

García-Olalla, A., Villa Sánchez, A., Aláez, M. y Romero-Yesa, S. (2022). Aplicación y resultados de un sistema para evaluar la calidad de la docencia universitaria en una década de experimentación. *Revista de Investigación Educativa*, 40(1), 51-68.

DOI: <http://dx.doi.org/10.6018/rie.401221>

Aplicación y resultados de un sistema para evaluar la calidad de la docencia universitaria en una década de experimentación

Implementation and results of a system to evaluate the quality of University teaching in a decade of experimentation

Ana García-Olalla¹, Aurelio Villa Sánchez, Marian Aláez y Susana Romero-Yesa

Universidad de Deusto, España

Resumen

La evaluación de la calidad de la docencia universitaria se ha convertido desde hace una década en un tópico de máxima relevancia por el impacto que tiene en el nuevo enfoque propuesto en el Espacio Europeo de Educación Superior (EEES) y por el cambio de papel que el profesorado debe desempeñar. Este artículo presenta un sistema pionero en la forma que se evalúa la calidad de la docencia a través de un doble Label de calidad: I Diseño y Planificación y II Puesta en práctica. Los análisis y resultados presentados se refieren exclusivamente al Label I, aplicado durante diez años. Se han generado 1930 informes del profesorado y a través de una muestra probabilística estratificada se han analizado 379. El método empleado ha sido fundamentalmente cuantitativo con un complemento cualitativo. Los resultados permiten concluir que se percibe una mejora de la calidad docente vinculada con una disminución de los errores a lo largo del proceso analizado. Esto nos lleva a debatir la relevancia de la evaluación como función de mejora y no exclusivamente sumativa.

Palabras clave: calidad docente; evaluación del profesorado; competencias docentes; estándares docentes.

¹ **Correspondencia:** Ana García-Olalla, ana.garciaolalla@deusto.es, Universidad de Deusto, Unidad de Innovación Docente, Avda. Universidades 24, 48007, Bilbao, España.

Abstract

The quality evaluation of university teaching has become a highly relevant topic in the last decade given its centrality on the European Higher Education Area's (EHEA) approach to higher education teaching and the shift towards new, comprehensive visions of the teacher's role nowadays. With this in mind, this article presents a pioneering method for assessing the quality of teaching through a double quality label: I Design and Planning and II Implementation. The analyses and results here presented refer exclusively to the first label, which has been applied for ten years. During this time, a total of 1930 teacher reports have been issued. 379 reports have been analysed through a stratified probabilistic sample, applying a mostly quantitative method with a qualitative complement. Results show that there has been an improvement in teacher quality linked to a decrease in errors, which brings us to discuss the relevance of assessment as an agent of improvement rather than, merely, a summative factor.

Keywords: teaching quality; teacher assessment; teaching competences; teaching standards.

Introducción

La evaluación del profesorado ha sido una temática muy estudiada por la importancia del factor docente en la calidad de la enseñanza. Como muestra, al introducir los términos “teaching standards” en el buscador Océano (herramienta de búsqueda de recursos bibliográficos) aparecen 985.871 resultados que, filtrados para Educación Superior, pasan a 72.492.

A finales de los 80 y década de los 90 emergieron una serie de autores de gran prestigio que marcaron las líneas de estudio e investigación sobre la evaluación del profesorado de finales del siglo XX e inicios del XXI. La mayoría de estos estudios se han centrado en la evaluación del docente como transmisor del conocimiento, considerando al buen profesor por la forma de impartir sus clases y tratando de identificar sus factores clave.

Con la incorporación del Espacio Europeo de Educación Superior (EEES) se propone un cambio de paradigma docente basado en un mayor protagonismo de los estudiantes y el cambio de rol del profesor a facilitador del aprendizaje. Es evidente que existe una vinculación relevante entre docencia y evaluación de la calidad (Muñoz-Cantero, Ríos y Abalde, 2002).

En este contexto se desarrollan paralelamente dos líneas de trabajo e investigación que a menudo confluyen e intentan definir cómo ha de ser una buena docencia:

- Una línea orientada al examen de un nuevo perfil docente en términos de competencias.
- Otra línea, ligada a la búsqueda de calidad, que describe las buenas prácticas docentes en términos de estándares de calidad.

Ambas líneas pueden considerarse complementarias. En nuestro contexto institucional se definieron un conjunto de competencias docentes deseables para promover el modelo de formación pretendido.

Se identificaron cinco dimensiones que fueron validadas a través de una estructura factorial hallada totalmente coincidente con la teórica (Villa, Campo, Arranz, Villa y

García-Olalla, 2013): *planificación, gestión del aprendizaje, tutoría y evaluación, revisión y mejora, y colegialidad o coordinación entre el profesorado.*

Estas dimensiones se han empleado para contrastar las competencias docentes propuestas por los distintos modelos, valorando la frecuencia de cada una de ellas. Siendo imposible la presentación de todos, se muestra una comparativa de 20 modelos, seleccionados por su adecuación al nuevo enfoque de enseñanza y su relevancia en las dos décadas estudiadas.

Se observa mayor presencia (más del 75%) de aquellas competencias que corresponden a tres dimensiones básicas de la buena docencia: planificación, gestión del aprendizaje y evaluación, además de las relaciones interpersonales. Por contraste, la menor presencia se da en competencias relacionadas con dimensiones emergentes de revisión y mejora, y de colegialidad (entre 5 y 45%).

El encargo de la Unión Europea a la ENQA de elaborar un marco de criterios y directrices para incorporar la calidad en los sistemas universitarios europeos (ENQA, 2005), así como que cada centro universitario sea responsable de definir su sistema de calidad y someterlo a acreditación posterior ante las agencias (Villa, 2008), han situado en un primer plano la definición de estándares de calidad docente en cada contexto institucional.

