
Tipologías de profesorado universitario en función de la evaluación aplicada a los futuros docentes¹

Clusters-Models of Teachers About Students' Assessment in Teacher Education

ANDRÉS PALACIOS
PICOS

Universidad de Valladolid
palacios@psi.uva.es

VÍCTOR MANUEL
LÓPEZ-PASTOR

Universidad de Valladolid
vlopez@mpc.uva.es

JOSÉ J. BARBA

Universidad de Valladolid
jjbarba@pdg.uva.es

Resumen: En el presente trabajo se analizan las diferentes tipologías de profesorado universitario en función de los sistemas de evaluación que utilizan. A partir de escalas fiables y válidas, se han realizado análisis multivariantes (clúster y ANOVAS) con una muestra de 217 profesores universitarios de Formación Inicial del Profesorado de 14 universidades. Los resultados sugieren la existencia de tres tipos de docentes, en función de los instrumentos de evaluación que utilizan: a) profesorado innovador, b) profesorado orientado a la calificación, y c) profesorado ecléctico con un fuerte componente práctico. Se analizan, por último, las implicaciones educativas que estos resultados pudieran tener.

Palabras clave: Formación del profesorado; evaluación formativa; análisis de conglomerados; perfiles docentes.

Abstract: This study analyzes diverse types of higher education teachers in terms of the systems and assessment tools they use. Reliable and valid attitudinal scales were developed and applied to a sample of 217 teachers from 14 Pre-service Teacher Education institutions in Spain. Data analysis was undertaken through multivariate analysis (cluster analysis and ANOVAs). The results suggest the existence of three types (clusters) of teachers, according to the assessment tools they use: a) innovator teachers, b) marks oriented teachers, and c) eclectic teachers, with a strong practical component. By the last, we analyze the pedagogical results prospective.

Keywords: Teacher education; formative assessment; cluster analysis; professors' profiles.

1 Este artículo forma parte de la ejecución del proyecto de investigación del Plan Nacional de I+D+I EDU 2010-19637 (Subprograma EDUC), titulado: *La evaluación en la formación inicial del profesorado de educación física*.

INTRODUCCIÓN

En los últimos años asistimos a un creciente interés por todo lo que tenga que ver con la evaluación en la formación inicial del profesorado, lo que se plasma en un aumento de investigaciones al respecto (Nieto, 2000; Palacios y López, 2012; Trillo y Porto, 2002; Zabalza, 2003). En ellas se constata que el uso de procesos de evaluación formativa y continua no está muy extendido en la universidad española, aunque se aprecia una ligera tendencia hacia su mayor uso. Ibarra y Rodríguez (2010) consideran que esta tendencia es consecuencia de la participación e implicación del profesorado universitario en proyectos piloto de innovación para la adecuación de las enseñanzas universitarias al Espacio Europeo de Educación Superior EEES.

No obstante esta tendencia, en la literatura sobre el tema pueden encontrarse severas críticas a cómo se desarrolla el proceso de evaluación en la actualidad, hasta el punto de considerar que se trata de uno de los mayores problemas en la docencia universitaria, principalmente debido a la falta de formación del profesorado universitario (Nieto, 2000; Trillo y Porto, 1999; Zabalza, 2003).

Pese a su escaso uso, pueden encontrarse numerosos trabajos que defienden la necesidad de desarrollar sistemas de evaluación formativa y compartida en educación superior, entre otros aspectos, por las mejoras que supone en el aprendizaje del alumnado (Boud y Falchikov, 2007; Brown y Glasner, 2003; López, 2009). No se puede entender la práctica de la docencia sin tener en cuenta la evaluación como una de las competencias básicas (Zabalza, 2003) y uno de los factores que más influyen en la calidad docente (Black y William, 1998). Tampoco podemos olvidar que la forma en que el docente planifica la evaluación determina los contenidos de aprendizaje, las estrategias y los esfuerzos del alumnado (Bonsón y Benito, 2005; Dochy, Segers y Dierick, 2002). Por todo ello, Dochy, Segers y Dierick (2002) consideran que las universidades tienen que afrontar el reto de pasar de una cultura del examen a una cultura de la evaluación; se trataría de situar la evaluación al servicio de los aprendizajes, desarrollando un alto grado de formación y de participación en los procesos. Bonsón y Benito (2005) entienden que la clave de este cambio está en dejar de entender la evaluación del aprendizaje, para practicar una evaluación para el aprendizaje. En este sentido, Carless, Joughin y Mok (2006) y Boud y Falchikov (2007) prefieren utilizar el concepto evaluación orientada al aprendizaje. Se trata, pues, de dos formas de entender y practicar la evaluación del alumnado: una evaluación orientada a la calificación (única o predominantemente sumativa y final) versus una evaluación orientada al aprendizaje (predominantemente formativa y continua).

