The role of students in the Docentia process after ten years of evaluation: Evaluating the teacher or the subject? That is the question

El papel del alumnado en el Docentia tras diez años de evaluación: ¿Evaluar al docente o la asignatura? Esa es la cuestión

Rosa ISLA-DÍAZ, PhD. Associate Professor. Universidad de La Laguna (*risladia@ull.edu.es*).
Stephany HESS-MEDLER, PhD. Associate Professor. Universidad de La Laguna (*sthess@ull.edu.es*).
Hipólito MARRERO-HERNÁNDEZ, PhD. Professor. Universidad de La Laguna (*hmarrero@ull.edu.es*).

Abstract:

This paper proposes separating the survey of students' satisfaction with their teachers from the survey of their satisfaction with the module, and presents the implementation of this process at the Universidad de La Laguna (ULL). Items linked to teacher performance were differentiated from those referring to the subject in the Docentia-ULL survey. The final teacher survey comprised 12 items. In order to compare it with the original 22-item survey, we compared the results from the overall assessment of teachers in the original survey with a simulation of the result if the reduced survey had been applied in the period 2012-2014 (n = 689), as well as with its actual application in the 2017-2019 rounds (n = 526). We observed an increase in Excellent teaching staff using the teacher survey and, to a lesser extent, Unfavourable teaching staff. We found that the teacher survey displayed a greater variance in overall student satisfaction compared to the subject survey. The results support the usefulness and validity of assessing student satisfaction through two different surveys. The two surveys would be relevant for different institutional assessment processes: the teacher survey for teacher promotion and

Revision accepted: 2021-06-25.
This is the English version of an article originally printed in Spanish in issue 280 of the **revista española de pedagogía**. For this reason, the abbreviation EV has been added to the page numbers. Please, cite this article as follows: Isla-Díaz, R., Hess-Medler, S. y Marrero-Hernández, H. (2021). El papel del alumnado en el Docentia tras diez años de evaluación: ¿evaluar al docente o la asignatura? Esa es la cuestión | *The role of students in the Docentia process after ten years of evaluation: Evaluating the teacher or the subject? That is the question. Revista Española de Pedagogía, 79* (280), 393-411.
https://doi.org/10.22550/REP79-3-2021-08
https://revistadepedagogia.org/ ISSN: 0034-9461 (Print), 2174-0909 (Online)

the subject survey for degree accreditation. Adoption of the teacher survey by the university would have a positive impact on teachers' intrinsic motivation, in particular by satisfying the needs for autonomy and competence by associating them with skills that they are in control of improving.

Keywords: satisfaction, students, survey, university, teacher appraisal.

Resumen:

Este trabajo propone diferenciar la evaluación de la satisfacción del alumnado con el docente de la satisfacción con la asignatura, y presenta su implementación en la Universidad de La Laguna. Se diferenciaron los ítems vinculados al desempeño del docente de los referidos a la asignatura en la encuesta Docentia-ULL. La encuesta del docente quedó finalmente compuesta por 12 ítems. Para compararla con la encuesta original, de 22 ítems, contrastamos los resultados de la evaluación del profesorado en la encuesta original con una simulación del resultado si se aplicara la encuesta reducida en su lugar en el periodo 2012-2014 (n = 689), así como con su aplicación real en las convocatorias de 2017-2019 (n = 526). Observamos un incremento del profesorado Excelente empleando la encuesta enfocada al docente y, en menor medida, del profesorado Desfavorable. Encontramos que la encuesta del docente explicaba una varianza mayor de la satisfacción general del alumnado en comparación con la encuesta de la asignatura. Los resultados apoyan la utilidad y validez de la evaluación de la satisfacción de alumnado mediante dos encuestas distintas, relevantes para distintos procesos de evaluación institucional: la del docente para la promoción del profesorado y la de la asignatura para la acreditación de los títulos. La incorporación de la encuesta del docente por parte de la universidad tendría un impacto positivo en la motivación intrínseca del profesorado, en particular en la satisfacción de las necesidades de autonomía y competencia, al asociarse con habilidades cuya mejora está «bajo su control».

Descriptores: satisfacción, estudiante, encuesta, universidad, evaluación del profesor.

1. Introduction

Improving the evaluation of teaching practice in Spanish universities is necessary when the quality criteria of the European Higher Education Area (EHEA) are adopted. In this context, Docentia, the instrument Spanish universities have used to do this, is a valuable tool that should enable teachers to evaluate the strengths and weaknesses of their teaching practice (Calderón & Escalera, 2008; Isla-Díaz et al., 2018). Teacher evaluation should be a stimulus for the adoption of good practices (Pozo et al., 2011) and for encouraging the adaptation of their competences to the objectives of the European Higher Education Area (Álvarez Rojo et al., 2009; Benito & Cruz, 2005; High Group on Modernization of Higher Education, 2013; Mayor, 2009; Murillo, 2008; Perales et al., 2014). At the same time, it must be accurate and have consequences

(Alfageme & Caballero, 2010; Tejedor & García-Valcárcel, 2010).

Docentia-ULL is used at the Universidad de La Laguna (ULL) to evaluate teachers by triangulating three sources of information: the teacher, the students, and the academic directors of the courses. It comprises three information collection instruments: the teacher self-report; the academic director's report; and the student satisfaction survey (Universidad de La Laguna, 2010). The fact that Docentia is an evaluation model that focuses on continuous improvement and is closely linked to the integration of a new university teaching approach centred on student learning means that periodic review of the model is necessary, in accordance with the Deming wheel.

With this in mind, and with the objective of optimising the evaluation system, we evaluated the results of the implementation of the Docentia-ULL model in the first three rounds, between 2010 and 2013. As a result of this analysis, various changes were proposed, modifying the weightings of the indicators and the weights of the model's dimensions (see Isla-Díaz et al., 2018). The proposal to review the student satisfaction survey, one of the dimensions of the Docentia model, derives from this. In this regard, we took into account the central importance for the teaching-learning model promoted by the EHEA of teachers individually accepting their evaluation, something that is dependant on the survey being valid in direct relationship with the teacher's individual performance. Changes can only be made if they have the support of the group entrusted with implementing them (Martínez & Esteban, 2005; Pozo et al., 2011; Valcárcel, 2005).

