedback over two years that those teachers implementing the “conceptual-change student-focused (CCSF) approach” in their classes (i.e., the CCSF approach was described by Prosser and Trigwell⁶³ in 1999 as a teaching approach in which teachers see students as active builders of their knowledge, being their role as teachers to help them in them in the process) were more responsive to the use of student feedback as guidance to improve their teaching. Given the large number of teachers considered in the present study and the diversity of subjects taught, it is expected to find all kinds of teachers' profiles and attitudes so that the explanations of the results obtained from the teacher evaluations could respond to different reasons. However, it seems that they do not respond to the expected patterns that would allow us to assert that major increases are mainly due to the reaction from teachers to SET ratings.

Overall, the results of the study contributed to a better understanding of the behaviour of major increases in SET ratings for those teachers that taught the same course or subject for at least two years in a row. The main results highlighted the three SET survey items captured by the dimension of teaching methodology as those most closely related to major increases in teacher evaluations and that there is no generalised response from teachers who experience major increases in SET ratings. However, and as is common in studies based on SET surveys, care must be taken when interpreting and extrapolating the results to other university educational contexts. Accordingly, as in all research, the findings of the present study should be interpreted based on a series of limitations. The instrument for collecting students' evaluations was not of those of standardised use in the international context and the items of which it was composed were grouped into three dimensions defined a priori. The study draws its sample from multiple years of a single institution (Spanish university, 5-year period). Also, further explanations for SET score changes, such as university administration's manipulation of

⁶¹ Fadia Nasser and Barbara Fresko, “Faculty views of student evaluation of college teaching,” *Assessment & Evaluation in Higher Education* 27, no. 2 (2002): 187-198, <https://doi.org/10.1080/02602930220128751>.

⁶² Graham D. Hendry, Patricia M. Lyon, and Cheryl Henderson-Smart, “Teachers' approaches to teaching and responses to student evaluation in a problem-based medical program,” *Assessment and Evaluation in Higher Education* 32, no. 2 (2007): 143-157, <https://doi.org/10.1080/02602930600801894>.

⁶³ Michael Prosser and Keith Trigwell, *Understanding Learning and Teaching: The experience in higher education*, SRHE and Open University Press: Buckingham, 1999.

courses arrangement, faculty competition, or tenure-track pressure, cannot be discarded. Thus, future research directions could further explore major score changes in teachers' performance evaluations trying (i.e., not only major increases but major dropdowns) to account for these potential limitations within the perspective of new multidimensional, long-term longitudinal and longitudinal international studies. Particularly and according to the aim of the present study, focusing specifically on those teaching components in which instructors' ratings significantly increase from year to year could contribute to deciphering the pathway towards excellence in teaching and learning, bringing with it the consequent benefit for students, teachers and institutions.

Bibliography

- Arthur, Linet. "From performativity to professionalism: Lecturer's responses to student feedback." *Teaching in Higher Education* 14, (2009): 441-454. <https://doi.org/10.1080/13562510903050228>.
- Avery, Rosemary J., Keith Bryant, Alan Mathios, Hyojin Kang, and Duncan Bell. "Electronic course evaluations: does an online delivery system influence student evaluations?" *The Journal of Economic Education* 36, no. 1 (2006): 21-37. <https://doi.org/10.3200/JECE.37.1.21-37>.
- Bacci, Silvia. "Longitudinal data: Different approaches in the context of item-response theory models." *Journal of Applied Statistics* 39, no. 9 (2012): 2047-2065. <https://doi.org/10.1080/02664763.2012.700451>.
- Ballantyne, Roy, Jill Borthwick, and Jan Packer. "Beyond student evaluation of teaching: Identifying and addressing academic staff development needs." *Assessment & Evaluation in Higher Education* 25, no. 3 (2000): 221-236. <https://doi.org/10.1080/713611430>.
- Bedggood, Rowan E., and Jerome D. Donovan. "University performance evaluations: what are we really measuring?" *Studies in Higher Education* 37, no. 7 (2012): 825-842. <https://doi.org/10.1080/03075079.2010.549221>.
- Beran, Tanya. N., and Jeniffer L. Rokosh. "Instructor's perspectives on the utility of student ratings of instruction". *Instructional Science* 37, (2009): 171-184. <https://doi.org/10.1007/s11251-007-9045-2>.
- Bolivar, Senior. "Student teaching evaluations: Options and concerns." *Journal of Construction Education* 5, no. 1 (2000): 20-29.
- British Educational Research Association (BERA). *Ethical Guidelines for Educational Research*. (London: BERA, 2018).
- Burden, Peter. "Does the end of semester evaluation forms represent teacher's views of teaching in a tertiary education context in Japan?" *Teaching and Teacher Education* 24, no. 6 (2008): 1463-1475. <https://doi.org/10.1016/j.tate.2007.11.012>.
- Burden, Peter. "Creating confusion or creative evaluation? The use of student evaluation of teaching surveys in Japanese tertiary education." *Educational*