LA EVALUACIÓN EN LA FORMACIÓN INICIAL DEL PROFESORADO

En el caso de la Formación Inicial de Profesorado (FIP), la evaluación orientada a la calificación es la que se encuentra más extendida (Nieto, 2000; Trillo y Porto, 1999). Trillo y Porto (1999) describen con detalle los sistemas de evaluación que han predominando en nuestro país en la formación inicial del profesorado (FIP) en los últimos 20 años, cuyas principales características son:

- Fuertes contradicciones entre la evaluación y los procesos de enseñanza-aprendizaje.
- El concepto “evaluación” suele entenderse como “calificación”, habitualmente mediante un examen final, predominando el examen escrito sin poder consultar ningún tipo de material.
- El profesor suele ser el único que participa en el proceso de evaluación.
- La mayoría de los profesores no explicitan los criterios de evaluación que aplican.
- Las funciones principales de la evaluación son la selección de alumnos y la certificación de resultados.

Por lo general, el concepto evaluación se interpreta como un sinónimo de calificación o de procedimientos de calificación. Estos modelos de evaluación tradicional también se caracterizan porque la evaluación es competencia única del profesorado. Situación que, en el caso que nos ocupa, adquiere tintes especiales, dado que el alumnado que está estudiando en la FIP son los docentes del futuro. En este contexto, la evaluación es, a la vez, contenido de aprendizaje, competencia profesional a adquirir y patrón de actuación en el desempeño de su futura labor docente fuera de la universidad.

En un trabajo de investigación reciente, López, Manrique y Vallés (2011) encuentran que la implantación de los nuevos estudios de Grado en FIP está generando cambios significativos en los procesos de evaluación:

- Suele ofrecerse al alumnado diferentes vías de aprendizaje y evaluación, dando prioridad a la vía de evaluación continua;
- Se diversifica el sistema de calificación del aprendizaje del alumnado, perdiendo mucho peso el examen, para repartirlo con otras actividades de aprendizaje;
- Se comienza a promover la participación del alumnado en el proceso de evaluación. En cambio, todavía se sigue identificando evaluación con calificación

en la mayor parte de los trabajos; en pocos casos de refleja el uso de sistemas de evaluación formativa.

Por otra parte, Palacios y López (2012), comprueban que en FIP sigue predominando la forma tradicional de evaluación-calificación, pero que comienzan a producirse algunos cambios en años recientes:

- La nota final proviene principalmente del examen, pero suele combinarse con otro tipo de trabajos o actividades;
- Algunas asignaturas utilizan sistemas de evaluación formativa de forma clara;
- También parecen existir asignaturas que favorecen la participación del alumnado en la evaluación.

En este sentido, cada vez aparecen más trabajos publicados sobre la utilización de de sistemas de evaluación formativa en la FIP y el desarrollo de experiencias de implicación del alumnado en los procesos de evaluación (Fernández-Balboa, 2006; López, 2009; Navarro y González, 2010; Santos, Martínez y López, 2009). En todos ellos, se plantean las razones y motivos pedagógicos para llevar a cabo este tipo de evaluación formativa y, en algunos casos, se ofrecen análisis genéricos sobre las ventajas y/o los inconvenientes que pueden surgir en este tipo de procesos evaluativos. En esta línea, Arribas (2012) y López, Pintor, Muros y Webb (2012) estudian la influencia de los sistemas de evaluación formativa en el rendimiento académico del alumnado en FIP, comprobando que la utilización de sistemas de evaluación formativa suele generar mejores rendimientos académicos. Por otra parte, Gutiérrez, Pérez, Pérez y Palacios (2011) y Hamodi y López (2012) encuentran que un porcentaje alto del alumnado de FIP considera que el desarrollo de procesos de evaluación formativa les ha ayudado en el desarrollo de sus competencias profesionales. Similares conclusiones pueden encontrarse en trabajos sobre experiencias piloto en FIP (López, 2009; Santos, Martínez y López, 2009). En todos estos casos se concluye que la utilización de sistemas y procesos de evaluación formativa y compartida mejoran la calidad del proceso enseñanza-aprendizaje y los resultados de aprendizaje, así como la percepción sobre su utilidad práctica en su futura labor docente.