Teachers' good performance will necessarily be reflected in the satisfaction it generates, principally among their students. Indeed, in the context of the EHEA, the basic tool for evaluating teacher performance is the satisfaction survey, which is administered to students. This is one of the most widely used quality indicators in higher-education institutions around the world (Darwin, 2017; Pozo et al., 2011). It uses a Likert-type scale (Matosas-López, 2019). In the context of the Spanish University System, the surveys follow the model of the Docentia programme of ANECA (Spain's national quality and accreditation evaluation agency), varying both in the competences included, and in the number of items they comprise. For example, Muñoz et al. (2002) developed a questionnaire of 40 items that measure 10 competences. The questionnaire developed by Casero (2008) features 92 items that measure teacher performance in competences similar to those of Muñoz et al. (2002). For its part, the questionnaire developed by Molero and Ruiz (2005) comprises 25 items for four competences. A shorter version of the questionnaire with 18 items was developed by Lizasoain-Hernández et al. (2017) for the Universidad del País Vasco.

In general, most of the surveys that follow ANECA's Docentia model, as is the case of the Docentia-ULL approved in 2010, have around 20 items measuring a set of competences relating to the three dimen-

sions of the model — planning of teaching, delivery of teaching, and results — covering capacities relating to teacher performance (Jerez et al., 2016), such as the establishment of an appropriate assessment system (Sinahuya & Sánchez-Tarazaga, 2018) or the match between the content delivered and the credits assigned to the module. The results dimension considers students' satisfaction with the teaching practice of their teachers and the satisfaction of the academic directors.

If we observe the content of the student satisfaction surveys, it is apparent that they contain two different classes of items. On the one hand, there are the items that refer to the module, such as the match between the content delivered and the credits allocated, or the establishment of an appropriate assessment system. These often involve more than one teacher and must include the requirements established in the programme approval reports. On the other hand, there are items that relate solely to the teacher's performance, such as the capacity to motivate students and the provision of appropriate tutoring, which are regarded as key elements in the profile of a good university teacher (Caballero & Bolívar, 2015; San Martín et al., 2014; Ruiz-Esteban & Santos del Cerro, 2020; Tejedor & García-Valcárcel, 2010; Zabalza, 2009) and are more important for evaluating teacher performance.

Accordingly, the Teaching Evaluation Committee of the ULL identified the objective of finding an evaluation model that is not only reliable but also valid. In this regard, the fact that the survey in use included items relating to the module and items relating to the teacher who is being evaluated was identified as a significant obstacle for achieving this objective. Obviously, evaluation of the module might involve more than one teacher and so have a positive or negative effect on the result of the evaluation of the particular teacher. As a result, it is necessary to ensure that the results of the teachers who are evaluated using the survey genuinely reflect each individual's performance, evaluated through competences that are "under his or her control" and are separated from the evaluation of the module, which is often shared by various teachers. In accordance with motivational theories, such as self-determination theory (Ryan & Deci, 2017), an evaluation of performance that is coherent with teachers' competences would, in the long term, facilitate the development of the perception of control, of agency, which is fundamental for intrinsic motivation.

The committee in question established that the appropriate strategy for achieving this objective entailed redefining evaluation and, following an objective analysis methodology, developing two different surveys: one that evaluates teachers and another that evaluates the module. Furthermore, we understand that it is important to distinguish between the purpose of each of them, and this is of great importance for the university system. The teacher evaluation survey is the one that is used by Docentia as a tool for evaluating the individual performance of teachers. In contrast, the module evaluation survey will essentially be used as evidence in the programme accreditation processes.

We can only be sure that teachers will willingly accept the results of their evaluation and make the most of the feedback received to improve their teaching practice if they identify the evaluation received with their own performance. Accordingly, the improvement process required a clear differentiation between the items from one survey and those from the other; in this article we propose a survey that comprises the items that only evaluate the performance of the teacher (the evaluation of the modules with regards to programme accreditations will be done separately). Furthermore, the improvement process also requires a reduction in the number of items (see Castro Morera et al., 2020), so that students are not overwhelmed in the surveying process, something that is also an objective of this article.

2. Method

2.1. Participants

The data are drawn from 33,349 surveys of students' satisfaction with teaching completed for 689 participating teachers in the third and fourth teaching activity evaluation call at the ULL in the 2012/2013 (n = 367) and 2013/2014 (n = 322) years. Table 1 shows the distribution of surveys and teachers evaluated by branch of knowledge.

TABLE 1. Number of surveys and teachers evaluated by branch of knowledge.

	2012/2013		2013/2014		Total	
Branch of knowledge	Surveys	Teachers	Surveys	Teachers	Surveys	Teachers
Arts and Humanities	1815	48	2086	62	3901	110
Sciences	2199	62	1524	46	3723	108
Health Sciences	3661	70	3279	76	6940	146
Social and Legal Sciences	8731	119	5633	103	14 364	222
Engineering and Architecture	3107	68	1314	35	4421	103
Total	19513	367	13836	322	33349	689

Source: Own elaboration.

2.2. Instruments

The student satisfaction survey comprises 21 items (plus one item relating to general satisfaction) and uses a five-point Likert-type answer scale (1=Strongly disagree, 5=Strongly agree). This instrument measures student satisfaction with the teaching competences of their teachers in relation to the organisation and presentation of content, methodological strategies used in the teaching-learning process, the teachers' degree of compliance with attendance and the timetable, the students' perceived level of learning, and general satisfaction, and other aspects relating to the organisation of the module. This survey was approved at the 24/06/2010 meeting of the Board of Governors of the ULL for use during the 2010/2011-2014/2015

five-year period and can be found on page 52 of the ULL's manual for evaluation of teaching activity (Universidad de La Laguna, 2010).

2.3. Procedure

The surveys of students' satisfaction with the teaching practice of the participating teachers were completed during the last two months of each term of the 2012/13 and 2013/14 years. This information collection was done in-person in class, using clickers provided by surveyors from the Technical Quality Unit of the ULL. A day and time were agreed in advance with the teachers involved to avoid potential changes in time or place and ensure an efficient data collection process.

