



POLITÉCNICA

Estudio estadístico de los indicadores de calidad en los títulos de edificación de la UPCT

Statistical study of quality indicators in UPCT building titles

Gemma Vázquez Arenas¹, Antonio García Martín², Josefina García-León¹, Teresa Montero Cases³

¹Departamento de Arquitectura y Tecnología de la Edificación, Escuela Técnica Superior de Arquitectura y Edificación, Universidad Politécnica de Cartagena

²Departamento de Ingeniería Minera y Civil, Escuela Técnica Superior de Ingeniería de Caminos, Canales y Puertos e Ingeniería de Minas, Universidad Politécnica de Cartagena

³Departamento de Matemática Aplicada, Escuela Técnica Superior de Ingeniería Industrial, Universidad Politécnica de Cartagena

Recibido: 21/10/2020 | Aceptado: 08/12/2020 | Fecha de publicación: 31/12/2020
DOI: 10.20868/abe.2020.3.4518

RESUMEN

El trabajo que se presenta ha consistido en un análisis de los datos de indicadores de calidad de la UPCT utilizando herramientas estadísticas sencillas, acompañado de una revisión bibliográfica de numerosas publicaciones sobre las encuestas de satisfacción y la carga de trabajo. Se han empleado los datos de los títulos de Arquitectura y Edificación de la UPCT durante el curso 2017-2018 y, para contextualizar los resultados, también los del conjunto de esa Universidad y su comparativa con dos cursos anteriores. El objetivo ha sido conocer la relación entre las variables disponibles, su evolución temporal y las posibles diferencias por cursos, tipos de asignatura, niveles, etc., para mejorar los procesos de gestión de la información de la UPCT. Se discuten los resultados obtenidos, entre ellos la no existencia de sesgos en las encuestas de satisfacción y las posibles implicaciones de cargas de trabajo elevadas en determinadas asignaturas.

Palabras clave: *Indicadores de calidad, Evaluación de la actividad docente, Carga de trabajo, Evaluación de títulos universitarios.*

ABSTRACT

The work presented has consisted of an analysis of the UPCT quality indicator data using simple statistical tools, accompanied by a bibliographic review of numerous publications on satisfaction surveys and workload. The data of the degrees in Architecture and Building of the UPCT during the academic year 2017-2018 have been used and, to contextualize the results, also those of the whole of that University and its comparison with two previous courses. The objective has been to know the relationship between the available variables, their temporal evolution and the possible differences by course, type of subject, levels, etc., in order to improve the information management processes of the UPCT. The results obtained are discussed, including the non-existence of biases in the satisfaction surveys and the possible implications of high workloads in certain subjects.

Keywords: *Quality indicators, Evaluation of teaching activity, Workload, Evaluation of university degrees.*

1. INTRODUCCIÓN

Uno de los requisitos esenciales del Espacio Europeo de Educación Superior es la necesidad que las universidades evalúen de forma periódica la calidad de los títulos que imparten [1, 2], de la investigación y de la docencia, entre otros aspectos. La evaluación de la calidad docente se justifica, por tanto, por dos razones: la rendición de cuentas y la mejora continua.

Las encuestas de satisfacción de los estudiantes con la actividad docente constituyen la herramienta más habitual de

evaluación docente universitaria, además de emplearse en los sistemas de garantía de calidad de los títulos y en los modelos de evaluación de la docencia. No obstante, su validez siempre ha sido objeto de un debate que se remonta casi a sus orígenes y que se ha materializado en un gran número de publicaciones, sobre todo en los países anglosajones. Muchas de las opiniones desfavorables al uso de las encuestas (students' evaluation of teaching, SET) se basan en que, según sus autores, no son una buena referencia para evaluar la calidad docente, ya que contienen sesgos inasumibles e imposibles de eliminar. Por esa razón, antes

de comenzar con el estudio estadístico se ha realizado una revisión bibliográfica sobre más de 80 referencias que nos facilite un conocimiento global de cuáles son los argumentos a favor y en contra de esta herramienta de evaluación.

La revisión bibliográfica pone de manifiesto que el debate sobre la validez de las encuestas como evaluadores de la calidad docente sigue abierto. Se ha observado que los trabajos sobre las encuestas empezaron a publicarse hace aproximadamente un siglo, pero todavía hay bastantes autores que reclaman más investigación sobre este tema, aunque no parece haber acuerdo respecto a cuáles serían los aspectos en los que resulta necesario profundizar. El listado que se incluye a continuación es un resumen de los aspectos más controvertidos de la evaluación docente, de acuerdo con la revisión bibliográfica. Dado que para muchos de ellos existen estudios recientes a favor y en contra, podemos concluir que no están definitivamente resueltos:

- Las variables más citadas cuando se habla de sesgo: carga de trabajo, tamaño de grupo, indulgencia en las calificaciones, género del profesor, entre otras [3-9].
- La capacidad de los estudiantes para evaluar la actividad docente [10-12].
- El debate sobre el uso de las SET: ¿Es interesante? ¿Está promovido por profesores y gestores a los que no les conviene que se hagan las encuestas? [13-14].
- El empleo de otros parámetros estadísticos distintos de la media aritmética (por ejemplo, la mediana) en los resultados de las encuestas SET [15-17].
- La evaluación del profesor frente a la de la docencia de calidad: ¿Son diferentes?

¿Cómo puede evaluarse la docencia de calidad? ¿Cuál es su relación con los resultados del aprendizaje y de estos con los resultados académicos? [18-22].

- La personalidad del profesor, frente a su popularidad, y la posible influencia de ambas en los resultados de las SET [23-25].
- El efecto de los resultados de las SET en la docencia: ¿Ayudan a mejorarla? ¿Influyen en la forma en que el profesorado enfoca cada asignatura? [26, 27].
- La evaluación docente formativa y la evaluación docente sumativa, especialmente la segunda [28, 29].
- La repetitividad de los resultados de las encuestas: ¿Se repiten las valoraciones de un mismo profesor en distintos grupos y cursos o hay mucha variabilidad? [30, 31].
- La multidimensionalidad de la calidad docente: ¿Puede medirse la calidad con una sola variable? [32, 33].
- El diseño y el uso de los cuestionarios: ¿Qué parámetros deben tenerse en cuenta? ¿Qué variables hay que introducir en ellos? ¿Es válido utilizar el mismo cuestionario para distintos niveles, disciplinas, etc. o debe adaptarse a cada caso? [34-36].
- La conveniencia de introducir una variable de control, como es la asistencia personal a clase, que, siempre suponiendo que el individuo que responde la encuesta es veraz, resulta ser un elemento clave para considerar la bondad y validez de la consulta [37].
- La forma y momento de pasar las encuestas, sus ventajas e inconvenientes: presenciales u online, antes o después de los exámenes finales, nominales o anónimas, para

todos los estudiantes o solo para los que hayan superado la asignatura, etc. [38-40].

- Las encuestas de satisfacción de los estudiantes y su aportación al modelo Docenteia [41-43].

- Los resultados de las encuestas en las carreras técnicas comparados con los de otras ramas, como la de Humanidades [44-46].

- La interpretación de los resultados de las SET: ¿Son capaces los profesores y los gestores universitarios de interpretar correctamente las valoraciones y de usarlas adecuadamente? [47-50].

- Las asignaturas obligatorias frente a las optativas: ¿Hay diferencia entre los resultados de las encuestas? ¿Es lógico que ocurra así? [46].

Con este trabajo se intenta dar respuesta a algunas de estas dudas, que siguen estando presentes cuando se analizan los resultados de las encuestas.