- Assessment, Evaluation and Accountability* 22, (2010): 97-117. <https://doi.org/10.1007/s11092-010-9093-z>.
- Capa-Aydin, Yesim. "Student evaluation of instruction: comparison between in-class and online methods." *Assessment & Evaluation in Higher Education* 41, no. 1 (2016): 112-126. <https://doi.org/10.1080/02602938.2014.987106>.
- Chan, Cecilia. K. Y., Lillian Y. Y. Luk, and Min Zeng. "Teachers' perceptions of student evaluations of teaching." *Educational Research and Evaluation* 20, no. 4 (2014): 275-289. <https://doi.org/10.1080/13803611.2014.932698>.
- Cheng, Jacqueline. H. S., and Herbert W. Marsh. "UK National Student Survey: Are differences between universities and courses reliable and meaningful." *Oxford Review of Education* 36, no. 6 (2010): 693-712. <https://doi.org/10.1080/03054985.2010.491179>.
- Cladera, Magdalena. "Let's ask our students what really matters to them." *Journal of Applied Research in Higher Education* 13, no. 1 (2021): 112-125. <https://doi.org/10.1108/JARHE-07-2019-0195>.
- Clayson, Dennis. E. "Student evaluations of teaching: are they related to what students learn? A meta-analysis and review of the literature." *Journal of Marketing Education* 31, no. 1 (2009): 16-30. <https://doi.org/10.1177/0273475308324086>.
- Cone, Catherine., Velliyyur Viswesh, Vasudha Gupta, and Elizabeth Unni. "Motivators, barriers, and strategies to improve response rate to student evaluation of teaching." *Currents in Pharmacy Teaching and Learning* 10, no. 12 (2018): 1543-1549. <https://doi.org/10.1016/j.cptl.2018.08.020>.
- Eurydice. "Teaching and learning in Primary Education. European Commission, Education Information Network in Europe." 2020. https://eacea.ec.europa.eu/national-policies/eurydice/content/teaching-and-learning-primary-education-42_en.
- Fan, Y., L. J. Shepherd, E. Slavich, D. Waters, M. Stone, M., R. Abel, and E. L. Johnston. "Gender and cultural bias in student evaluations: Why representation matters." *PLoS ONE* 14, no. 2 (2019): e0209749. <https://doi.org/10.1371/journal.pone.0209749>.
- Hendry, Graham D., Patricia M. Lyon, and Cheryl Henderson-Smart. "Teachers' approaches to teaching and responses to student evaluation in a problem-based medical program". *Assessment and Evaluation in Higher Education* 32, no. 2 (2007): 143-157. <https://doi.org/10.1080/02602930600801894>.
- Johnson, Rachel. "The authority of the student evaluation questionnaire." *Teaching in Higher Education* 5, no. 4 (2000): 419-434. <https://doi.org/10.1080/713699176>.
- Jones, Sandra. "Measuring the quality of higher education: Linking teaching quality measures at the delivery level to administrative measures at the university level." *Quality in Higher Education* 9, no. 3 (2003): 223-229. <https://doi.org/10.1080/1353832032000151094>.
- Kulik, James A. "Student ratings: Validity, utility and controversy". *New Directions for Institutional Research* 27, (2002): 9-25. <http://dx.doi.org/10.1002/ir.1>.
- Marsh, Herbert W. "Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness." In *The scholarship of*