The modification of the survey (identifying the items that relate to the module and those that relate to the teacher) with a view to implementing it in the second Docentia-ULL five year period (2015–2020) was done under the leadership of the Teaching Evaluation Committee with the participation of experts[1].

2.4. Data analysis

We calculated descriptive statistics, principal component analyses, reliability by internal consistency, correlations, Anova and Chi-squared using the SPSS v.21 (IBM, 2012) statistics package. After selecting the items, we carried out a dimensional analysis of the psychometric properties of the teacher evaluation survey using exploratory factor analysis and we examined the possible effect of the branch of knowledge. We compared the teacher scores obtained from the original survey with the scores that would have been obtained using the "simulation" of a reduced survey with the data from the 2012 to 2014 years. In addition, we compared the results from the original survey (2012-2014 rounds) with the published results from the 2017-2019 rounds, from the second five-year evaluation period, where the reduced survey was used, slightly modified by the Evaluation Committee for this five-year period (Universidad de La Laguna, 2017).

3. Results

First we performed a study of the completed surveys (33,349), analysing possible response tendencies by the students, in other words, detecting surveys where one of the options has been selected too often according to the procedure described by Correa & Camacho (1993). This left 32,297 surveys. In addition, we excluded students who reported not attending class at least "somewhat" (3). Accordingly, we eventually analysed 28,965 valid surveys. In addition, for the analyses that include items 13 ("It is easy to access the teacher in his/her tutorial times") and 14 ("The help received in tutorials is effective for learning"), we excluded students who reported not attending tutorials at least "somewhat" (3).

3.1. Cleansing and structure of the questionnaire

Next, we analysed the content of the 21 items (without taking into account item 22 on general satisfaction) to re-

move from the survey any for which the teacher is not solely responsible but which instead relate to organisational or regulatory questions regarding the modules and so will be included in the process of evaluation of programmes for their accreditation.

As a result, we eliminated 8 items (the numbering corresponds to that of the original survey from 2010), as well as item 9, as in the university context of the EHEA it is not viable to reduce the requirements established in the programme approval reports. Similarly, and to avoid highly redundant items, we calculated the Pearson correlations and observed that item 20 has a correlation of at least 0.7 with items 11, 19, and 21, and so we decided to eliminate this as well.

1. The information provided by the teacher in the module guide (or handbook) is accessible and useful.

2. The planned assignments relate to what the teacher wants us to learn in the teaching activity.

3. In the delivery of the teaching activity there is no overlap with the content of other activities.

4. The theoretical and practical classes are coordinated.

5. The credits assigned to the teaching activity are in proportion with the volume of content and proposed assignments.

6. The effort this teaching activity requires corresponds with that set out in the module guide.

16. The bibliography recommended by the teacher is useful for developing the topics.

18. The teacher applies the evaluation criteria contained in the module guide appropriately.

19. The teacher adapts the programme depending on students' prior level of knowledge.

20. The teacher has facilitated my learning. With his/her help I have been able to improve my knowledge, skills, or way of approaching particular topics.

Table 2 shows the descriptors of the 11 remaining items from the original survey, which will be subjected to principal component analysis.

To determine the internal structure of the retained items, we performed a principal component analysis. The KMO value is 0.93, and so the data are appropriate for a factorial model. The model captures 63.49% of the variance in two components with an eigenvalue greater than 1.

We carried out a varimax rotation, which showed good loadings ranging between .50 and .83 for the first component and greater than .82 for the second. Given that this second component only consists of the two items that refer to the teacher's compliance

with attendance and the timetable (items 7 and 8) and that the scree plot recommends a single factor model, we repeated the analysis limiting the model to one component.

This single factor model captures 52.42% of the variance. Table 2 shows the weights of the resulting component matrix, which range between .439 and .836.

TABLE 2. Descriptors of the selected items.

Items		Mean (*SD*)	*n* analysed (*n* excluded)	Rotated components weight
7.	The teacher complies with the set timetable.	4.11 (1.20)	28 570 (381)	.517
8.	The teacher attends class regularly.	4.62 (0.81)	28 656 (295)	.439
10.	The teacher does a good job of preparing, organising, and structuring the activities or assignments done in class (or laboratory, workshop, fieldwork, seminar, etc.).	3.56 (1.23)	28 491 (460)	.795
11.	The teacher explains clearly.	3.51 (1.35)	28 653 (298)	.813
12.	The teacher solves doubts and guides students in the completion of their assignments	3.63 (1.23)	28 422 (529)	.836
13.	It is easy to access the teacher in his/her tutorial times	3.50 (1.24)	14 724 (14 227)	.625
14.	The help received in tutorials is effective for learning	3.47 (1.20)	11 711 (17 240)	.764
15.	The teacher uses teaching resources effectively to facilitate learning	3.46 (1.20)	28 008 (943)	.803
17.	The teacher favours the participation of the student in the delivery of the teaching activity	3.61 (1.22)	28 115 (836)	.701
19.	The teacher manages to inspire interest in the different topics covered in the delivery of the teaching activity	3.27 (1.35)	28 108 (843)	.799
21.	I have improved on my starting level, with relation to the competences listed in the module guide	3.41 (1.16)	28 058 (893)	.751

Source: Own elaboration.

The internal consistency analysis gave a Cronbach's alpha of .915, increasing slightly to .916 or .918, if items 7 and 8 are eliminated. For this reason, and also because they are the items with the weakest relationship with the component and both refer to the teacher's compliance with attendance and the timetable, we suggest combining them into a single item: "The teacher attends class regularly and complies with his/her timetable".

Furthermore, we suggest reformulating the original items 10, 12, 13, 14, and 15 to facilitate their comprehension by involving the students' personal experience (see Appendix).

With the objective of supplementing the last item in the survey, regarding general satisfaction ("In general I am satisfied with this teacher's teaching work"), we suggest adding a new item "I would take another module with this teacher", but we recommend separating it spatially from the former in the presentation of the survey.

Ultimately, we propose a single factor questionnaire with 10 items as well as two items relating to general satisfaction (see Appendix).