2. OBJETIVOS

Los objetivos de este trabajo se centran en varios aspectos:

En primer lugar, se pretende cuantificar las posibles relaciones entre los valores de diversos indicadores de calidad (satisfacción con la actividad docente, carga de trabajo percibida, resultados académicos, características de los grupos docentes y del profesorado, etc.), además de contribuir a mejorar los procesos de gestión de la información de la UPCT desde la perspectiva de los procedimientos internos y externos de aseguramiento de la calidad.

En segundo lugar, se va a incluir en el análisis una componente temporal, que hasta ahora no había sido posible, ya que disponemos de datos completos de tres cursos académicos consecutivos.

En tercer lugar, se pretende contribuir a aumentar la confianza de todos los estamentos de la Universidad Politécnica de Cartagena (UPCT) en los indicadores de calidad.

Por último, se intenta aportar criterios de utilidad para los procesos que trabajen con esta información: renovación de la acreditación de los títulos, evaluación del profesorado, etc.

3. METODOLOGÍA

3.1. Obtención de datos

Entre los datos que se recogen cada curso académico en la UPCT, y que gestiona la Oficina de Prospección y Análisis de Datos (OPADA), se encuentran las encuestas de satisfacción de los estudiantes con la actividad docente, que valoran distintos aspectos de la docencia. Además, se cuantifica también la carga de trabajo percibida por sus estudiantes, mediante una encuesta ad hoc, con el objetivo de mejorar la coordinación horizontal.

La base de datos de la OPADA también incluye los resultados académicos de todas las asignaturas impartidas (tasas de presentados, éxito y rendimiento, calificación media, etc.), la información sobre las asignaturas (número de matriculados, tipo, etc.) y las características del profesorado que las imparte. La información empleada en el estudio corresponde al curso 2017-18, si bien se han hecho algunas comparativas con los datos que se obtuvieron en los cursos 2016-17 y 2015-16. El hecho de disponer de datos de tres cursos consecutivos permite analizar la evolución temporal de algunas de las variables.

Encuesta de satisfacción de la actividad docente.

El cuestionario que emplea la UPCT para la encuesta de satisfacción con la actividad docente consta de 11 ítems. Los estudiantes valoran cada uno de los aspectos de la actividad docente mediante una escala Likert de 1 a 5 (siendo 1 la valoración más desfavorable y 5 la más favorable). En los informes que recogen los resultados de las encuestas se indica la media aritmética de las valoraciones en cada uno de los ítems de la encuesta, la muestra y la desviación típica. También se indican, para que sirvan de referencia, los valores medios del curso en el que se sitúa la asignatura, del título y del departamento al que pertenece el profesor, entre otros. Como suele ser habitual en este tipo de encuestas, el último ítem sirve para valorar globalmente la labor docente del profesor. No existe en la UPCT un modelo de evaluación de la actividad docente tipo DOCENTIA, por lo que la encuesta es la única información disponible actualmente y su uso es casi exclusivamente formativo.

Los datos correspondientes al curso 2017-18 recopilan las 1.566 encuestas de satisfacción realizadas, incluyendo las medias aritméticas de los valores por cada uno de los 11 ítems de cada encuesta, el número de estudiantes que las respondieron y las características del profesorado evaluado. El número es similar al del curso anterior, 2016-17, en el que se recogieron 1.561 encuestas.

Carga de trabajo percibida

La carga de trabajo es, para muchos autores, uno de los aspectos del contexto académico que en mayor medida afectan al comportamiento académico del estudiante universitario. La carga de trabajo puede llegar a

condicionar el enfoque de aprendizaje adoptado por él, además de incidir de forma determinante en sus resultados académicos. Un exceso de carga de trabajo, o una distribución inadecuada de la misma, en una o varias asignaturas puede repercutir en las restantes y dificultar la organización de las actividades docentes y de evaluación del curso académico.

En la UPCT se empezó a cuantificar la carga de trabajo percibida durante el curso 2014-2015. Para ello se pasa una encuesta en todas las asignaturas obligatorias de títulos de Grado y de títulos de Máster habilitante. La encuesta no se pasa para cada profesor, como ocurre con la satisfacción con la actividad docente, sino para cada uno de los grupos docentes de cada asignatura. Se emplea una escala Likert de 1 (carga muy baja) a 5 (carga excesiva). El valor medio de la escala (3) corresponde a un grupo docente cuya carga de trabajo no presencial coincide sensiblemente con la que tiene asignada la asignatura en el plan de estudios. Los valores iguales o superiores a 4 se consideran anómalos (por carga excesiva) y también los iguales o inferiores a 2 (por carga escasa o insuficiente). El cuestionario es específico de cada cuatrimestre/curso/título, ya que recoge todas las asignaturas que el estudiante no repetidor cursa simultáneamente, y permite también que los estudiantes incluyan comentarios relativos a la carga de trabajo [51].

Base de datos general

La estructura de la base de datos empleada mayoritariamente en este estudio se adapta a la variable de referencia, que es la valoración obtenida por el profesorado en la encuesta de satisfacción con la actividad docente. Por consiguiente, cuenta con un registro por cada encuesta de satisfacción del curso 2017-18, es decir por cada profesor de cada grupo docente

impartido. En la base de datos figuran también los datos que identifican la asignatura (título, código, nombre de la asignatura, etc.) y, en su caso, cada uno de los grupos docentes que se formaron el cuatrimestre en el que se impartió cada asignatura y si es obligatoria u optativa.

Entre la información sobre el profesorado recogida en la base de datos está el género, la categoría académica, el departamento, el área de conocimiento y la antigüedad (desde la creación de la UPCT). También se dispone de datos sobre los resultados académicos de cada grupo docente de cada asignatura: calificación media obtenida por los estudiantes, tasas de presentados, éxito y rendimiento, etc. Si el grupo docente ha sido impartido por varios profesores, todos sus registros tendrán asociados los mismos datos de resultados académicos.

En la base de datos figuran algunos registros incompletos. Existen unos pocos casos en los que se dispone de la información sobre la carga de trabajo percibida, pero no llegó a pasarse la encuesta de satisfacción. Además, las asignaturas optativas y las de Máster no habilitante disponen de datos de satisfacción, pero no de carga de trabajo, por lo que los registros correspondientes tienen en blanco los campos correspondientes a esta variable. No se dispone de la información de algunos profesores no pertenecientes a la UPCT que impartieron docencia en centros adscritos. Finalmente, pueden faltar datos de resultados académicos en alguno de los registros.

En esta base de datos, buena parte de la información corresponde al profesor, mientras que otra parte de ella corresponde al grupo docente; en estos casos, se han tomado los mismos valores para todos los profesores que impartieron el grupo docente. La base de datos definitiva consta de 1.661 registros.

3.2. Principios de diseño

La fiabilidad es una característica de los resultados de una encuesta, y expresa el grado de confianza que se puede tener de las valoraciones del cuestionario. La fiabilidad se puede analizar mediante el coeficiente alfa de Cronbach, que indica la consistencia interna del cuestionario. El coeficiente se calcula con la siguiente expresión, a partir de las varianzas de los ítems del cuestionario:

$$\alpha = \left| \frac{k}{k-1} \right| \left| 1 - \frac{\sum_{i=1}^k S_i^2}{S_t^2} \right|$$

donde:

- K es el número de ítems del cuestionario (11 en nuestro caso)
- S_i^2 es la varianza del ítem i
- S_t^2 es la varianza de las sumas de los valores observados.

El valor del coeficiente de Cronbach para los datos del curso 2017-18, es de $\alpha = 0,978$. Cuanto más se aproxime α a la unidad, mayor es la fiabilidad de la escala. Suele considerarse que valores de α superiores a 0,7 o 0,8 son suficientes para garantizar la fiabilidad. El valor obtenido es muy próximo a la unidad, y es idéntico al obtenido con los datos de los cursos 2015-16 y 2016-17.