- teaching and learning in higher education: An evidence-based perspective*, ed. R. P. Perry and J. C. Smart (New York: Springer, 2007a), 319-383.
- Marsh, Herbert W. "Do university teachers become more effective with experience? A multilevel growth model of students' evaluations of teaching over 13 years." *Journal of Educational Psychology* 99, no. 4 (2007b): 775-790. <https://doi.org/10.1037/0022-0663.99.4.775>.
- Marsh, Herbert W., and L. A. Roche. "Effects of grading leniency and low workload on students' evaluations of teaching: Popular myth, bias, validity, or innocent bystanders?" *Journal of Educational Psychology* 92, no. 1 (2000): 202–228. <https://doi.org/10.1037/0022-0663.92.1.202>.
- Moreno-Murcia, Juan Antonio, Yolanda Silveira, and Noelia Belando. "Questionnaire Evaluating Teaching Competencies in the University Environment. Evaluation of Teaching Competencies in the University." *Journal of New Approaches in Educational Research* 4, no. 1, (2015), 54-61. <https://doi.org/10.7821/naer.2015.1.106>.
- Nasser, Fadia, and Barbara Fresko. "Faculty views of student evaluation of college teaching." *Assessment & Evaluation in Higher Education* 27, no. 2 (2002): 187-198. <https://doi.org/10.1080/02602930220128751>.
- Ory, John C. "Faculty thoughts and concerns about student ratings." *New Directions for Teaching and Learning* 87, (2001): 3-15. <http://dx.doi.org/10.1002/tl.23>.
- Roche, Lawrence A., and Herbert W. Marsh. "Multiple dimensions of university teacher self-concept." *Instructional Science* 28, (2000): 439-468. <http://dx.doi.org/10.1023/A:1026576404113>.
- Spooren, Pieter, Bert Brockx, and Dimitri Mortelmans. "On the validity of student evaluation of teaching: The state of the art." *Review of Educational Research* 83, no. 4 (2013): 598-642. <https://doi.org/10.3102/0034654313496870>.
- Stein, Sarah. J., Dorothy Spiller, Stuart Terry, Trudy Harris, Lynley Deaker, and Jo Kennedy. *Unlocking the impact of tertiary teachers' perceptions of student evaluation of teaching* Wellington, New Zealand: Ako Aotearoa National Centre for Tertiary Teaching Excellence, 2012.
- Sulis, Isabella, Mariano Porcu., and Vincenza Capursi. "On the Use of Student Evaluation of Teaching: A Longitudinal Analysis Combining Measurement Issues and Implications of the Exercise." *Social Indicators Research* 142, (2019): 1305-1331. <https://doi.org/10.1007/s11205-018-1946-8>.
- Surgenor, P. W. G. "Obstacles and opportunities: Addressing the growing pains of summative student evaluation of teaching." *Assessment & Evaluation in Higher Education* 38, (2013): 363-376. <https://doi.org/10.1080/02602938.2011.635247>.
- Tran, Thi Thu Trang, and Truong Xuan Do. "Student evaluation of teaching: do teacher age, seniority, gender, and qualification matter?" *Educational Studies*, (2020). <https://doi.org/10.1080/03055698.2020.1771545>.
- Tucker, Beatrice, Sue Jones, Leon Straker, and Joan Cole. "Course evaluation on the web: Facilitating student and teacher reflection to improve learning." *New Directions for Teaching and Learning* 96, (2003): 81-94. <http://dx.doi.org/10.1002/tl.125>.

- Winchester, Tiffany. M., and Maxwell Winchester. "A longitudinal investigation of the impact of faculty reflective practices on students' evaluations of teaching." *British Journal of Educational Technology* 45, no. 1 (2014): 11-124. <https://doi.org/10.1111/bjet.12019>.
- Yao, Yuankun, and Marilyn Grady. "How do faculty make formative use of student evaluation feedback? A multiple case study." *Journal of Personnel Evaluation in Education* 18, (2005): 107-126. <https://doi.org/10.1007/s11092-006-9000-9>.
- Zabaleta, Francisco. "The use and misuse of student evaluation of teaching." *Teaching in Higher Education* 12, no. 1 (2007): 55-76. <https://doi.org/10.1080/13562510601102131>.
- Zhao, Jing, and Dorinda J. Gallant. "Student evaluation of instruction in higher education: Exploring issues of validity and reliability." *Assessment & Evaluation in Higher Education* 37, no. 2 (2012): 227-235. <https://doi.org/10.1080/02602938.2010.523819>.

About the authors

JAIME PRIETO (jaime.prieto@urjc.es) is a lecturer on information and communications technology at the Rey Juan Carlos University (Spain). He holds a PhD by the Technical University of Madrid. His research interests focus on the field of ICT and education, with particular interest to the field of higher education.

ROCÍO GUEDE-CID (rocio.guede@urjc.es) is a lecturer on Mathematics and Statistics at the Rey Juan Carlos University (Spain). She holds a PhD by the Rey Juan Carlos University. Her research interests focus on the field of innovation and assessment in higher education, as well as on the analysis of the technology transfer process.

ANA I. CID-CID (ana.cid@urjc.es) is a senior lecturer on Mathematics and Statistics at the Rey Juan Carlos University (Spain). She holds a PhD in Business and Economics by the Complutense University of Madrid. Her research interests focus on the assessment of both students' and teachers' performance in the higher education context.

SANTIAGO LEGUEY (corresponding author, santiago.leguey@urjc.es), PhD, is a senior lecturer on Mathematics and Statistics at the Rey Juan Carlos University (Spain). His research interests include applied statistics and education, with particular interest to the field of higher education. He is the Chief Manager of the University Center for Applied Social Studies CUESA.