3.2. Comparison of the structure across branches of knowledge

To confirm that the structure remains stable independently of the branch of knowledge, we performed a principal component analysis for each branch: arts and humanities (AaH), sciences (S), health sciences (HS), social and legal sciences (SaLS), and engineering and architecture (EaA).

The structure is maintained almost perfectly across the five branches of knowledge. Percentages of explained variability of between 50.37 and 55.87% are obtained. Table 3 shows the loadings of the items in the component as well as the place the component occupies as a function of this loading. The internal consistency according to Cronbach's alpha is greater than 0.9 in all cases. It is notable that item 12 is always in first place and that items 17, 13, 7, and 8 always occupy the last four places.

TABLE 3. Comparison of loadings and place across the different branches of knowledge.

Rama / **% de Varianza explicada**	**General** **52.42**		**AyH** **52.33**		**C** **51.73**		**Cdis** **55.87**		**CsyJ** **50.37**		**IyA** **54.71**	
Ítem	lugar	peso	lugar	peso	lugar	peso	lugar	peso	lugar	peso	lugar	peso
R12 solves doubts	1	.84	1	.83	1	.84	1	.85	1	.82	1	.85
R11 explains clearly	2	.82	3	.81	2	.82	2	.83	2	.81	2	.82
R15 appropriate use of teaching resources	3	.81	2	.81	6	.78	4	.82	3	.8	4	.81
R19 motivates	4	.80	4	.79	3	.81	3	.82	4	.79	5	.79
R10 prepares activities well	5	.80	5	.78	5	.80	5	.81	5	.78	3	.82
R14 tutorials help effectively	6	.77	6	.78	4	.80	6	.77	6	.75	7	.76
R21 I have improved my competences	7	.75	7	.75	7	.76	7	.75	7	.75	6	.78
R17 favours participation	8	.70	8	.70	8	.67	8	.73	8	.69	8	.74
R13 accessible tutorials	9	.63	9	.64	9	.61	9	.67	9	.61	9	.59
R7 complying with timetable	10	.52	10	.53	10	.49	10	.57	10	.45	10	.57
R8 attending class	11	.45	11	.40	11	.37	11	.51	11	.40	11	.51
Cronbach's Alpha	0.92		0.92		0.92		0.93		0.90		0.92	

Source: Own elaboration.

3.3. Comparison of teacher scores from the original survey with the "simulation" of the reduced one (2012-2014)

To compare the real scores obtained in the evaluation with the score that would have been obtained with the 11 items selected for the new survey (without combining items 7 and 8), we performed a 2 × 5 mixed Anova with the *score* factor (original score with 21 items vs simulation with the 11 selected items) and the *branch of knowledge* factor.

The principal effects of both factors are significant: the score obtained with the reduced survey (M = 3.56; SD = 0.82) is slightly higher than the original (M = 3.67; SD = 0.87) in all of the branches as a group ($F_{1, 28946}$ = 4801.27; $p < .001$; η_p^2 = .142) (see Table 4).

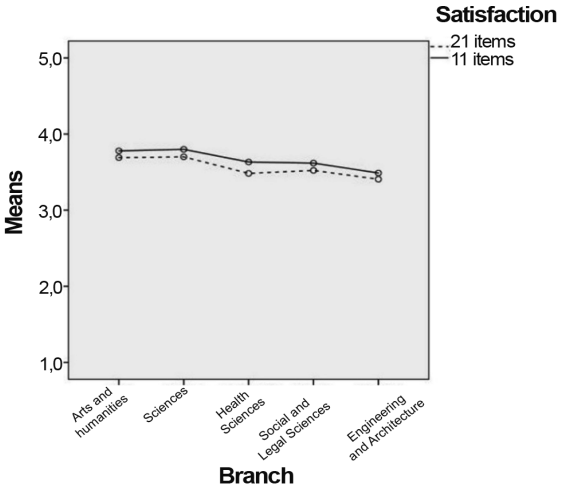
TABLE 4. Descriptive statistics for the complete 21-item survey and the simulated 11-item version, by branch and version.

Branch		21 items	11 items
Arts and Humanities	Mean	3.72	3.81
	SD	0.83	0.86
	n	3535	3535
Sciences	Mean	3.67	3.77
	SD	0.77	0.81
	n	3419	3419
Health Sciences	Mean	3.50	3.65
	SD	0.85	0.91
	n	5641	5641
Social and Legal Sciences	Mean	3.58	3.68
	SD	0.78	0.83
	n	12335	12335
Engineering and Architecture	Mean	3.34	3.42
	SD	0.88	0.94
	n	4021	4021
Total	Mean	3.56	3.67
	SD	0.82	0.87
		28951	28951

Source: Own elaboration.

By comparison, the scores differ slightly by branch ($F_{1, 28946}$ = 127.25; $p < .001$; η_p^2 = .017), although the mean score is greater than 3 in all branches (see Graph 1).

GRAPH 1. Comparison of the scores in the complete 21-item survey and the simulated 11-item survey by branch and version.



Source: Own elaboration.

We performed a Chi-square test of independence to compare the distributions of the teachers in the 4 possible categories of results in the 12-14 rounds (n = 598) from the complete survey (21 items — without the general satisfaction item) compared with the reduced 11-item survey. Although this is significant (χ^2_3 = 14.0; p = .003; Cramér's V = .08), Cramér's V confirms that there are no substantive differences between the two. Nonetheless, it is apparent (see Table 5) that the reduced survey has a greater percentage of teaching and research staff in the *Excellent* category (from 21.3% to 29.9%) and in the *Unfavourable* category (from 4.3% to 5.5%). This trend is repeated across all of the branches to a greater or lesser extent.

TABLE 5. Distribution of the teachers in the categories of results in the 2012-2014 rounds of the mean score of the complete 21-item survey compared with the simulated 11-item survey: % observed by columns (n).

Result category	Score	21 items	11 items
Excellent	4 o más	21.3 % (125)	29.9 % (179)
Very favourable	3.25 - 3.99	50.2 % (295)	44.1 % (264)
Favourable	2.5 - 3.24	24.3 % (143)	20.4 % (122)
Unfavourable	less than 2.5	4.3 % (25)	5.5 % (33)
	Total	100 % (588)	100 % (598)

Source: Own elaboration.