Se ha comprobado si el ítem 11 de la encuesta puede ser utilizado como valor representativo del conjunto de los restantes ítems, para poder realizar una simplificación de los resultados de la encuesta a efectos estadísticos. Para ello, se calculó el coeficiente de correlación entre los valores del ítem 11 y las medias aritméticas de los valores de los otros 10 ítems, obteniéndose un valor $\rho = 0,96$, muy próximo a la unidad. Por

tanto, entendemos que el ítem 11 se puede considerar como representativo de la satisfacción de los estudiantes con la actividad docente.

4. RESULTADOS

Antes de comenzar a analizar los resultados del trabajo se muestra la tabla 1, en la que aparecen todos los acrónimos relativos a las variables de la tabla de datos completa que se han utilizado en el estudio.

4.1. Estadística descriptiva de los valores de la base de datos

En esta primera fase se han empleado todos los registros disponibles de la base de datos, incluyendo los de todos los títulos que imparte la UPCT, pero teniendo en cuenta que algunos registros, como ya se ha indicado, estaban incompletos. Los valores de carga de trabajo (CT) que se muestran en la tabla 2 no tienen por qué coincidir con los obtenidos de la base de datos específica de la encuesta de CT, ya que en la base de datos general se han adaptado los datos de CT a los de satisfacción, asignando a todos los docentes de cada grupo docente el valor de CT de la encuesta, que es única para el grupo docente.

En la tabla 2 se observa que la calificación media (CM) obtenida por los estudiantes en las 1.651 asignaturas de las que se dispone de información es 6,27, muy parecida a la del curso anterior (6,20). Los valores medios de las tasas TP, TE y TR son relativamente altos. Se sabe que en títulos de Grado van mejorando con el curso, y que en títulos de Máster, incluso en los habilitantes, esos valores medios suelen ser superiores a los de los títulos de Grado. Los valores más bajos corresponden a asignaturas con muy pocos matriculados, en las que

ningún estudiante se presentó a examen o no hubo aprobados.

El valor medio de carga de trabajo percibida CT con esta base de datos completa es de 3,47, mientras que con la base de datos específica para CT es de 3,45. Ya hemos indicado que esto es debido a que en la base de datos completa están repetidos los valores de asignaturas/grupos impartidas por varios profesores. La diferencia, muy pequeña, entre los valores puede deberse a que la percepción de carga de trabajo tiende a ser mayor en asignaturas con más profesores.

La tabla 3 y la figura 1 muestran el histograma de frecuencias de la variable S (ítem 11). Para calcularlo se han empleado las 1.566 encuestas disponibles, sin tener en cuenta el tamaño de muestra de cada una de ellas.

Hay 222 encuestas de satisfacción con valores de S inferiores o iguales a 3. Eso supone un 14,17 % de las encuestas, mientras que en 2016-17 el porcentaje de valores bajos fue del 16,08 % y en 2015-16 fue del 14,11 %. El restante 85,83 % de las encuestas corresponde a valores superiores a 3, mientras que en el curso 2016-17 fueron un 83,92 % y en 2015-16 fueron un 85,89 %.

Se han estudiado los valores de satisfacción (S) por cursos, solo para las asignaturas de títulos de Grado. Se ha optado por hacer el cálculo con y sin asignaturas optativas, ya que estas se concentran en los cursos superiores y podrían afectar a los valores medios de esos cursos. Además, se ha tenido en cuenta que existen asignaturas definidas como optativas pero que realmente son asignaturas de especialidad, obligatorias para alguna de las menciones que se dan en determinados títulos. Los resultados se muestran en la tabla 4.

La satisfacción media es, efectivamente, mayor en las asignaturas optativas que en las obligatorias, como cabía esperar. Si consideramos todas las asignaturas, optativas incluidas, el curso con mayor valor medio de S es el de 4º, en el que se concentra la mayor

Tabla 1. Acrónimos empleados para las variables en estudio.

Acrónimo	Variable
N	Tamaño de grupo (número de matriculados en la asignatura/grupo docente)
S	Resultado del ítem 11 de la encuesta de satisfacción con la actividad docente
NS	Número de respuestas a la encuesta de satisfacción
CT	Carga de trabajo percibida por los estudiantes de la asignatura/grupo docente
NCT	Número de respuestas a la encuesta de carga de trabajo percibida
TP	Tasa de presentados de la asignatura/grupo
TE	Tasa de éxito de la asignatura/grupo
TR	Tasa de rendimiento de la asignatura/grupo
CM	Calificación media obtenida por los estudiantes de la asignatura
CPA	Número de convocatorias presentadas hasta aprobar la asignatura
MPA	Número de matrículas realizadas hasta aprobar
CP1	Convocatorias hasta que se presenta a examen por primera vez (o hasta la fecha)
MP1	Matrículas transcurridas hasta que el estudiante se presenta (o hasta la fecha)
P1M	Porcentaje de estudiantes de 1ª matrícula
P1MA	Porcentaje de estudiantes de 1ª matrícula aprobados
PSP	Porcentaje de suspensos sobre presentado

Tabla 2. Estadísticos elementales de algunas de las variables (datos 2017-18).

Variable	Media	Muestra	Varianza	Máximo	Mínimo
N	27,74	1.661	357,24	107	1
CPA	1,47	1.651	0,24	3,88	1,00
MPA	1,27	1.661	0,10	2,91	1,00
CP1	2,31	1.661	1,54	7,91	1,00
MP1	1,27	1.661	0,10	2,91	1,00
CM	6,27	1.651	1,86	10,00	2,99
TP	84,03	1.661	266,47	100,00	0,00
TE	84,63	1.651	298,79	100,00	26,67
TR	72,53	1.661	569,06	100,00	0,00
P1M	81,86	1.661	363,52	100,00	10,71
P1MA	73,37	1.661	632,36	100,00	0,00
PSP	15,37	1.651	298,79	73,33	0,00
NCT	15,71	1.297	95,19	65,00	1
CT	3,47	1.297	0,33	5,00	2,00

Tabla 3. Histograma de satisfacción S, ítem 11 (curso 2017-18).

	1,00-1,50	1,51-2,00	2,01-2,50	2,51-3,00	3,01-3,50	3,51-4,00	4,01-4,50	4,51-5,00	Total
Nº	6	25	50	141	256	463	396	229	1.566
%	0,38	1,60	3,19	9,00	16,35	29,57	25,29	14,62	100

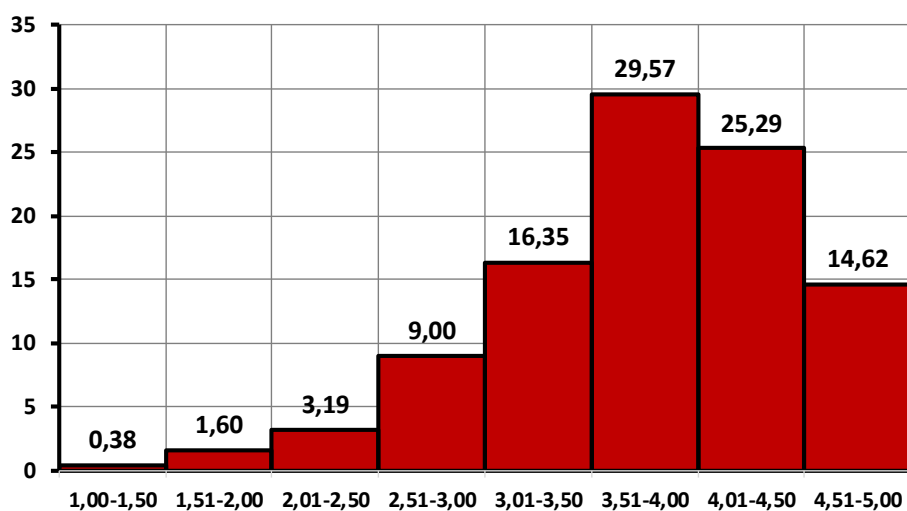


Fig. 1: Histograma de satisfacción S, ítem 11 (curso 2017-18).