La garantía de calidad en la educación superior se ha centrado históricamente en instalaciones, políticas, disposiciones, entornos y procesos, más que en los resultados del aprendizaje (Sadler, 2017). Así, el enfoque del criterio 4 relativo al profesorado estaba muy centrado en los procedimientos (ENQA, 2005). Sin embargo, en 2015 (ENQA, 2015) se produce un notable cambio hacia un enfoque más cualitativo en los dos criterios relativos a la docencia, focalizados ahora en promover una enseñanza centrada en el estudiante y una evaluación coherente.

Quince años después, el sistema de calidad docente de gran parte de las universidades españolas sigue el Programa Docencia de ANECA (ANECA, 2006). No obstante, no resulta claro que la mayoría esté empleando estándares que definan cualitativamente la buena docencia. Por ello resulta relevante clarificar el concepto y utilidad de los estándares, así como dar cuenta de experiencias contrastadas de buenas prácticas en su empleo.

Tabla 1

Correspondencia entre los criterios para la excelencia docente de Villa y García-Olalla (2006) con 19 modelos de competencias docentes. Fuente: adaptado de Lucía Campo (2016).

Villa y García 2006	Alegre y Villar 2006	Álvarez-Rojo 2009	Apodaca y Grad 2005	Cano 2005	García-Valcárcel 2003	Gargallo et al. 2011	González y López 2010	Loredo y Romero 2012	Mas y Ruiz 2007	Medina et al. 2011	Molenaar et al. 2009	Ortega 2010	Rodríguez-Espinar 2003	Rueda-Beltrán 2009	Saravia 2011	Tejada 2009	Tigelaar et al. 2004	Torra et al. 2012	Zabalza 2003	Sumatorio (N= 19)	Porcentaje
Gestión del aprendizaje																				17	89.47
Planificación																				15	78.95
Relaciones interpersonales																				15	78.95
Evaluación del aprendizaje																				14	73.68
Orientación/ apoyo al estudiante																				12	63.16
Trabajo en equipo																				10	52.63
Innovación pedagógica																				9	47.37
Orientación al desarrollo profesional																				9	47.37
Manejo de las TIC																				8	42.11
Creación de un clima favorable al aprendizaje																				7	36.84
Gestión del tiempo																				4	21.05
Liderazgo																				3	15.79
Autoevaluación																				3	15.79
Compromiso institucional																				2	10.53
Compromiso ético																				1	5.26
Orientación a la calidad																				1	5.26
Orientación a los resultados																				0	0.00

Los primeros estándares de evaluación en educación fueron aplicados por el Comité Conjunto para evaluar programas (Sanders, 1998), evaluación del personal (Stufflebeam, 2001), y para la mejora de los sistemas de evaluación (McKenna, Nevo, Stufflebeam y Thomas, 1998). Refiriéndonos a la evaluación de personal, este Comité concibe el estándar

dar como “la evaluación sistemática de la actuación profesional y/o cualificaciones de una persona en relación a un cometido profesional y algún fin institucional justificado y defendible” (Stufflebeam, 2001, p. 26).

A nivel internacional se pueden encontrar numerosas propuestas, siendo tradicionales las de los países anglosajones, que definen estándares para la selección de su profesorado y para la formación pre-servicio, además de la evaluación de la actuación docente. Algunas representativas de distintos ámbitos territoriales son:

- En Estados Unidos la NBPTS, Junta Nacional para Estándares de Enseñanza Profesional (2016) publica una propuesta con tres componentes: cinco grandes estándares (compromiso con los estudiantes y su aprendizaje, conocimiento de la materia y su enseñanza, gestión y seguimiento del aprendizaje, reflexión sobre su práctica, y participación como miembros de comunidades de aprendizaje), el proceso de evaluación/certificación y el componente de desarrollo profesional (Serafini, 2002); convirtiéndose en el marco más utilizado por las universidades americanas, entre los más conocidos: Modelo In TASC (CCSSO, 2013), modelo de la Texas Education Agency (2011) y de las instituciones de Formación del profesorado de la NCATE (2008).
- En el Reino Unido la Academia de Educación Superior ha definido un marco de estándares profesionales, el UKPSF, United Kingdom Professional Standards Framework (HEA, 2011), con 15 estándares en torno a tres dimensiones: áreas de actividad (diseño y programas, enseñar y apoyar el aprendizaje, evaluación y retroalimentación, entornos de aprendizaje efectivos, apoyo y orientación, desarrollo profesional continuo), conocimientos básicos (asignatura, métodos de enseñanza y aprendizaje, cómo aprenden los estudiantes, uso apropiado de tecnologías, métodos de evaluación, mejora y calidad de la docencia) y valores profesionales (respeto individual y a la diversidad, participación e igualdad de oportunidades, enfoques basados en la evidencia y la investigación, visión del contexto para la práctica profesional); diferenciando 4 niveles de desarrollo profesional (Hibbert y Semler, 2016).
- Australia extiende en su amplio territorio estándares que difieren en cada área geográfica: los estándares NPQTL, los de Queensland, los Victorianos y los de Westero (Louden, 2000). Merece especial atención la experiencia desarrollada conjuntamente por un grupo de universidades australianas y neozelandesas (Chalmers, 2008), como la de Auckland (2011) en Nueva Zelanda, y las de Western (2009) y Macquarie (2010) en Australia.
- En España destacamos un estudio interuniversitario que realiza una amplia revisión de modelos de estándares de evaluación docente en educación superior a nivel internacional (Marbán, 2012).

De la revisión de los modelos de estándares se desprenden algunas conclusiones respecto al tipo de dimensiones y contenidos evaluados:

- Los estándares más frecuentes son los referidos a: el dominio de la materia, conocimientos y competencias psicopedagógicas, la planificación de la enseñanza, las metodologías de enseñanza-aprendizaje, los procedimientos de evaluación,

la elaboración y empleo de materiales y recursos, la relación con los estudiantes, la motivación y el clima de aprendizaje, y los resultados de aprendizaje.