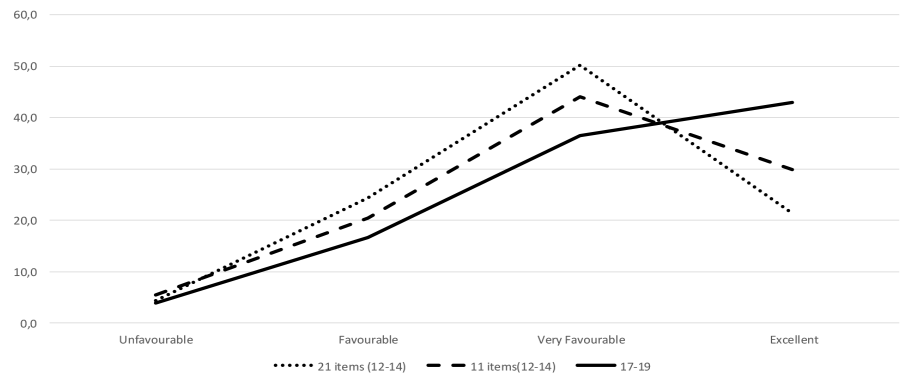
To compare the correlation between general satisfaction and mean satisfaction with the teachers (11 items) (r = .825; p< .001) and its correlation with the mean satisfaction with the module (8 items) (r = .719; p < .001), we performed Hotelling's test (t(31335) = 54.63; p < .001), with the result that the correlation between general satisfaction and mean satisfaction with the teacher is greater, sharing 68.06% of their variability compared with the 51.7% it shares with the mean module score (this difference would also be significant with just 50 participants). Furthermore, the mean score of the 8 items referring to the module and the mean score of the 11 items referring to the teacher share 63.84% of their variance (r = .799; p < .001). In other words, despite being closely related, they do not measure the same thing.

3.4. Comparison of the real results of the 2012-2014 call and the 2017-2019 call

Of the 589 teachers who participated in the 2012-2014 rounds (n = 367+322), 93.58% coincide with those evaluated again in the 2017-2019 rounds (n = 285+241) applying the reduced survey. Therefore, we compared the distribution in the different student satisfaction result categories, combining the 2012-2013 and 2013-2014 rounds (referred to as 2012–2014) with the results from the combination of the 2017-2018 and 2018-2019 rounds (referred to as 2017-2019). This comparison between the 11 items from the 2012–2014 rounds and the survey applied in the 2017–2019 rounds is significant (χ^2_3 = 20.5; p < .001), although the difference between the distributions is not substantive (V = .10). Nonetheless, an increase in the *Excellent* category can be seen (from 29.9% to 43%) (see Table 6).

Graph 2 shows the distributions of teaching and research staff in the results categories in the 2012-2014 rounds with the survey of 21 items, with the simulation

GRAPH 2. Distribution of teachers in the categories of results in the 2012-2014 rounds of the complete 21-item survey, the 11-item simulated survey, and the real 11-item survey applied in the 2017-2019 rounds.



Source: Own elaboration.

of 11 items referring to the teachers, and in the 2017-2019 rounds actually applying the reduced survey. It is apparent that the percentage of *Very favourable* teachers in the 2012-2014 rounds falls somewhat in the 11-item simulation and even more so in the real 11-item survey of the 2017-2019 rounds, while at the same time the percentage in the *Excellent* category increases.

Furthermore, the distribution between results from the general satisfaction items from the 2012-2014 rounds compared with the 2017-2019 rounds is significant ($\chi^2_3 = 10.8$; $p = .001$), although the difference between the distributions again is not substantive ($V = .07$). Nonetheless, an increase is also apparent in the *Excellent* category (from 41.4% to 50.8%) (see Table 6).

TABLE 6. Distribution of teachers in the categories of results of the mean score obtained in the 11 items and general satisfaction items from the reduced questionnaire by rounds: % observed by columns (*n*).

		11 items		**Satisfaction Items**	
Result category	Score	2012-2014	2017-2019	2012-2014	2017-2019
Excellent	4 or more	41.4 % (242)	50.8 % (259)	29.9 % (179)	42.9 % (219)
Very favourable	3.25 - 3.99	30.7 % (179)	26.9 % (137)	44.1 % (264)	36.5 % (186)
Favourable	2.5 - 3.24	20.4 % (119)	15.1 % (77)	20.4 % (122)	16.7 % (85)
Unfavourable	less than 2.5	7.5 % (44)	7.3 % (37)	5.5 % (33)	3.9 % (20)
	Total	100 % (581)	100 % (510)	100 % (598)	100 % (510)

Source: own elaboration

4. Discussion

In this study we have examined the survey of students' satisfaction with the teaching they receive, which is of central important in the Docentia process and for the accreditation processes of the programmes. We have proposed a change, separating evaluation of the teachers' performance from evaluation of the module. Optimal evaluation of teaching is, undoubtedly, an important issue for the quality of higher-education institutions (Álvarez Rojo et al., 2009; Benito & Cruz, 2005; High Group on Modernization of Higher Education, 2013; Mayor, 2009; Murillo, 2008; Perales et al., 2014). To the best of our knowledge, this work is an innovative proposal, as it has enabled us to fulfil the objective of separating the teaching quality survey from content that evaluates the quality of a module, without any resulting loss of relevant information with regards to the original survey of the Docentia-ULL model.

The cleansing and study of the structure of the questionnaire and its comparison across the branches of knowledge lead us to propose a version with 10 items relating to performance for which the teacher evaluated is solely responsible. We also added two items relating to students' general satisfaction with the teaching practice of the teachers. This survey uses a single-factor model with high internal consistency that captures more than half of the variance, on the same lines as what is obtained by Castro Morera et al. (2020), and which is stable across the branches of knowledge.