Tabla 4. Valores de satisfacción S, ítem 11, según cursos y tipo de asignaturas (curso 2017-18). Solo títulos de Grado.

Curso	S (todas)	Nº datos	S (solo obligatorias)	Nº datos	S (obligatorias + especialidad)	Nº datos
1º	3,84	306	3,84	306	3,84	306
2º	3,68	376	3,68	376	3,68	376
3º	3,75	294	3,71	234	3,73	270
4º	3,90	256	3,61	125	3,73	154
5º (GFA)	3,57	19	3,57	19	3,57	19
Grados	3,78	1.251	3,72	1.060	3,74	1.125

Tabla 5. Histograma de carga de trabajo percibida, CT (curso 2017-18).

	1,00-1,50	1,51-2,00	2,01-2,50	2,51-3,00	3,01-3,50	3,51-4,00	4,01-4,50	4,51-5,00	Total	
Nº	0	6	29	210	282	245	108	39	919	
%	0	0,65	3,16	22,85	30,69	26,66	11,75	4,24	100	

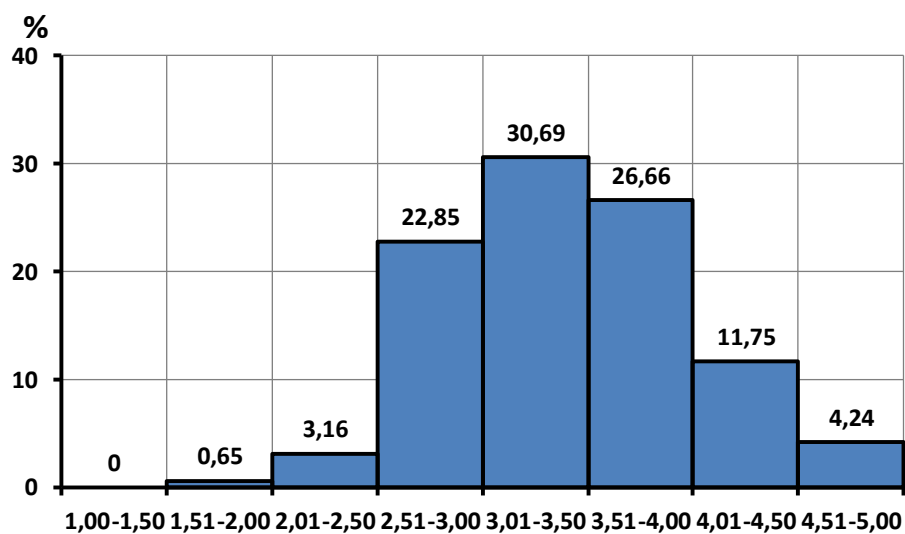


Fig. 2: Histograma de carga de trabajo percibida, CT (curso 2017-18).

Tabla 6. Carga de trabajo percibida CT por títulos y centros (curso 2017-18).

Centro	Título	Nº datos	Media CT	Varianza
Adscrito	GIOI	76	3,57	0,29
EICIM	GIC	36	3,40	0,31
	GIRME	36	3,28	0,33
	MUICCP	16	3,64	0,36
ETSAE	GFA	82	3,52	0,53
	GIDE	41	3,45	0,43
ETSIA	GIASB	51	3,46	0,24
	MUIA	13	3,43	0,26
ETSII	GIE	34	3,51	0,24
	GIEIA	57	3,42	0,26
	GIM	86	3,42	0,37
	GIQI	35	3,27	0,32
	GITI	66	3,43	0,44
	MUII	17	3,55	0,43
ETSINO	GANISM	34	3,52	0,31
	MUINO	16	3,67	0,46
ETSIT	GIST+GIT*	110	3,52	0,35
	MUIT	9	3,51	0,15
FCE	GADE	102	3,30	0,27
	GT**	2	3,57	0,14
UPCT	Grado+Máster	919	3,45	0,35

Tabla 7: Valores medios por cursos en asignaturas obligatorias en títulos de Grado y Máster habilitante (2017-2018).

Curso	S	CT	CM	TP	TE	TR	N
1º	3,78	3,44	5,71	77,87	77,77	62,01	34,93
2º	3,73	3,38	5,86	81,99	78,64	65,89	30,76
3º	3,73	3,58	6,19	85,32	86,15	74,47	32,15
4º	3,74	3,58	6,65	88,04	91,09	81,04	22,12
5º (GFA)	3,57	3,83	6,41	77,62	85,09	66,90	3,24

optatividad en todos los títulos de Grado de la UPCT. No obstante, si solo se consideran las asignaturas obligatorias y las optativas de especialidad, el valor medio de 4º curso baja apreciablemente hasta equipararse al de 3º.

La carga de trabajo percibida (CT) por los estudiantes se ha analizado a partir de la base de datos específica de CT. La base de datos de CT contiene 919 registros de asignaturas obligatorias, de los que 848 corresponden a títulos de Grado y 71 corresponden a títulos de Máster habilitantes. La encuesta de CT, como se ha indicado anteriormente, no se pasa en asignaturas optativas ni en títulos de Máster no habilitantes. La tabla 5 y la figura 2, muestran el histograma de esta variable.

El valor medio de la carga de trabajo percibida CT en el curso 2017-18 fue 3,45, en una escala Likert 1-5, siendo la varianza 0,35. El intervalo más frecuente es el de carga normal/carga ligeramente elevada (3,01-3,5), seguido de los intervalos siguiente y anterior. No hubo asignaturas con valores de CT inferiores a 1,5 y solo hubo 35 asignaturas con valores inferiores a 2,5. Los valores de CT superiores a 4 se consideran elevados y los superiores a 4,5 representan cargas de trabajo que, en opinión de los estudiantes, resultan excesivas; se han dado 147 encuestas con valores superiores a 4, de las que 39 superan el valor de 4,5.

Los valores de CT por títulos y centros de la UPCT para el curso académico 2017-18 se muestran en la tabla 6. Se han destacado los datos de la Escuela Técnica Superior de Arquitectura y Edificación (ETSAE).

Todos los títulos presentan valores de CT superiores al valor medio de la escala Likert (3), y los valores más elevados corresponden a los títulos de Máster, todos ellos habilitantes. En el

caso de la ETSAE se observan valores de CT similares a los de otros títulos de Grado de la UPCT para el Grado en Ingeniería de Edificación (GIDE) y algo más altos para el Grado en Fundamentos de Arquitectura (GFA). Esto puede deberse a la considerable carga práctica que tiene el GFA y es una situación que viene observándose desde que se empezó a cuantificar esta variable. No se recoge en la tabla 6 ningún título de Máster de la ETSAE, ya que la encuesta de CT solo se pasa, como se ha indicado, en los que son habilitantes y durante el curso analizado la ETSAE no ofertó ningún título de este tipo.