- Se constata una evolución en la última década, incluyendo dimensiones como: el seguimiento, la orientación y tutoría, el empleo de las TIC, la colaboración y colegialidad, el aprendizaje profesional, la reflexión sistemática, la autoevaluación y revisión docente, la innovación y mejora, y el desarrollo profesional.
- Aparecen ciertos aspectos diferenciales contemplados en función del modelo educativo de una institución determinada o un grupo de ellas que trabajan en colaboración, incorporando dimensiones relacionadas con: el compromiso con la institución, los valores profesionales, el compromiso ético y/o social, la interculturalidad, la adaptación y respuesta a la diversidad, el liderazgo, la comunicación con las familias y otros miembros de la comunidad, entre otras.

Como contribución más reciente, la Red Española de Docencia Universitaria (REDU) ha publicado un trabajo que viene desarrollando en los tres últimos años para proponer un Marco de Desarrollo Académico Docente (Paricio, Fernández y Fernández, 2019) que, si bien se inspira en los precedentes modelos de Australia y Reino Unido, difiere en la concepción de su finalidad y empleo, orientándolo más al desarrollo profesional que a la acreditación. Sin duda esta propuesta tiene un valioso potencial para abrir un amplio horizonte de desarrollo profesional en el ámbito estatal.

Para definir el perfil profesional en la institución objeto de estudio (Universidad de Deusto, UD en adelante), se analizó el proceso de docencia definiendo cinco fases. Tras la selección y definición de diecisiete competencias docentes alineadas con ellas, se precisaron criterios e indicadores que permitieran su evaluación, tanto formativa como acreditativa -Label-, desarrollada en tres niveles: planificación, puesta en práctica, y revisión y mejora.

El procedimiento de evaluación acreditativa Label 1, de *calidad docente en planificación*, hace referencia a asegurar, de acuerdo con el Modelo de Formación de la UD (MFUD), un aprendizaje autónomo y significativo por parte de los estudiantes para el desarrollo de las competencias académico-profesionales de la asignatura. Para ello, se definen y organizan los objetivos de aprendizaje, contenidos, metodología, tiempos, recursos materiales, sistemas de tutoría y evaluación, reflejando de forma organizada estos elementos en el programa público de la asignatura, y documentando las especificaciones requeridas para su desarrollo en la guía de aprendizaje del estudiante, que será evaluada cualitativamente por un comité (Villa y García-Olalla, 2014). Los criterios e indicadores de evaluación se recogen en la Tabla 2 (UD, 2007).

Tabla 2

Criterios e indicadores de evaluación del Label 1

CRITERIOS	INDICADORES
1.1 Estándar <i>contribución al perfil y competencias</i> :	1.2.1 <i>Contribución al perfil</i> . Contextualiza y justifica la asignatura describiendo su contribución al Perfil Académico-Profesional de la Titulación.
1.2 Contextualiza y justifica la asignatura en función del Perfil académico-profesional de la titulación, y define los resultados esperados del aprendizaje en términos de Competencias Genéricas y Específicas que se van a desarrollar.	1.2.2 <i>Competencias genéricas</i> . Selecciona la Competencia Genérica trabajada en la asignatura y define el nivel de desarrollo, de acuerdo al MFUD y al Mapa de Competencias de la Titulación. 1.2.3 <i>Competencias específicas</i> . Selecciona y define de manera correcta las Competencias Específicas que desarrolla la asignatura.
1.3 Estándar <i>estrategia de enseñanza-aprendizaje y adecuación al ECTS</i> : Diseña una estrategia didáctica coherente con el MFUD y las competencias que se pretende desarrollar, y planifica la actividad de aprendizaje de los estudiantes en términos del ECTS	1.3.1 <i>Estrategia didáctica</i> . Diseña una estrategia didáctica coherente con las competencias que pretende desarrollar y acorde a los principios de autonomía y significatividad del aprendizaje de los estudiantes, de acuerdo al MAUD. 1.3.2 <i>ECTS</i> . Planifica la actividad de aprendizaje de los estudiantes exigiendo el cumplimiento de los ECTS asignados a la asignatura, detallando las actividades a realizar y los tiempos previstos para las mismas. 1.3.3 <i>Documentación y recursos</i> . Explicita la documentación y los recursos de soporte que se van a utilizar para el correcto seguimiento de la asignatura y la realización de las actividades.

- 1.4 Estándar sistema de evaluación:** Diseña un sistema de evaluación para la asignatura que define los indicadores de evaluación de las competencias trabajadas, las técnicas que se van a utilizar, y el sistema de calificación que se va a aplicar.
- 1.4.1 Seguimiento y tutoría.** Establece los procedimientos, tiempos y espacios para realizar el seguimiento, la tutoría y orientación del aprendizaje de los estudiantes.
- 1.4.2 Indicadores.** Formula los indicadores a partir de los cuales puede obtenerse evidencias relevantes y significativas de logro de las competencias por los estudiantes.
- 1.4.3 Técnicas e instrumentos.** Selecciona las técnicas e instrumentos que se van a utilizar para recoger la información relativa a los indicadores seleccionados, a lo largo del proceso y al final del mismo.
- 1.4.4 Sistema de calificación.** Define y detalla el sistema de calificación que se va a aplicar para realizar la ponderación cuantitativa que refleje el grado de desarrollo de las competencias a lo largo del proceso, y que se expresará en la calificación final, tanto para la convocatoria ordinaria como extraordinaria.
-

Tras la obtención de la acreditación en planificación, el profesor puede presentarse al Label 2 de excelencia docente en la puesta en práctica. Un comité de evaluación valorará su docencia mediante el análisis cualitativo de un portafolio docente elaborado a este efecto, a través de 3 criterios que se concretan en 10 indicadores: gestión del aprendizaje (metodología, gestión del tiempo y recursos), orientación y tutoría (orientación y sistema de tutoría), y evaluación del aprendizaje (información, técnicas, retroalimentación y resultados de aprendizaje).

Posteriormente, la renovación de la acreditación dependerá de cómo el profesorado desarrolle la revisión y mejora de su práctica, así como de sus competencias de colegialidad docente.