The three-way comparison of the scores obtained by the teachers (original 21-item survey, simulated 11-item survey, and real application of the new proposed survey) makes it possible to ensure that information has not been lost in the process of improvement. There are important differences in the evaluation of teachers depending on whether one survey or the other is used, which we believe endorses the need and opportunity to have two student satisfaction surveys: one on the teacher and one on the module. In this regard, we have found that both in the simulation carried out with the short survey with the sample of teachers from the 2012-2014 rounds, and in the one already applied in the 2017-2019 rounds, there is a substantial increase in *Excellent* teachers when the short survey is applied without the items referring to the module. This is similar to what happens when the distributions of the general satisfaction item are studied, which display a greater percentage of *Excellents* compared with the corresponding averages of items. The implementation of a short survey with indicators of personal teaching skills, which are directly controlled by the actions of the teacher being surveyed, makes it possible to increase the percentage of teaching staff in the *Excellent* category, at the cost of a reduction in the *Favourable* and *Very favourable* categories, which García Martín et al. (2020) suggested was liable to happen.

According to previous literature (Caballero & Bolívar, 2015; Ruiz-Esteban & Santos del Cerro, 2020; San Martín et al., 2014; Tejedor & García-Valcárcel, 2010; Zabalza, 2009), the items from the teacher survey actually measure competences that are of high value in teaching performance and, in particular, they measure competences that are under the teacher's control. As has been indicated, items such as the organisation of credits of the module and the evaluation system are imposed on the teacher by the programme approval reports; in contrast, items such as motivation, inspiring interest, and appropriate tutoring directly relate to the teacher's own performance. The fact a teacher has control over these competences entails, in accordance with motivational theories (Ryan & Deci, 2017), modifying the context of the teacher to give him or her more autonomy. So, a performance evaluation that is better linked to each teacher's teaching practice would, in the long term, facilitate the development of agency, which is fundamental for intrinsic motivation (Gámez & Marrero, 2006; Gámez et al. 2021). Furthermore, the fact that some colleagues obtain a high degree of student satisfaction with the teaching they provide, linked to competences that the teacher might learn and develop, increases motivation

to improve their teaching, resulting in a healthy degree of competitiveness. This could explain the higher percentage of *Excellents* obtained with the real application of the proposed survey in comparison with the simulated model. Both surveys were applied to virtually the same teachers, and so it appears that these teachers have motivated themselves between one call and the next to improve competences that are under their control, something reflected in the perception of their students. Teachers are aware that they will be periodically evaluated and seem to have made an effort to improve their performance. This supports the role of Docentia in incentivising teachers to improve their teaching quality.

Teachers deserve to be evaluated on their own individual performance merits. They should not be evaluated on the performance of the teaching team responsible for delivering the module or on the performance of other people involved in the design of the programme approval model, which the teacher might not have been involved in preparing. It is important to note that the long survey makes it possible to mask teachers with less satisfactory teaching performance, as they sometimes benefit from better organisation of the module done jointly with their colleagues. So, the simulated model also displayed an increase in the percentage of *Unfavourables*, which is in line with this. The teacher survey is, therefore, fairer and more equitable, with a greater ecological validity than the long survey given that, as noted above, the long survey masks teachers with less satisfactory performance. It is also more parsimonious as it has fewer items and these clearly relate to the object of the evaluation — the teacher — facilitating its completion by the students and, ultimately, its validity.

One important result of the division of the survey into groups of items relating to the teacher and items relating to the module is the relationship of each of them with the general satisfaction item: the correlation between general satisfaction and the mean satisfaction with the teacher is considerably greater, sharing 68.06% of variability compared with the 51.7% shared with the mean score for the module. The perception of teachers' skills and their attitude in the classroom are the most significant aspects with regards to student satisfaction (Guevara & Stewart, 2011; Leguey-Galán et al., 2018; Ruiz-Esteban & Santos del Cerro, 2020). This result supports the convergent validity of the new survey, given that the items referring to the teacher are better related to student satisfaction, which is the type of measurement we are seeking within Docentia as an evaluation model to follow. Being evaluated on what one does *individually* instead of being evaluated for what *others* do seems fairer.

Ultimately, separating the student satisfaction survey into two surveys has a noteworthy institutional value. These surveys are appropriate for particular quality evaluation process. The quality of the programmes directly relates to students' satisfaction with the modules (that is to say: what is delivered) and it is the type of measurement that should be taken into account in accreditation processes. Conversely, the survey of satisfaction with teachers is the appropriate type of

measurement to evaluate the performance of the teacher (in other words: how the module is delivered) and it is an important instrument in promotion and financial incentivisation processes.

As well as the institutional importance, it is also necessary to mention the job performance evaluation aspect, and the fact is that all workers — and university teachers are workers — have the right to fair evaluation, meaning that the evaluation is directly related to their personal performance. We do not want to say that coordination in the module is not an element to consider in the evaluation, but the fact is that, for now, the good or bad disposition of the teachers of the module as a group to deliver good teaching, does not just depend on each individual teacher. Nonetheless, we believe that this element of the cooperation of the teacher in coordinating the module deserves reflection with the object of incorporating it in the performance evaluation, albeit clearly not through the student satisfaction survey, but rather through the academic directors' reports.

In this context, we hope that the results we have obtained will contribute to the reform of evaluation of student satisfaction within the Spanish university system, linking each type of survey to the corresponding quality evaluation process. In summary, the review of the verified Docentia-ULL model offered an opportunity to guide teaching activity in line with the quality objectives that correspond to a higher-education institution, strategically incentivising the evaluation of teachers' performance in a truly individual way.

Note

[1] We would like to thank the members of the Teaching Quality Evaluation Committee (2012-2015) and of the Technical Quality Unit of the ULL for their close collaboration and significant involvement, in particular: Marcos Blanco-Freijó, Victoria Pérez-Monteverde, Severo Acosta-Rodríguez, Francisco Jiménez-Moreno, Aixa Noda-Ramos, Teresa Ramos-Domínguez, and Isabel Soriano-Torres.

Appendix

Final proposal for a 10-item survey plus two general satisfaction items.

Current	Original	Statement
1		***I would take another module with this teacher*
2	7 and 8	*The teacher attends class regularly and complies with his/her timetable.*
3	10	The teacher prepares, *and* organises the *teaching* activities done in class well (or in laboratories, workshops, field work, seminars, etc.).
4	11	The teacher explains clearly.
5	12	The teacher solves *my* doubts and guides *me* in the completion of my assignments.
6	13	*I find it it* easy to access the teacher *through* tutorials.
7	14	I find the help *and support I receive from the teacher useful* for learning.