En la tabla se muestra la evolución de los valores medios de CT de toda la UPCT entre los cursos 2014-15 y 2017-18. La carga de trabajo percibida va aumentando con el tiempo, pero no es fácil deducir cuáles pueden ser las causas de esta continua subida. Comparando los resultados con los del curso 2016-17 se observa que la subida se ha producido, en media, tanto en los títulos de Grado como en los de Máster. Sin embargo, por títulos, se observa que algunos han subido y otros han bajado, a veces de forma sustancial. Por lo demás, se mantienen las tendencias observadas con los datos de cursos anteriores: los títulos de Máster (habilitante) presentan mayor CT media que los de Grado; los cursos superiores de los títulos de Grado presentan mayor CT media; no se aprecian diferencias entre cuatrimestres, pero sí entre las asignaturas cuatrimestrales y las anuales.

Para completar el estudio estadístico básico se van a relacionar los valores medios de número de matriculados (N), calificación media (CM), tasa de presentados (TP), de éxito (TR) y de rendimiento por cursos, con S y con CT. Los datos que se muestran en la tabla 7 se han calculado solo para asignaturas obligatorias, y

en esta ocasión sí se han incluido los datos de títulos de Máster. Se observa que la satisfacción S mantiene un valor muy estable en los cuatro primeros cursos; el quinto curso solo existe en el Grado en Fundamentos de Arquitectura y no es significativo. Los indicadores de resultados académicos (calificación media CM y tasas de presentados, éxito y rendimiento) van creciendo con el curso, salvo para quinto.

En el caso de la ETSAE se muestran los mismos valores por cursos en la tabla 8, pero en este caso solo con los títulos GFA y GIDE.

La satisfacción (S) en los títulos de Grado de la ETSAE se encuentra por debajo de la media en los cursos de 2º a 4º, mientras que la CT se mantiene en los valores de los globales de la UPCT, o es incluso un poco superior. Los valores más dispares respecto a los globales de la UPCT se encuentran en las tasas de presentados, éxito y rendimiento. Además, el número de matriculados a partir de 2º es mucho menor que la media.

4.2. Análisis de correlación entre variables

Los coeficientes de correlación entre algunas de las variables numéricas recogidas en la base de datos se muestran en la tabla 9.

La satisfacción S y la carga de trabajo percibida CT muestran una correlación muy débil ($\rho = -0,16$) pero que tiene el signo que cabía esperar. Los coeficientes de correlación de S y CT con las restantes variables consideradas presentan valores bajos o muy bajos, aunque hay que destacar que el signo de ρ siempre es el esperado. En todos los casos, si la correlación es positiva con S, será negativa con CT y viceversa. El mayor valor del coeficiente de correlación ρ para S es con la calificación media CM; aun así el valor es bajo (0,20).

Las variables que representan tasas y resultados académicos se correlacionan poco, como hemos dicho, con S y CT pero, como es lógico, presentan entre sí valores moderados o fuertes de ρ , siempre con los signos esperados.

La tabla 9 muestra también los coeficientes de correlación correspondientes a los cursos 2016-17 y 2015-16. Los valores del coeficiente de correlación son parecidos: valores débiles de S y CT con las variables de resultados académicos y moderados o fuertes de estas variables entre sí. Los signos de todos los coeficientes coinciden en los tres cursos. El coeficiente de correlación entre S y CT obtenido con los datos de 2014-15 fue de -0,11.

Las figuras 3 a 6 muestran gráficamente lo que se deduce de los valores de los correspondientes coeficientes de correlación: gráficos de dispersión con formas poco definidas y sin una tendencia clara.

Se han elegido para las figuras las variables CT, CM, N y TE porque están entre las que algunos autores citan como posibles causas de sesgo en las encuestas de satisfacción con la actividad docente. Los valores de los coeficientes de correlación son débiles o muy débiles, pero significativos, lo que indica que existe cierta influencia, por pequeña que sea.

Por otra parte, el coeficiente de correlación entre la antigüedad de cada profesor, contada desde la creación de la UPCT, y la satisfacción es -0,05, casi nulo y parecido al obtenido en el curso 2016-17 (-0,06). Llama la atención el signo negativo de este valor, que indicaría que la satisfacción tiende a disminuir cuando aumenta la antigüedad y, por tanto, la experiencia. No obstante, la correlación es muy pequeña y, además, es posible que el valor se vea afectado por el hecho de que los datos de antigüedad

Tabla 8: Valores medios por cursos en asignaturas obligatorias en títulos de Grado de la ETSAE (2017-2018).

Curso	S	CT	CM	TP	TE	TR	N
1º	3,88	3,57	5,05	70,71	68,84	48,74	36,71
2º	3,71	3,31	5,60	81,88	82,77	68,18	22,74
3º	3,65	3,65	5,76	78,43	82,74	65,72	22,22
4º	3,64	3,76	6,32	78,83	87,65	70,70	10,31
5º (GFA)	3,57	3,83	6,41	77,62	85,09	66,90	3,24

Tabla 9. Coeficientes de correlación entre algunas de las variables (curso 2017-18).

2017-18	Con S	Con CT	Con CM	Con TE	Con TR
CT	-0,16	1	-0,05	-0,03	-0,09
N	-0,15	0,10	-0,47	-0,45	-0,39
CPA	-0,14	0,07	-0,72	-0,70	-0,69
MPA	-0,15	0,15	-0,60	-0,57	-0,65
CP1	-0,17	0,16	-0,71	-0,68	-0,77
MP1	-0,15	0,15	-0,60	-0,57	-0,65
CM	0,20	-0,05	1	0,84	0,81
TP	0,15	-0,14	0,57	0,55	0,87
TE	0,14	-0,03	0,84	1	0,88
TR	0,16	-0,09	0,81	0,88	1

2016-17	Con S	Con CT	Con CM	Con TE	Con TR
CT	-0,12	1	-0,11	-0,10	-0,13
N	-0,18	0,06	-0,51	-0,44	-0,44
CM	0,22	-0,11	1	0,83	0,79
TP	0,15	-0,15	0,52	0,55	0,85
TE	0,18	-0,10	0,83	1	0,89
TR	0,19	-0,13	0,79	0,89	1

2015-16	Con S	Con CT	Con CM	Con TE	Con TR
CT	-0,15	1	-0,14	-0,16	-0,17
N	-0,15	0,11	-0,41	-0,35	-0,43
CM	0,22	-0,14	1	0,83	0,81
TP	0,18	-0,13	0,53	0,49	0,81
TE	0,14	-0,16	0,83	1	0,90
TR	0,17	-0,17	0,81	0,90	1

están truncados, ya que solo se cuentan desde la fundación de la UPCT.

4.3. Sesgo de Género

El análisis de la posible existencia de un sesgo de género en las encuestas de satisfacción con la actividad docente, que se realizó en su momento con datos de los cursos 2015-16 [52] y 2016-17 [53], se ha completado con los del curso 2017-18, lo que ha permitido incorporar una perspectiva temporal en el estudio. Se dispone de 1.566 encuestas de satisfacción del curso 2017-18. No todas disponen de datos sobre el profesorado, ya que algunas de las de Centros adscritos corresponden a asignaturas no impartidas por profesorado UPCT.

Si los eliminamos, quedan 1.420 registros, de los que 357 (25,14 %) corresponden a encuestas de profesoras y 1.063 a encuestas de profesores varones. En la tabla 10 se muestran los datos del profesorado agrupado por categorías, siendo el término "Resto LOU" el que engloba a los profesores Ayudantes, Ayudantes Doctores y Contratados Doctores, mientras que "Resto" incluye las categorías anteriores a la LOU, como Titular de Escuela Universitaria, Colaboradores, Docentes de sustitución, Titulado Superior y otras.

Como ya se observó en los estudios de 2016-17 y 2015-16, sigue existiendo un mayor número de profesores varones que de profesoras en la UPCT, y existen diferencias en la consolidación de puestos de Catedrático, aunque no en Titulares de Universidad. En la ETSAE se dan, por el momento, diferencias notables tanto en la categoría de Catedráticos como en la de Titulares de Universidad (tabla 11).