Esta doble situación en relación a los procesos de evaluación en la FIP –alto interés teórico en el uso de procesos orientados hacia el aprendizaje, escaso uso práctico– puede que tenga su correlato vivencial en las aulas universitarias, con la presencia de diferentes modelos de actuación y con la presencia de diferentes tipologías de profesorado a la hora de realizar procesos de evaluación del alumnado. Al

menos esta es la tesis que plantean Palacios y López (2012), quienes encuentran que el profesorado de las facultades de FIP de Primaria puede agruparse en tres perfiles, que se diferencian por sus actitudes hacia la evaluación y los sistemas de evaluación que utilizan. Sus resultados sugieren la existencia de tres tipos de docentes (conglomerados). Un *Profesorado Innovador* utilizaría sistemas y procesos de evaluación formativa y continua, tendería a implicar al alumnado en los procesos de evaluación, con diversidad de instrumentos y técnicas de calificación (portafolios, por ejemplo) y obtendría la calificación a partir de las actividades de aprendizaje que realiza el alumnado. El *Profesorado Tradicional*, segundo de los perfiles de profesor, se caracterizaría por la utilización de sistemas de evaluación sumativa y final, sin implicar al alumnado en los procesos de evaluación y que obtendría la calificación a través de un único examen final o, como mucho, con algún trabajo monográfico con poco peso en la nota definitiva. Por último, el *Profesorado Ecléctico* utilizaría pequeños procesos de evaluación formativa con alguna implicación del alumnado en los procesos de evaluación y combinando el examen final con otras actividades de aprendizaje y evaluación, a las que otorga cierto peso en la calificación definitiva.

En este estudio pretendemos analizar las características de los sistemas de evaluación que lleva a cabo el profesorado de diferentes centros de formación de futuros docentes y cuantificar cuál o cuáles son las estrategias presentes en mayor medida, así como de los instrumentos y sistemas de los que estos profesionales se sirve para calificar a sus alumnos. Señalamos, por último, como un objetivo más, el análisis de los procesos de evaluación entre iguales en la FIP.

OBJETIVOS DEL ESTUDIO

Podemos concretar los objetivos del estudio en:

1. Identificar las posibles tipologías de profesorado de los centros de Formación Inicial del Profesorado a partir de los instrumentos y técnicas de evaluación que utilizan.
2. Estudiar las características de las diferentes tipologías de profesorado encontradas.
3. Describir el uso de los diferentes instrumentos de evaluación y calificación.
4. Analizar la presencia de procesos de coevaluación y evaluación ente iguales en la FIP.

MATERIAL Y MÉTODO

Participantes

El estudio se ha llevado a cabo con una muestra de 217 docentes de las Facultades de Ciencias de la Actividad Física y el Deporte (47%) y las Escuelas de Magisterio (53%), de las universidades: Autónoma de Madrid, Valladolid, Granada, Zaragoza, Lleida, Autónoma de Barcelona, Salamanca, Huelva, Barcelona, León, La Coruña, Murcia, Católica de Murcia y Tenerife (Tabla 1).

Tabla 1. Distribución de la muestra por universidades

	FRECUENCIA	PORCENTAJE	PORCENTAJE VÁLIDO	PORCENTAJE ACUMULADO
Autónoma de Madrid	18	8,3	8,3	8,3
Valladolid	26	12,0	12,0	20,3
Granada	19	8,8	8,8	29,0
Zaragoza	16	7,4	7,4	36,4
Lleida	16	7,4	7,4	43,8
Autónoma de Barcelona	26	12,0	12,0	55,8
Salamanca	3	1,4	1,4	57,1
Huelva	13	6,0	6,0	63,1
Barcelona	6	2,8	2,8	65,9
León	24	11,1	11,1	77,0
La Coruña	12	5,5	5,5	82,5
Murcia	5	2,3	2,3	84,8
UCAM	18	8,3	8,3	93,1
Tenerife	15	6,9	6,9	100,0
Total	217	100,0	100,0	

La muestra presenta un predominio de hombres (63%) y con una edad media de 44 años.

Instrumentos

La toma de datos se realizó mediante una escala tipo Likert relacionada con el uso de instrumentos y procedimientos de evaluación, así como de la participación del alumnado en este proceso. Dado que se trata de medir la percepción de uso y no de una constatación real de lo realizado, la versión final de la escala ha sido denominada Escala de Percepción de Competencia de la Evaluación Formativa y Participación del Alumnado (EPCEF-PA). La versión final consta de 17 preguntas en las que el docente responde según el grado de competencia percibida en una escala tipo Likert de cinco puntos (valores de 0 a 4).