8	15	*The teaching resources the teacher uses facilitate my learning.*
9	17	The teacher favours the participation of the student in the implementation of the teaching activity.
10	19	The teacher manages to inspire *my* interest in the different topics covered in the delivery of the teaching activity.
11	21	I have improved on my starting level, *in* relation to the competences listed in the module guide.
12	22	**In general I am satisfied with this teacher's teaching practice.

The changes to wording and new proposal are in italics.
**General satisfaction item.
Source: Own elaboration.

References

Alfageme, M. B., & Caballero, K. (2010). Evaluación y profesorado en la universidad española [Assessment and faculty in the Spanish university]. *Educatio Siglo xxi*, *28* (1), 271-294. https://t.ly/viNO

Álvarez-Rojo, V., Asensio-Muñoz, I., Clares, J., Del-Frago, R., García-Lupión, B., García-Nieto, N., García-García, M., Gil, J., González-González, D., Guardia, S., Ibarra, M., López-Fuentes, R., Rodríguez-Diéguez, A., Rodríguez-Gómez, G., Rodríguez-Santero, J., Romero, S., & Salmerón, P. (2009). Perfiles docentes para el espacio europeo de educación superior (EEES) en el ámbito universitario español [Teaching profiles for the European Higher Education Area (EHEA) in Spanish universities]. *RELIEVE, 15* (1). https://doi.org/10.7203/relieve.15.1.4187

Benito, A., & Cruz, A. (2005). *Nuevas claves para la docencia universitaria en el Espacio Europeo de Educación Superior [New keys to university teaching in the European Higher Education Area]*. Narcea.

Caballero, K., & Bolívar, A. (2015). El profesorado universitario como docente: hacia una identidad profesional que integre docencia e investigación [Academics' identity as teachers: towards the Scholarship of Teaching and Learning]. *REDU: Revista de Docencia Universitaria*, *13*, 57-77. https://t.ly/ZRq3

Calderón, C., & Escalera, G. (2008). La evaluación de la docencia ante el reto del Espacio Europeo de Educación Superior (EEES) [Teaching assessment against the challenge of the EUROPEAN highereducation area]. *Educación XX1*, *11*, 237-256. https://t.ly/iiKN

Casero, A. (2008). Propuesta de un cuestionario de evaluación de la calidad docente universitaria consensuado entre alumnos y profesores [Proposal of a questionnaire for the evaluation of university teaching quality agreed upon by students and professors]. *Revista de Investigación Educativa*, *26* (1), 25-44. https://t.ly/gzk9

Castro Morera, M., Navarro-Asencio, E., & Blanco-Blanco, Á. (2020). La calidad de la docencia percibida por el alumnado y el profesorado universitarios: análisis de la dimensionalidad de un cuestionario de evaluación docente [The quality of teaching as perceived by university students and faculty: dimensionality analysis of a teaching evaluation questionnaire]. *Educación XX1*, *23* (2), 41-65. https://doi.org/10.5944/educxx1.25711

Correa, A. D., & Camacho, J. (1993). Diseño de una metodología para el estudio de las teorías implícitas [Design of a methodology for the study of implicit theories]. In M. J. Rodrigo, A. Rodríguez, & J. Marrero. *Las teorías implícitas. Una aproximación al conocimiento cotidiano* (pp. 123-165). Visor.

Darwin, S. (2017). What contemporary work are student ratings actually doing in higher education? *Studies in Educational Evaluation*, *54* (Supplement C), 13-21. https://t.ly/WD5f

Gámez, E., & Marrero, H. (2006). *¿Por qué hacemos lo que hacemos? Dimensiones básicas de la motivación [Why do we do what we do? Basic dimensions of motivation]*. Idea.

Gámez, E., Díaz-Gómez, J. M., Ardévol, A., & Marrero, H. (2021). *Dimensiones Básicas de la Motivación Humana [Basic Dimensions of Human Motivation]*. Ed. Pirámide.

García Martín, A., Montero Cases, T., García León, J., & Vázquez Arenas, G. (2020). Validity of student satisfaction surveys to assess teaching quality: the UPCT case study (Cartagena, Spain). *REDU. Revista de Docencia Universitaria, 18* (1), 275-290. https://doi.org/10.4995/redu.2020.12996

Guevara, C., & Stewart, S. (2011). Do student evaluations match alumni expectations? *Managerial Finance*, *37* (7), 610-623. https://doi.org/10.1108/03074351111140252

High Group on Modernization of Higher Education (2013). *Improving the quality of teaching and learning in Europe's higher education institutions*. https://goo.gl/rlc3Sr

IBM Corp. Released 2012. *IBM SPSS Statistics for Windows, Version 21.0*. IBM Corp.

Isla-Díaz, R., Marrero-Hernández, H., Hess-Medler, S., Soriano, M., Acosta-Rodríguez, S., Pérez-Monteverde, M. V., & Blanco-Freijo, M. (2018). Una mirada longitudinal: ¿Es el «Docentia» útil para la evaluación del profesorado universitario? [A Longitudinal Look: Is "Docentia" an useful instrument for university teacher's evaluation?] *RELIEVE, Revista Electrónica de Investigación y Evaluación Educativa, 24* (2). https://doi.org/10.7203/relieve.24.2.12142

Jerez, Ó., Orsini, C., & Hasbún, B. (2016). Atributos de una docencia de calidad en la educación superior: una revisión sistemática [Attributes of quality teaching in higher education: a systematic review]. *Estudios Pedagógicos*, *42* (3), 483-506. https://doi.org/10.4067/S0718-07052016000400026

Leguey-Galán, S., Leguey-Galán, S., & Matosas-López, L. (2018). ¿De qué depende la satisfacción del alumnado con la actividad docente? [What drives students' satisfaction with instruction?] *Revista Espacios*, *39* (17), 13-29. https://t.ly/NkpV