Respecto al género del alumnado, se comprueba que la diferencia de porcentajes de alumnas entre unos centros de la UPCT y otros son muy grandes, siendo la ETSAE el que tiene mayor porcentaje de matriculadas en Grado, con un 46,21 %, seguido por la Facultad de Ciencias de la Empresa (FCE). El número de alumnas en el resto de centros es bastante inferior. La ETSAE es también el centro con mayor número de matriculadas en Máster.

Conocidos estos datos se estudia la posibilidad de que exista sesgo de género en las encuestas de satisfacción, por discriminación del género del profesorado o por el género del alumnado, como se indicaba en la revisión bibliográfica. Analizados los valores medios de satisfacción obtenidos de las 1.063 encuestas al profesorado masculino y de las 357 al profesorado femenino se obtuvieron valores medios del ítem 11 de la encuesta de 3,81 y 3,80, respectivamente, mientras que las medias de los ítems 1 a 10 fueron de 3,68 y 3,72, respectivamente. Las diferencias son mínimas y, además, de distinto signo. Valores similares se repiten haciendo el estudio por cursos y tipos de asignaturas.

Respecto a la ETSAE, los valores obtenidos son mejores para las profesoras: los profesores varones obtienen un 3,72 como media del ítem 11, frente a un 3,87 de las profesoras, y un 3,61 en la media de los ítems 1 a 10 frente a un 3,79 de las profesoras.

En la figura 7 se muestran las diferencias del valor de S entre profesoras y profesores varones en función del porcentaje de alumnas en cada centro. Se aprecia que un porcentaje mayor de alumnas en un centro no implica que la valoración obtenida por las profesoras sea mejor que la de los profesores varones en ese

mismo centro. De hecho, el centro con mayor porcentaje de alumnas, la ETSAE, es el que obtiene mejor valoración para sus profesoras,

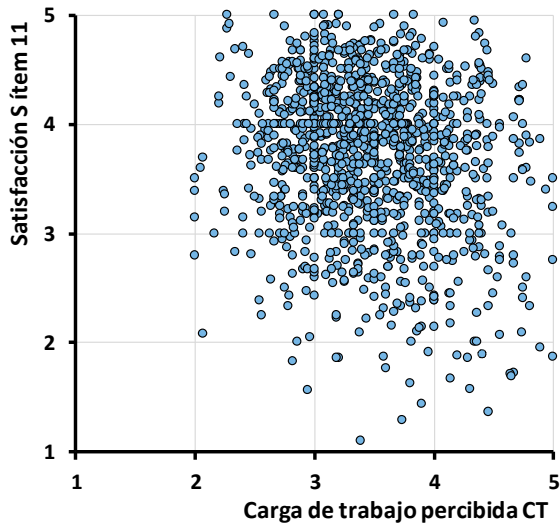


Fig. 3. Gráfico de dispersión entre S y CT

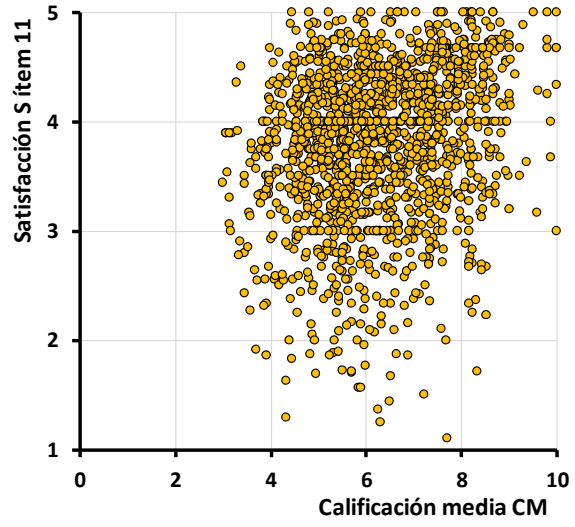


Fig. 4. Gráfico de dispersión entre S y CM

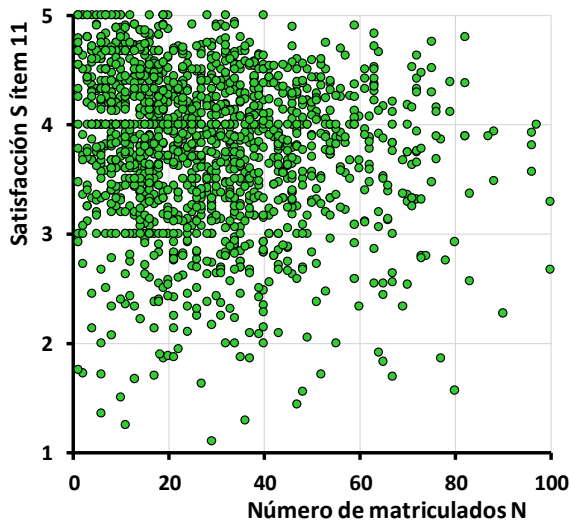


Fig. 5. Gráfico de dispersión entre S y N

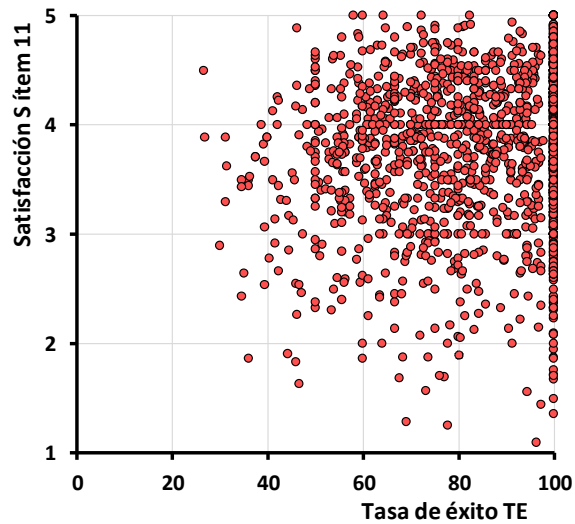


Fig. 6. Gráfico de dispersión entre S y TE

Tabla 10: Porcentaje y número de profesores por categorías agrupadas por géneros UPCT.

Categoría	Mujeres %	Varones %
Catedrático Universidad	5,32	12,51
Titular Universidad	34,45	32,74
Asociado	12,89	16,27
Resto LOU (AYU+AYUDOC+PCDOC)	23,53	16,84
Resto	23,81	21,64
	357 encuestas profesoras	1.063 encuestas profesores

Tabla 11: Porcentaje y número de profesores por categorías agrupadas por géneros ETSAE.