Pueden verse varias aportaciones de su descripción y puesta en práctica en el conjunto de la universidad (García-Olalla, 2014; García-Olalla y Auzmendi, 2012; Villa y García-Olalla, 2006, 2014) y en algunos de sus centros, como Ingeniería (Romero, 2014), Económicas (Eizaguirre, García-Olalla, Del Orden y García-Feijoo, 2017) y Derecho (Aláez, García-Olalla, García-Feijoo y Eizaguirre, 2018).

Método

El presente trabajo parte de la premisa de que un modelo de desarrollo profesional, con una buena práctica docente definida en términos de estándares, debiera constatar una mejora significativa de dicha práctica en línea con los estándares definidos.

Tras diez años de aplicación de dos Label de calidad docente, es necesario evaluar esta experiencia y extraer algunas conclusiones que ayuden a mejorar. La importancia del estudio radica en la consecuencia que tienen los resultados en el perfeccionamiento y promoción del profesorado, además de lograr un sistema de acreditación de la calidad por parte de la universidad. Esto significa contar con un sistema válido para valorar la calidad docente en un modelo basado en competencias y en la organización del crédito europeo.

Objetivos

Se plantea el objetivo de contrastar en qué medida la aplicación durante una década de un modelo sistemático de estándares docentes ha permitido una mejora significativa en su cumplimiento. Más concretamente, centraremos la atención en la evolución y mejora de la planificación de la docencia, tratando de responder a las siguientes cuestiones:

- ¿En qué medida se mejora en los resultados de la planificación docente de acuerdo con los estándares planteados?
- ¿Cuáles son los criterios e indicadores que han revelado un mayor nivel de dificultad?
- ¿Cuáles son las principales dificultades halladas por el profesorado?
- ¿Cuáles son los criterios e indicadores en los que se ha constatado una mayor evolución en el tiempo?
- ¿Qué retos se afrontan tras una década de implantación?

Población y Muestra

El estudio toma como población objeto de análisis el conjunto de 1930 expedientes evaluados durante las diez convocatorias de evaluación acreditativa Label 1 entre 2008 y 2018. Se han diferenciado cuatro periodos:

- Periodo 1: tras la aprobación del nuevo modelo de planificación (2006) en 2008/09 se realizó la primera convocatoria experimental Label 1, en la que participó más del 60% del profesorado. Este sería el punto de partida del proceso.
- Periodo 2: comprende las cuatro siguientes convocatorias, entre 2009/10 y 2012/13, en las que se presentaba la planificación de las asignaturas de los cursos que se iban implantando progresivamente en los nuevos grados.
- Periodo 3: tras la experiencia del periodo anterior, las tres convocatorias celebradas entre 2013/14 y 2015/16 se podrían considerar de consolidación y mejora, tanto del procedimiento como de la propia planificación.

- Periodo 4: tras ocho años de implantación, en el curso 2015/16 se realizó un proceso de reflexión institucional para detectar los puntos fuertes y débiles del modelo, fruto del cual se desarrolló una nueva versión de guía de aprendizaje, así como la adaptación de su evaluación, que entraron en vigor en 2016/17.

Tabla 3

Población (P) y muestra (M) de informes Label 1 analizados

Centro	Periodo 1		Periodo 2		Periodo 3		Periodo 4		TOTAL	
	P	M	P	M	P	M	P	M	P	M
Psicología y Educación	37	7	137	27	46	9	40	8	260	51
Ingeniería	66	13	171	34	14	4	30	6	281	57
Derecho	69	13	123	23	73	15	30	6	295	57
CC EE y Empresariales	77	15	136	26	87	15	29	6	329	62
CC Sociales y Humanas	107	21	299	57	107	21	46	9	559	108
Teología			23	5	1	1	5	1	29	7
Formación Humana y Valores			37	8	4	1			41	9
Escuela de Formación Profesorado			83	16	46	10	7	2	136	28
TOTAL	356	69	1009	196	378	76	187	38	1930	379

Se ha utilizado una muestra probabilística estratificada para considerar, proporcionalmente, las diferentes áreas presentadas a evaluación. La muestra está basada en el universo de 1930 profesor-asignatura presentados a evaluación durante diez años. Aplicando un margen de error del 5% y un nivel de confianza del 95%, se ha requerido una muestra de 372.

La estratificación aumenta la precisión de la muestra e implica el uso deliberado de diferentes tamaños muestrales para cada estrato, a fin de lograr reducir la varianza. Siguiendo a Kish (1995), cada estrato se multiplicará por 0,192 para obtener la muestra necesaria:

$$ksh = \frac{n}{N} = \frac{372}{1930} = 0,192.$$

Aunque la muestra necesaria era 372 se han analizado 379 por redondeo.

Procedimiento de recogida y análisis de datos

La evaluación se lleva a cabo por cuatro comités formados por pares de expertos -externos e internos- en las materias presentadas a evaluación, los externos son evaluadores reconocidos por ANECA y los internos por su acreditada trayectoria en innovación y liderazgo pedagógico.

Analizan el programa y la guía de aprendizaje mediante tres instrumentos: (a) un protocolo de evaluación que detalla para cada criterio e indicador de manera descriptiva y cualitativa las evidencias pertinentes para justificar su cumplimiento; (b) una hoja de respuestas en la que se evalúa, cuantitativa y cualitativamente el cumplimiento de cada criterio e indicador, según el protocolo establecido ; y (c) un informe en el que se indica el resultado global de la evaluación y las observaciones que expresan para cada criterio las modificaciones necesarias para su aprobación y/o las recomendaciones para su mejora.

Cada expediente es revisado por dos evaluadores, uno externo y otro interno, primero individualmente y luego conjuntamente. Cuando hay disconformidad en alguno de los criterios, se lleva a la reunión del comité sometiéndolo a la revisión del presidente. Los casos en los que hayan surgido divergencias se tratan en la reunión general de los cuatro comités, celebrada al final del proceso, en aras de la uniformidad en la aplicación de los criterios.