Lizasoain-Hernández, L., Etxeberria-Murgiondo, J., & Lukas-Mujika, J. F. (2017). Propuesta de un nuevo cuestionario de evaluación de los profesores de la Universidad del País Vasco. Estudio psicométrico, dimensional y diferencial [A proposal for a new questionnaire for the evaluation of teachers at the University of the Basque Country. Dimensional, differential and psychometric study]. *RELIEVE - Revista Electrónica de Investigación y Evaluación Educativa*, *23* (1). https://doi.org/10.7203/relieve.23.2.10436

Martínez, M., & Esteban, F. (2005). Una propuesta de formación ciudadana para el EEES [A proposal for citizenship education for the EHEA]. **revista española de pedagogía**, *63* (230), 63-73. https://cutt.ly/qkYoUf0

Matosas-López, L. (2019). Diferencias en las puntuaciones de las encuestas de valoración del profesorado en función del tipo de cuestionario: comparativa cuestionarios Likert vs cuestionarios BARS [Differences in teacher evaluation survey scores according to the type of questionnaire: Comparison of Likert vs. BARS questionnaires]. *Revista Infancia, Educación y Aprendizaje*, *5*, 371- 378.

Mayor, C. (2009). Nuevos retos para una Universidad en proceso de cambio: ¿Pueden ser los profesores principiantes los protagonistas? [New challenges for a changing university: can junior faculty be the protagonists?] *Revista de Currículum y Formación del Profesorado*, *13* (1), 61-77. https://goo.gl/JrS8lV

Molero, D., & Ruiz Carrascosa, J. (2005). La evaluación de la docencia universitaria. Dimensiones y variables más relevantes [The evaluation of university teaching. Dimensions and most relevant variables]. *Revista de Investigación Educativa*, *23* (1), 57-84.

Muñoz, J. M., Ríos, M. P., & Abalde, E. (2002). Evaluación docente vs. Evaluación de la calidad [Teacher evaluation vs. quality evaluation]. *Revista electrónica de Investigación y Evaluación Educativa*, *8* (2), 103-134. https://t.ly/sZW8

Murillo, J. (2008). La evaluación del Profesorado Universitario en España [The assessment of university teachers in Spain]. *Revista Iberoamericana de Evaluación Educativa*, *1* (3), 29-45. https://t.ly/fLaW

Perales, M. J., Jornet, J. M., & González, J. (2014). Tendencias en las políticas de formación y evaluación del profesorado en la Educación Superior en España [Policy trends in training and evaluation of teachers in Higher Education in Spain]. *Revista Iberoamericana de Evaluación Educativa*, *2*, 53-64. https://t.ly/shYn

Pozo, C., Bretones, B., Martos, M. J., & Alonso, E. (2011). Evaluación de la actividad docente en el Espacio Europeo de Educación Superior: un estudio comparativo de indicadores de calidad en universidades europeas [Evaluation of teaching activity in the European Higher Education Area: A com-

parative study of quality indicators in European universities]. **revista española de pedagogía**, *248* (69), 145-163. https://t.ly/SBJh

Ruiz Esteban, C. M., & Santos del Cerro, J. (2020). Teaching quality: The satisfaction of university students with their professors. *Anales De Psicología / Annals of Psychology*, *36* (2), 304-312. https://doi.org/10.6018/analesps.335431

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. The Guilford Press.

San Martín, S., Santamaría, M., Hoyuelos, F. J., Ibáñez, J., & Jerónimo, E. (2014). Variables definitorias del perfil del profesor/a universitario/a ideal desde la perspectiva de los estudiantes pre-universitarios/as [Defining variables of the profile of the ideal university teacher from the perspective of pre-university students]. *Educación XX1*, *17* (2), 193-215. https://t.ly/9H9J

Sinahuya, A., & Sánchez-Tarazaga, V. (2018). La competencia evaluativa de los docentes: formación, dominio y puesta en práctica en el aula [The assessment competence of teachers: training, expertise and classroom implementation]. *Revista Iberoamericana de Educación*, *76* (2), 95-116. https://doi.org/10.35362/rie7623072

Tejedor, F. J., & García-Valcárcel, A. (2010). Evaluación del desempeño docente [Teacher performance evaluation]. **revista española de pedagogía**, *68* (247), 439-460. https://bit.ly/2VXU69g

Universidad de La Laguna (2010). *Manual para la Evaluación de la Actividad Docente (DOCENTIA-ULL).* http://cutt.ly/PfClayO

Universidad de La Laguna (2017). *Manual de la Universidad de La Laguna para la evaluación de la actividad docente [Manual of the University of La Laguna for the evaluation of teaching activities].* https://cutt.ly/ZfCldkg

Valcárcel, M. (2005). *La preparación del profesorado universitario español para la Convergencia Europea en Educación Superior [The preparation of Spanish university teachers for the European Convergence in Higher Education].* Proyecto EA2003-0040. https://goo.gl/iGdTTO

Zabalza (2009). Ser profesor universitario hoy [Being a university teacher today]. *La Cuestión Universitaria, 5*, 69-81.

Authors' biographies

Rosa Isla-Díaz is Associate Professor in the area of social psychology at Universidad de La Laguna, teaching on the programmes in Psychology, Pedagogy, and Employment Relations, among others . Her main lines of research include the role of psychosocial factors in organisational security or the factors contributing to the evaluation of performance at work.

https://orcid.org/0000-0003-0252-1873

Stephany Hess-Medler is Associate Professor in the area of Methodology of Behavioural Sciences at the Universidad de La Laguna, where she took her doctorate. She teaches on the Bachelor's in Psychology, the University Masters' degrees in HR Development and Management, and in General Health Psychology. She is a member of the Environmental Psychology and Neuropsychology research teams.

https://orcid.org/0000-0002-0289-8796

Hipólito Marrero-Hernández is a Full Professor in Basic Psychology. He is head of the Motivation and Brain research group in the University Neurosciences Institute. He teaches on Motivation and Emotion and on doctorate of excellence programmes. His research centres on approach and avoidance affect, language and social cognition, motivated approach/avoidance reasoning, and university students' motivation when choosing a degree.

https://orcid.org/0000-0002-6008-3587