Categoría	Mujeres %	Varones %
Catedrático Universidad	0,00	2,96
Titular Universidad	8,00	14,07
Asociado	40,0	31,11
Resto LOU (AYU+AYUDOC+PCDOC)	36,0	35,56
Resto	16,0	16,30
	50 encuestas profesoras	145 encuestas profesores

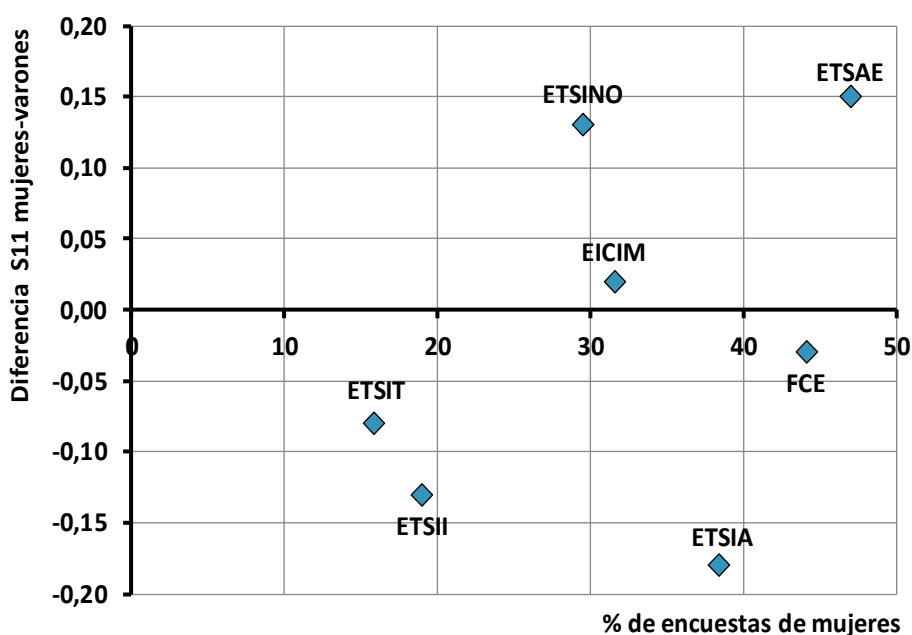


Fig. 7: Relación entre la diferencia existente entre los resultados del ítem 11 de profesores y profesoras y el porcentaje de alumnas por Centros.

pero no ocurre lo mismo con los dos siguientes (FCE y ETSIA), en los que los profesores varones obtienen mayor valoración. Hay que tener en cuenta, además, que las diferencias pueden cambiar de signo de un curso a otro, tal y como se ha observado al comparar los datos con los de 2016-17 y 2015-16, mientras que los porcentajes de alumnas son relativamente estables de un curso a otro.

En consecuencia, no puede establecerse ninguna tendencia temporal y parece claro que las pequeñas fluctuaciones observadas en los resultados no dependen del género del profesorado evaluado. Las encuestas de satisfacción con la actividad docente de la UPCT no muestran, por tanto, que exista un sesgo de género, a pesar del porcentaje reducido de alumnas y profesoras en muchos de sus centros.

El hecho de que las encuestas sean anónimas, y no indiquen el género de los estudiantes que las rellenan, impide profundizar más en el análisis de ciertos aspectos del sesgo de género. Por lo demás, y a la vista de sus resultados, el estudio se alinea con los de los autores que defienden que no existe sesgo de género en las SET [10, 21, 46].

5. CONCLUSIONES

Del estudio bibliográfico se concluye que el debate sobre las evaluaciones docentes y su validez sigue abierto; muchas de las dudas planteadas hace décadas no se han resuelto aún, en buena medida porque siguen apareciendo estudios con resultados contradictorios sobre muchos de los aspectos relevantes relacionados con este tema.

La correlación entre la satisfacción con la actividad docente y las variables más

significativas de nuestra base de datos es débil o muy débil, aunque el signo suele ser el esperado. Así, variables como la carga de trabajo estimada, la dificultad de la asignatura o la calificación obtenida, influyen poco o muy poco en los resultados de las encuestas de satisfacción. Por tanto las variables más citadas como sesgos en la literatura sobre las SET tienen, realmente, una influencia escasa o nula sobre las encuestas. En particular, no se aprecia sesgo por género.

La principal conclusión de nuestro estudio es que los resultados de las encuestas de satisfacción no son una herramienta perfecta, pero sí son suficientemente confiables. Por lo tanto, consideramos que constituyen un buen indicador para medir la calidad docente, al que habría que dar un peso importante en cualquier modelo de evaluación tipo Docentia.

Se puede afirmar que el cuestionario empleado en la UPCT para la encuesta de satisfacción con la actividad docente es bastante fiable. No obstante, la información que ofrece actualmente podría mejorarse, completándola: por ejemplo, incluyendo la mediana, además de la media aritmética, en cada uno de los 11 ítems.

Hemos llegado a la convicción de que es importante enseñar, a gestores y profesorado, a interpretar las encuestas y a utilizarlas para mejorar la docencia.

Por otra parte, los resultados académicos de cualquier asignatura dependen mucho de cómo funcionen las otras asignaturas del curso, de que no excedan la carga de trabajo establecida en la memoria verificada y de que exista una coordinación horizontal efectiva.

Agradecimientos

A la Oficina de Prospección y Análisis de Datos (OPADA) y al Portal de Transparencia de la Universidad Politécnica de Cartagena, por ceder los datos que han sido necesarios para la realización de este estudio. Al Servicio de Gestión de la Calidad de la UPCT por su colaboración.

REFERENCIAS

- [1] ANECA, "Guía de apoyo para la elaboración de la memoria de verificación de títulos oficiales universitarios (Grado y Máster)", pp. 20-24, 2015.
- [2] López-Aguado, M. "La evaluación de la calidad de títulos universitarios. Dificultades percibidas por los responsables de los sistemas de garantía de calidad", *Educación XX1*, 21(1), 263-284,
- [3] García Martín, A. y García-León, J. "Una experiencia de medición de la carga de trabajo percibida por los estudiantes para facilitar la coordinación horizontal", *REDU, Revista de Docencia Universitaria*, 15(1), pp. 81-104, 2017.
- [4] Boring, A. "Gender biases in student evaluations of teaching" *Journal of Public Economics*, 145, pp. 27-41, 2017.
- [5] Vázquez Arenas, Gemma et al. "¿Existe sesgo de género en la valoración de la actividad docente de títulos de Edificación? = Is there a gender bias in the valuation of the teaching activity in Building Engineering degrees?", *Advances in Building Education*, v. 2, no. 3, pp. 80-93, 2018.
- [6] Wachtel, H.K. "Student evaluation of college teaching effectiveness: a brief review". *Assessment & Evaluation in Higher Education*, 23(2), pp. 191-212, 1998.
- [7] Nasser, F. y Fresko, B. "Faculty views of student evaluation of college teaching", *Assessment & Evaluation in Higher Education*, 27(2), pp. 187-198, 2002.
- [8] Merritt, D.J. "Bias, the brain, and student evaluations of teaching", *St. John's Law Review*, 82, pp. 235-287, 2008.
- [9] MacNell, L., Driscoll, A., y Hunt, A.N. "What's in a name: exposing gender bias in student ratings of teaching" *Journal of Collective Bargaining in the Academy*, Article 52, 2015.
- [10] Aleamoni, L.M. "Student rating myths versus research facts from 1924 to 1998", *Journal of Personnel Evaluation in Education*, 13(2), pp. 153-166, 1999.
- [11] Benton, S.L. y Cashin, W.E. "Student ratings of instruction in college and university courses". En M.B. Paulsen (ed.), *Higher Education: Handbook of Theory and Research*, 29, pp. 279-326, 2013.
- [12] Culver, S. "Course grades, quality of student engagement, and students' evaluation of instructor", *International Journal of Teaching and Learning in Higher Education*, 22(3), pp. 331-336, 2010.
- [13] Tolosa, P. "Evaluación de la calidad docente". *Academia. Revista sobre enseñanza del Derecho*, 7, pp. 411-418, 2006.
- [14] De Carvalho Andrade, E. y De Paula Rocha, B. "Factors affecting the student evaluation of teaching scores: Evidence from panel data estimation". *Estudos Econômicos*, 42(1), pp. 129-150. 2012.
- [15] Apodaca, P. y Grad, H. "The dimensionality of student ratings of teaching: integration of uni and multidimensional models",