El proceso de evaluación se desarrolla en dos fases cuyo resultado puede ser:

- Fase 1: *favorable*, si se han cumplido todos los criterios e indicadores del protocolo; *favorable con recomendaciones* (sugerencias de mejora que se considerarán en Label 2); o *modificaciones necesarias*, en cuyo caso se deberán introducir y presentar en la fase 2 en el plazo de un mes, recibiendo el asesoramiento oportuno.
- Fase 2: *favorable*, *favorable con recomendaciones*, o *no favorable* -si no se han introducido las modificaciones solicitadas en fase 1-.

En este estudio se han analizado 379 informes, considerando tres aspectos:

- El resultado global de la evaluación: *favorable* o *no favorable*, tanto en fase 1 como en fase 2.
- El resultado *favorable* o *no favorable* en cada uno de los criterios e indicadores.
- Un análisis cualitativo de dificultades y/o errores que aparecían en las modificaciones propuestas por los evaluadores. Se identificaron un total de 60 errores típicos, y el análisis de su presencia en cada uno de los informes y periodos permitió identificar las principales dificultades y analizar su evolución.

Resultados

En la primera aplicación experimental del procedimiento de acreditación (periodo 1) se convocó al conjunto del profesorado con dedicación y algo más de la mitad del profesorado presentado obtuvo una evaluación final favorable (53,62%), frente al 46,38% que no consiguió cumplir satisfactoriamente los tres criterios. Este dato revela la dificultad del momento inicial, a pesar de la formación recibida.

Ya en el periodo 2, que abarca los cuatro primeros cursos de implantación de los grados, el porcentaje de favorables aumenta a un 86,73%, aunque en la primera fase del proceso se produjeran incluso más fallos que en el periodo experimental.

Los periodos 3 y 4 muestran una clara consolidación de la mejora iniciada en el periodo anterior, con un 89,47% y un 94,59% de guías favorables respectivamente.

Son periodos en los que el mapa de errores muestra que se superan algunas dificultades de los inicios y, sin embargo, persisten otras.

Criterio I. Contribución al Perfil y Competencias

En la aplicación experimental el criterio 1 de *contribución al perfil y competencias* fue el que obtuvo una valoración más favorable, siendo superado por el 63,77% del profesorado. En periodos posteriores este porcentaje llegaría a niveles por encima del 94%.

Tabla 4

Porcentaje de favorables en cada fase y de no favorables sobre total de presentados para cada indicador y criterio

CRITERIO INDICADOR	PERIODO 1			PERIODO 2			PERIODO 3			PERIODO 4		
	% FAV F1	% FAV F2	% NF	% FAV F1	% FAV F2	% NF	% FAV F1	% FAV F2	% NF	% FAV F1	% FAV F2	% NF
	Total informes	39,13	14,49	46,38	54,59	32,14	13,27	55,26	34,21	10,53	13,51	81,08
Criterio 1:	40,58	23,19	36,23	76,53	18,37	5,10	90,79	6,58	2,63	86,49	10,81	2,70
1.1 Contribución al perfil	40,58	31,88	27,54	90,31	9,69	0,00	93,42	6,58	0,00	94,59	5,41	0,00
1.2 Competencias genéricas	60,87	5,80	33,33	83,67	13,78	2,55	94,74	3,95	1,32	94,59	2,70	2,70
1.3 Competencias específicas	56,52	7,25	36,23	88,78	7,65	3,57	93,42	3,95	2,63	91,89	8,11	0,00
Criterio 2:	46,38	10,14	43,48	58,67	33,16	8,16	73,68	17,11	9,21	59,46	35,14	5,41
2.1 Estrategia didáctica	47,83	8,70	43,48	73,47	18,88	7,65	73,68	22,37	3,95	64,86	29,73	5,41
2.2 ECTS	56,52	5,80	37,68	78,06	19,39	2,55	93,42	2,63	3,95	86,49	10,81	2,70
2.3 Documentación y recursos	79,71	0,00	20,29	78,06	18,37	3,57	94,74	3,95	1,32	100,0	0,00	0,00
Criterio 3:	40,58	13,04	46,38	63,27	29,08	7,65	68,42	22,37	9,21	56,76	37,84	5,41
3.1 Seguimiento y tutoría	62,32	5,80	31,88	86,22	12,24	1,53	97,37	1,32	1,32	78,38	18,92	2,70
3.2 Indicadores	55,07	4,35	40,58	84,18	13,27	2,55	90,79	7,89	1,32	91,89	8,11	0,00
3.3 Técnicas e instrumentos	53,62	5,80	40,58	80,10	15,82	4,08	93,42	6,58	0,00	97,30	2,70	0,00
3.4 Sistema de calificación	49,28	7,25	43,48	80,10	17,86	2,04	73,68	21,05	5,26	64,86	29,7	5,41

Un análisis por indicadores nos lleva a concluir que el referido a la *contribución al perfil* fue el mejor logrado (72,46%) en el primer periodo, aunque un 27,54% del profesorado no hacía aún referencia al perfil profesional de la titulación o lo hacía

insuficientemente. En los periodos siguientes ninguna guía terminó con errores en este indicador.

En el periodo 1, el 66,67% del profesorado superó el indicador relativo a las *competencias genéricas*, a pesar de su novedad. Los principales errores fueron: formularlas incorrectamente, no seguir el modelo UD, no indicar el nivel de dominio y omitir estas competencias. Estos problemas pueden darse por superados a partir del periodo 2 ya que desde ahí hasta el último el porcentaje de profesores que incumplen no supera el 3%.

El indicador referido a la formulación de las *competencias específicas* fue el más problemático en el periodo 1: un 36,23% del profesorado no las formulaba correctamente o lo hacía en un número excesivo. Este porcentaje se redujo drásticamente en el periodo 2 (3,57%) y continuó haciéndolo en el periodo 3 (2,63%) hasta desaparecer totalmente en el periodo 4.

Criterio 2. Estrategia de enseñanza-aprendizaje y adecuación al ECTS

El Criterio 2 supuso una notable dificultad para el profesorado en el periodo inicial, ya que un 43,48% no fue capaz de diseñar en aquel momento una estrategia de enseñanza-aprendizaje adecuada al contexto de la innovación esperada ni al sistema de créditos del espacio europeo. Este porcentaje bajaría en los periodos posteriores, reduciéndose la proporción de profesores que no lo supera hasta un máximo del 5,5%.

El análisis por indicadores permite identificar el más problemático: la *estrategia didáctica*. En el primer periodo casi un 44% del profesorado no lo supera, siendo los principales motivos: falta de concreción y precisión, diseñar una estrategia en torno a los contenidos y no a las competencias, no plantear una estrategia coherente en el conjunto de actividades, no describir suficientemente las actividades y no diferenciar las fases de acuerdo al modelo de aprendizaje autónomo y significativo de la UD. Este indicador será uno de los pocos que se mantendrá con problemas en periodos posteriores, si bien su incidencia se reduce drásticamente respecto al periodo 1, cayendo al 7,65% y manteniéndose en un nivel por debajo del 5,5% en el periodo 4.

El 37,68% de profesorado halló dificultades en planificar la estrategia en términos del *ECTS*, por falta de precisión, por no detallar el tiempo estimado para cada actividad, o por no aportar un procedimiento para el seguimiento del plan de trabajo por los estudiantes. Los profesores van aprendiendo de estos errores a lo largo de los periodos restantes hasta hacerlo prácticamente desaparecer (2,70% en el cuarto periodo).

El tercer indicador de este criterio, relativo a la *documentación y recursos*, fue superado por casi un 80% del profesorado en el periodo 1, siendo el mejor valorado del protocolo. En el periodo 2 se siguen cometiendo pequeños fallos en el detalle de la *documentación y recursos* (3,57%) que van reduciéndose (1,32% en periodo 3) hasta llegar al 0% en el periodo 4.

Criterio 3. Sistema de Evaluación

Este fue el criterio de mayor dificultad en el primer periodo, donde el 46,38% del profesorado no logró definir un *sistema de evaluación* correcto. En los periodos posteriores se aprecia el aprendizaje del profesorado y los no acreditados caen al 7,65% en

el periodo 2, llegan al 9,21% en el periodo 3 y concluyen en 5,41% en el periodo 4. El diseño del *sistema de evaluación* sigue siendo una de las dificultades que los profesores encuentran hoy en día para acreditarse en la primera fase de evaluación: en el cuarto periodo tan solo el 56,76% lo logra en esta fase.

El primer indicador de este criterio, *seguimiento y tutoría*, es el menos problemático en todos los periodos. A pesar de ello, casi un 32% no estableció un procedimiento de *seguimiento y tutoría* adecuados en el primer periodo. Posteriormente, la gran mayoría de los profesores superó este indicador incluso en la primera fase de evaluación, con la excepción del cuarto periodo en el que, si bien un 98,68% lo aprueban, llama la atención que un 18,92% lo hagan en segunda fase. La explicación responde al cambio de modelo de guía que se introduce en el cuarto periodo y que, entre otras cosas, establece la integración de las actividades de *seguimiento y tutoría* en el plan de trabajo del estudiante. Los profesores no especificaban cómo y cuándo realizarían la orientación y el seguimiento y darían feedback de las actividades, a pesar de la relevancia del feedback señalada por la investigación (Ajjawi y Boud, 2017).

Centrándonos en el segundo indicador, en el periodo 1 más del 40% del profesorado no supo formular correctamente los *indicadores* para la evaluación de las competencias trabajadas, o los formuló en un número insuficiente o excesivo, o no los estableció para las competencias genéricas. Lo que había supuesto un 40,58% en el periodo 1 dejó de ser problema en los siguientes periodos, con cifras inferiores al 3%.

Idéntico nivel de dificultad se encontró en el tercer indicador para el primer periodo, fundamentalmente por no detallar suficientemente las *técnicas* para evaluar las competencias, pero también por no ser coherentes estas con las competencias que pretenden evaluar, por plantear poca variedad o un excesivo peso del examen final. En el periodo 2 los problemas con este indicador (4,08%) se concentran en la coherencia en cuanto a la definición de las *técnicas e instrumentos de evaluación*, y en los periodos 3 y 4 ya no se detectan errores.

La mayor dificultad del primer periodo en todo el protocolo de evaluación estuvo en el indicador referido al *sistema de calificación*, desfavorable para el 43,48% del profesorado. El motivo más grave fue no distribuir la calificación de la asignatura por competencias, pero también diseñar un sistema de evaluación confuso, no detallar el valor de cada técnica en la calificación final, un sistema de evaluación poco coherente con la estrategia de enseñanza-aprendizaje o no planificar la evaluación de la competencia genérica. El logro de este indicador mejora notablemente a partir del siguiente periodo hasta el final. Sin embargo, la asignación de calificaciones por competencias sigue siendo el principal motivo por el que los profesores no superan a la primera este indicador (en el periodo 4 casi un 30% lo aprueba en la segunda fase).

Recomendaciones

Los evaluadores han realizado numerosas recomendaciones que sin duda contribuyen a la mejora docente. De estos comentarios de los evaluadores se deduce la necesidad de trasladar explícitamente al estudiante la justificación de la asignatura en la titulación, a través de una explicación más completa de las funciones y ámbitos profesionales para los que forma la asignatura; la necesidad de incluir un sistema de

seguimiento de tiempos de las actividades de los estudiantes fuera del aula que permita un mayor ajuste de los tiempos planificados y realmente invertidos; y, por último, se les pide mejorar la formulación de indicadores de evaluación para un juicio más certero del grado de avance en cada competencia trabajada.

Discusión y conclusiones

En los últimos 20 años, la educación terciaria ha enfrentado desafíos importantes. En un entorno cada vez más globalizado y competitivo, los proveedores de educación han tenido que reevaluar las actitudes tradicionales hacia la enseñanza y el aprendizaje (Devonshire y Brailsford, 2012). Se han revisado los procedimientos y el rol del profesor para evaluar su desempeño y promover su desarrollo profesional. Sin embargo, como afirman Navarro y Ramírez (2018) tras realizar una revisión sistemática de la literatura sobre evaluación docente, existe aún un déficit de investigación que estudie el impacto de la evaluación en el desarrollo profesional docente.