- Studies in Higher Education, 30(6), pp. 723–748, 2005.
- [16] Shook, B. y Greer, M. "The misanalysis, misinterpretation, and misuse of student end-of-course evaluation data", *National Social Science Journal*, 44(2), pp.89-97, 2015.
- [17] Dorta-González, P. y Dorta-González, M.I. "The student evaluation of teaching and the competence of students as evaluators". *Anales de ASEPUMA* nº 21, 4012 p., 2013.
- [18] Barth, M.M. "Deciphering student evaluations of teaching: A factor analysis approach". *Journal of Education for Business*, 84(1), pp. 40-46, 2008.
- [19] Pounder, J. "Is student evaluation of teaching worthwhile? An analytical framework for answering the question". *Quality Assurance in Education*. 15(2), pp. 178-191, 2007.
- [20] Molero, D. y Ruiz, J. "La evaluación de la docencia universitaria. Dimensiones y variables más relevantes", *Revista de Investigación Educativa*, 23(1), pp. 57-84, 2005.
- [21] Marsh, H. y Roche, L. "Making students' evaluations of teaching effectiveness effective", *American Psychologist*, 52(11), pp. 1187-1197, 1997.
- [22] Uttl, B., White, C.A. y Gonzalez, D.W. "Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related", *Studies in Educational Evaluation* 54, pp. 22-42, 2017.
- [23] Abadía, A.R., Bueno, C., Ubieto-Artur, M.I., Márquez, M.D., Sabaté, S., Jorba, H. y Pagès, T. "Competencias del buen docente universitario. Opinión de los estudiantes", *REDU-Revista de Docencia Universitaria*, 13(2), pp. 363-390, 2015.
- [24] Zabaleta, F. "The use and misuse of student evaluations of teaching", *Teaching in Higher Education*, 12(1), pp. 55-76, 2007.
- [25] Kim, C., Damewood, E. y Hodge, N. " Professor attitude: its effect on teaching evaluations". *Journal of Management Education*, 24(4), pp. 458-473, 2000.
- [26] Álvarez, V., García, E. y Gil, J. " La calidad de la enseñanza universitaria desde la perspectiva de los profesores mejor valorados por los alumnos", *Revista de Educación*, 319, pp. 273-290, 1999.
- [27] Marsh, H. y Roche, L. " Effects of grading leniency and low workload on students' evaluations of teaching: popular myth, bias, validity, or innocent bystanders?", *Journal of Educational Psychology*, 92(1), pp. 202-228, 2000.
- [28] Berk, R.A. "Should student outcomes be used to evaluate teaching?", *Journal of Faculty Development*, 28(2), pp. 87-96, 2014.
- [29] Marsh, H. " Students' evaluations of university teaching: dimensionality, reliability, validity, potential biases and usefulness"., en R.P. Perry and J.C. Smart (eds.), *The Scholarship of Teaching and Learning in Higher Education: An Evidence-Based Perspective*, pp. 319-383, 2007.
- [30] Murray, H.G. "Student evaluation of teaching: has it made a difference?" Paper presented at the Annual Meeting of the Society for Teaching and Learning in Higher Education, Charlottetown, Prince Edward Island, 2005.
- [31] Boring, A., Ottoboni, K., y Stark, P. " Student evaluations of teaching (mostly) do not

measure teaching effectiveness". *Science Open Research*, 2016.

[32] Penny, A.R. " Changing the agenda for research into students' views about university teaching: Four shortcomings of SRT research", *Teaching in Higher Education*, 8(3), pp. 399-411, 2003.

[33] Luna, E. y Torquemada, A. " Los cuestionarios de evaluación de la docencia por los alumnos: balance y perspectivas de su agenda", *Revista Electrónica de Investigación Educativa*, número especial, 2008.

[34] Abrami, P.C., d'Apollonia, S. y Rosenfield, S. "The dimensionality of student ratings of instruction: What we know and what we do not", en R.P. Perry and J.C. Smart (eds.), *The Scholarship of Teaching and Learning in Higher Education: An Evidence-Based Perspective*, pp. 385-456, 2007.

[35] Kramp, J. "The Effect of Workload on Student Evaluations of Teaching", ETD Archive. Paper 455. Cleveland State University, 2010.

[36] Hajdin, G. y Pažur, K, " Differentiating between student evaluation of teacher and teaching effectiveness", *Journal of Information and Organizational Sciences*, 36(2), pp. 123-134, 2012.

[37] Wolbring, T. y Treischl, E. "Selection bias in students' evaluation of teaching. Causes of student absenteeism and its consequences for course ratings and rankings", *Research in Higher Education*, 57, pp. 51–71, 2016.

[38] Arnold, I.J.M. "Do examinations influence student evaluations?", *International Journal of Educational Research*, 48, pp. 215–224, 2009.

[39] Spooren, P. y Van Loon, F. "Who participates (not)? A non-response analysis on

students' evaluations of teaching", *Procedia - Social and Behavioral Sciences*, 69, pp. 990-996, 2012.

[40] Carle, A.C. "Evaluating college students' evaluations of a professor's teaching effectiveness across time and instruction mode (online vs. face-to-face) using a multilevel growth modeling approach", *Computers & Education*, 53, pp. 429–435, 2009.

[41] Sánchez, F., Rubio, R., Alonso, E. y Retamal, K. " La valoración de la actividad docente. Algo más que la opinión de los estudiantes", *Boletín de Psicología*, 97, pp. 71-92, 2009.

[42] Mas, O. " Las competencias del docente universitario: la percepción del alumno, de los expertos y del propio protagonista", *REDU Revista de docencia Universitaria*, 10 (2), pp. 299-318, 2012.

[43] Berk, R.A. "Survey of 12 strategies to measure teaching effectiveness", *International Journal of Teaching and Learning in Higher Education*, 17(1), pp. 48-62, 2005.

[44] Clayson, D.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, pp. 16–30, 2009.

[45] Fernández Rico, J. E., Fernández Fernández, S., Álvarez Suárez, A. y Martínez Cambor, P. " Éxito Académico y satisfacción de estudiantes con la enseñanza universitaria" *RELIEVE, Revista Electrónica de Investigación y Evaluación Educativa*. 13(2), pp. 203-214, 2007.

[46] Centra, J. A. "Differences in responses to the student instructional report: Is it bias?", Princeton: Educational Testing Service, 2009.

- [47] Tran, N.D. " Reconceptualisation of approaches to teaching evaluation in higher education", *Issues in Educational Research*, 25(1), pp. 50-61, 2015.
- [48] Benton, S.L. y Li, D. "Judging the credibility of quantitative student ratings of instruction research" *IDEA Paper*. Nº. 80, 2019.
- [49] Boysen, G.A. "Statistical knowledge and the over-interpretation of student evaluations of teaching", *Assessment & Evaluation in Higher Education*, 42(7), pp. 1095-1102, 2017.
- [50] Berk, R.A. "Should student outcomes be used to evaluate teaching?", *Journal of Faculty Development*, 28(2), pp. 87-96, 2014.
- [51] García Martín, A. y García-León, J. " Una experiencia de medición de la carga de trabajo percibida por los estudiantes para facilitar la coordinación horizontal", *REDU, Revista de Docencia Universitaria*, 15(1), pp. 81-104, 2017.
- [52] García Martín, A., Conesa, J.F. y Blancafort, J., coordinadores, *La innovación docente aplicada a la calidad de los títulos universitarios en la UPCT (curso 2017-18)*. CRAI UPCT Ediciones, pp. 73-106, 2018.
- [53] García Martín, A. y Conesa Pastor, J.F., coordinadores, *Acciones para la mejora de la calidad de los títulos de la UPCT: proyecto de innovación docente (curso 2018-19)*. CRAI UPCT Ediciones, pp. 35-108, 2019.