La primera conclusión que se extrae del análisis de los resultados es una notable mejora en la competencia de planificación del profesorado, resultado de una década de trabajo en un esquema de desarrollo profesional que integra de manera sistemática las prácticas de innovación, evaluación, formación, práctica, revisión y mejora. Sin duda, la transparencia de la planificación docente contribuye a esta mejora, tal como afirman Álvarez, López, Peña y González (2019).

Los criterios 2- estrategia de enseñanza-aprendizaje y adecuación al ECTS- y 3 -sistema de evaluación- son los que concentran la mayor parte de los problemas detectados en las guías de aprendizaje en todos los periodos analizados. De tal manera que las modificaciones necesarias versan fundamentalmente sobre carencias en los indicadores, estrategia didáctica y sistema de calificación, respectivamente. En el primero, los evaluadores señalan que, en muchos casos, de la secuencia de actividades que los profesores presentan en su plan de trabajo, no se deduce una estrategia de enseñanza-aprendizaje coherente: no hay una buena articulación entre competencias, actividades planteadas para su desarrollo y estrategia de evaluación, siendo expresión de una falta de alineamiento constructivo (Biggs y Tang, 2015) y de conciencia respecto a la importancia de una retroalimentación de calidad (Ajjawi y Boud, 2017). El otro problema grave es que las competencias genéricas no se planifican o lo hacen insuficientemente aunque, curiosamente, en la mayoría de los casos, sí se contempla su evaluación.

El sistema de calificación es el otro gran protagonista de las carencias en las guías. Los evaluadores reclaman una mayor claridad de las técnicas y agentes de evaluación que se emplean para medir el avance en cada competencia, así como que se llegue a una máxima coherencia entre lo expuesto en el plan de trabajo del estudiante, las tablas resumen de la evaluación y el programa de la asignatura. La investigación confirma la necesidad de “encontrar un enfoque adecuado para medir o determinar el logro de competencias” (Manríquez, 2018).

A modo de balance cabe destacar cómo, a pesar de las resistencias iniciales del profesorado ante la falta de experiencia en procesos internos de acreditación docente y la tensión entre la actividad docente e investigadora (Theurillat y Gareca, 2015), actualmente un 64% del total del profesorado está acreditado con el Label 1, y este número

se eleva progresivamente al 73% si consideramos sólo al profesorado con dedicación, al 89% con titularidad y al 100% con cátedra. Es valorable también la mejora constatada por la revisión del modelo de guía producida del periodo 4, en la dinamización de los procesos de elaboración, evaluación y empleo de la guía.

Referencias

- Ajjawi, R. & Boud, D. (2017). Researching feedback dialogue: an interactional analysis approach. *Assessment & Evaluation in Higher Education*, 42(2), 252-265. doi:10.1080/02602938.2015.1102863
- Aláez, M., García-Olalla, A., García-Feijoo, M. & Eizaguirre, A. (2018). Evaluación de la calidad docente en una facultad de derecho. *Revista de Educación y Derecho*, 17, 1-28. doi:10.1344/REYD2018.17.23482
- Álvarez, P.R., López, D., Peña, R. y González, M.O. (2019). La guía docente en la planificación y desarrollo de la enseñanza universitaria. *Revista electrónica Investigación y Docencia*, 21, 7-24. doi: 10.17561/reid.n21.1
- ANECA (2006). *DOCENTIA, Programa de apoyo para la evaluación de la actividad docente del profesorado universitario. Modelo de evaluación*. Recuperado de <http://www.aneca.es>
- Biggs, J. & Tang, C. (2011). *Calidad del aprendizaje universitario*. Madrid: Narcea.
- Campo, L. (2016). *Evaluación de las competencias docentes en el ámbito universitario: un sistema de 360º. Aplicación en dos universidades chilenas* (Tesis doctoral inédita). Universidad de Deusto, Bilbao.
- Chalmers, D. (2008). Teaching and learning quality indicators in Australian universities. University of Western Australia. En OECD, *Outcomes of Higher Education: Quality relevance and impact*. Paper for the IMHE 2008 General Conference, París, 8-10 septiembre 2008. Recuperado de <http://www.oecd.org/site/eduimhe08/41216416.pdf>
- Council of Chief State School Officers (2013). *InTASC Model core teaching standards and learning progressions for teachers*. Washington: CCSSO. Recuperado de <http://programs.ccsso.org/content/pdfs/corestrd.pdf>
- Devonshire, P. & Brailsford, I. (2012). Defining pedagogical standards and benchmarks for teaching performance in law schools: contrasting models in New Zealand and the United Kingdom. *The Law Teacher*, 46(1), 50-64. doi: 10.1080/03069400.2012.644970
- Eizaguirre, A., García-Olalla, A., Del Orden, O. & García-Feijoo, M. (2017). Principios básicos de la Pedagogía Ignaciana y su implantación en la Deusto Business School a través del Modelo Deusto de Formación. *Estudios Empresariales*, 1-2(50-51), 80-98.
- ENQA (2005). *Standards and guidelines for Quality Assurance in European Higher Education Area*. Helsinki: ENQA. Recuperado de https://enqa.eu/wp-content/uploads/2015/09/ESG_3edition.pdf
- ENQA (2015). *Standards and Guidelines for Quality Assurance in the European Higher Education Area (ESG)*. Brussels: EURASHE. Recuperado de https://enqa.eu/wp-content/uploads/2015/11/ESG_2015.pdf
- García-Olalla, A. (2014). El portafolio docente: un instrumento para evaluación y mejora de la práctica docente. *Revista CIDUI*, 2, 1-13. Recuperado de <https://www.cidui.org/revistacidui/index.php/cidui/article/view/550>

- García-Olalla, A. & Auzmendi, E. (2012). Evaluación de la planificación de la docencia en las asignaturas de grado. *Revista CIDUI*, 1, 1-19. Recuperado de <https://www.cidui.org/revistacidui/index.php/cidui/article/view/126>
- Hibbert, P. & Semler, M. (2016). Faculty development in teaching and learning: the UK framework and current debates. *Innovations in Education and Teaching International*, 53(6), 581-591. doi: 10.1080/14703297.2015.1022201
- Higher Education Academy (2011). *The UK Professional Standards Framework for Teaching and Learning in Higher Education*, Recuperado de <https://www.heacademy.ac.uk>
- Kish, L. (1995). *Diseño estadístico para la investigación*. Madrid: Centro de Investigaciones Sociológicas.
- Louden, W. (2000). Standards for standards: The development of Australian professional standards for teaching. *Australian Journal of Education*, 44(2), 118-134. doi: 10.1177/000494410004400203
- Macquarie University (2010). *Teacher standards-individual level*. Recuperado de http://mq.edu.au/learning-and-teaching/activities/research/docs/TeachingStandards_Individual.doc
- Manrriquez, L. (2012). ¿Evaluación en competencias? *Estudios pedagógicos*, 38(1), 353-366. doi: 10.4067/S0718-07052012000100022
- Marbán, J. M. (Coord.) (2012). *Análisis de las herramientas de evaluación de la Calidad Docente mediante contrastes basados en Estándares Internacionales de Excelencia*. Madrid: Ministerio de Educación, Cultura y Deporte, Proyecto EA2011-0113.
- McKenna, B., Nevo, D., Stufflebeam, D. & Thomas, R. (1998). *Guía profesional para la mejora de los sistemas de evaluación del profesorado*. Bilbao: ICE de la Universidad de Deusto-INCE.
- Muñoz-Cantero, J.M., Ríos, M.P. & Abalde, E. (2002). Evaluación docente vs. evaluación de la calidad. *Revista Electrónica de Investigación y Evaluación Educativa (RELIEVE)*, 8(2) 103-134. Recuperado de http://www.uv.es/RELIEVE/v8n2/RELIEVEv8n2_4.htm
- National Board for Professional Teaching Standards (2016). *What teachers should know and be able to do*. Arlington: NBPTS. Recuperado de <http://accomplishedteacher.org/wp-content/uploads/2016/12/NBPTS-What-Teachers-Should-Know-and-Be-Able-to-Do-.pdf>
- National Council for Accreditation of Teacher Education (2008). *Professional standards for the accreditation of teacher preparation institutions*. Washington. Recuperado de www.ncate.org
- Navarro C. & Ramírez, M.S. (2018). Mapeo sistemático de la literatura sobre evaluación docente (2013-2017). *Educação e Pesquisa*, 44, 1-23. doi: 10.1590/s1678-4634201844185677
- Paricio, J., Fernández, A. & Fernández, I. (Eds.) (2019). *Cartografía de la buena docencia Universitaria*. Madrid: Narcea.
- Romero, S. (2014). Label 2: el portafolio del profesor como herramienta para la medida de la calidad docente. En actas del XI Congreso de Tecnologías, Aprendizaje y Enseñanza de la Electrónica (TAEE 2014), Bilbao: Universidad de Deusto. http://taee.etsist.upm.es/actas/2014/Actas_Congreso_TAEE_2014.pdf
- Sadler, R. (2017). Academic achievements standards and quality assurance. *Quality in Higher Education*, 23(2), 81-99. doi: 10.1080/13538322.2017.1356614
- Sanders, J.M. (1998). *Estándares para evaluación de programas*. Bilbao: Mensajero.

- Serafini, F. (2002). Possibilities and challenges. The national board for professional teaching standars. *Journal of Teacher Education*, 53(4), 316-327. doi: 10.1177/0022487102053004004
- Stufflebeam, D.L. (Dir) (2001). *Estándares de Evaluación de Personal*. Bilbao: Mensajero.
- Texas Education Agency (2011). *Texas educator certification preparation manual*. Texas. Recuperado de <https://tea.texas.gov/>
- Theurillat, D. & Gareca, B. (2015). Organización de docencia e investigación en universidades: Una exploración al caso chileno. *Calidad en la Educación*, 42, 120-160. doi:10.4067/S0718-45652015000100005
- Universidad de Deusto (2007). *Manual para la garantía interna de calidad docente: modelo de desarrollo profesional en la UD*. Bilbao: Universidad de Deusto.
- University of Auckland (2011). *Guidelines for effective teaching. Teaching and learning quality committee*. Auckland. Recuperado de <http://www.auckland.ac.nz/webdav/site/central/shared/about/teaching-and-learning/policies-guidelines-procedures/documents/Guidelines-for-effective-teaching-2011.pdf>
- University of Western Australia (2009). *Teaching criteria framework. Teaching quality indicators project. Centre for advancement of teaching and learning*. Recuperado de <http://www.catl.uwa.edu.au/tcf>
- Villa, A., Campo, L., Arranz, S., Villa, O. & García-Olalla, A. (2013). Valoración del profesorado de magisterio sobre el aprendizaje basado en competencias implantado. *Profesorado, Revista de Currículum y Formación del Profesorado*, 17(3), 35-55. Recuperado de <https://www.redalyc.org/articulo.oa?id=56729527003>
- Villa, A. & García-Olalla, A. (2006). Garantía interna de calidad docente: propuesta de un modelo de desarrollo profesional en la Universidad de Deusto. En E. Arregui y otros (Coords.), *Políticas de gestión de las organizaciones educativas. Implicaciones para el currículum universitario en el marco europeo*. Oviedo: Universidad de Oviedo.
- Villa, A. & García-Olalla, A. (2014). Un sistema de garantía de calidad de la docencia: un estudio de caso. *Revista Electrónica Interuniversitaria de Formación del Profesorado*, 17(3), 65-78. doi: 10.6018/reifop.17.3.204061

Fecha de recepción: 28 de octubre de 2019.

Fecha de revisión: 29 de noviembre de 2019.

Fecha de aceptación: 1 de mayo de 2020.