Para su construcción se han tenido en cuenta instrumentos ya validados con temática parecida (Prieto, 2005; Gutiérrez *et al.*, 2011; Palacios y López, 2012). En una primera fase se recogió una amplia muestra de preguntas sobre estos tres aspectos, que fueron evaluadas por distintos expertos, profesionales de la FIP. De todas ellas, fueron seleccionadas las más pertinentes por su relevancia (los ítems deberían estar claramente relacionados con el objeto de estudio) y claridad (fácilmente comprensibles, con afirmaciones simples) con las que se realizaron los primeros pre-test.

Para el cálculo de la fiabilidad de la EPCEF-PA se ha utilizado el Alfa de Cronbach, obteniendo un valor de 0,80 (17 ítems), que puede ser considerado un valor alto, lo que nos asegura que nuestras medidas y sus resultados sean fiables.

Por lo que respecta a la validez de contenido, se procedió a realizar un Análisis Factorial de Componentes Principales con los resultados que resumimos en la Tabla 2.

Tabla 2. Análisis Factorial de Componentes Principales

	COMPONENTE		FACTORES
	1	2	
Examen de preguntas cerradas	,737		F1
Examen de preguntas cortas	,736		Percepción de competencias de instrumentos de evaluación
Examen de preguntas abiertas	,675		
Exámenes orales	,668		
Exámenes escritos con documentos	,653		
Ensayos a partir de textos o materiales audiovisuales	,608		
Examen tipo test	,587		

	COMPONENTE		FACTORES
	1	2	
Informes o trabajos escritos	,535		F1
Cuadernos de campo	,469		Percepción de competencias de instrumentos de evaluación
Observación del profesor en clase	,460		
Pruebas prácticas de carácter físico	,452		
Control de la participación en el aula	.347		
Se califica a partir de la autoevaluación (parcial o totalmente)		,827	F2
El alumno se autocalifica (parcial o totalmente)		,801	Participación del alumnado
Se califica de forma dialogada y consensuada (entre profesorado y alumnado) (parcial o totalmente)		,799	
Se califica a partir de la coevaluación (entre compañeros) (parcial o totalmente)		,674	
Portafolios	,437	,522	
Autovalores	4,47	3,19	
% Varianza explicada	26,32	18,80	
% Varianza explicada acumulada	26,32	45,12	
KMO: ,80			
Prueba Esfericidad de Barlett: Chi-cuadrado aproximado: 1151,69 gl. 136; sig: ,000			

Método de extracción: Análisis de componentes principales. Método de rotación: Normalización Varimax con Kaiser.

Se procedió por último, al Análisis Factorial Confirmatorio (AFC) con los 17 ítems del cuestionario. Para su cálculo, se utilizaron la matriz de covarianzas y el algoritmo de cálculo de los mínimos cuadrados no ponderados, pues con dicha estrategia de cálculo se realizan estimaciones consistentes sin necesidad de asumir multinormalidad de las variables (Ruiz, 2000). Concretamente, se ha contrastado la hipótesis de la existencia de dos factores correspondientes a las dos vertientes del cuestionario antes mencionado (Tabla 3).

Tabla 3. Evaluación mediante AFC de modelo bifactorial de la EAEF-PA

MODELO	S-B _(CHI-CUADRADO) (GL) (P)	RMSEA	NFI	NNFI	CFI	AGFI	AIC
Dos factores	57,88 (43) (p = ,064)	,077	,90	,96	,87	,81	590,99

Como podemos comprobar, el modelo alcanza buenos ajustes. Además, la matriz LAMBDA-X del modelo presenta todos los valores significativos con $p < 0,05$, lo que proporciona evidencias añadidas de la validez de las preguntas de la escala. Un dato interesante a resaltar es la alta correlación de 0,79 entre ambos factores del cuestionario: las actitudes hacia la evaluación formativa y continua y la participación del alumnado en dicho proceso.

Como complemento a la escala anterior, se realizaron preguntas relativas a los centros de trabajo, la edad, el sexo y otros aspectos académicos de interés relacionados con la evaluación y los sistemas de calificación.

Procedimiento de recogida de información

Los cuestionarios fueron recogidos por el equipo de investigación y por profesores colaboradores durante los cursos académicos 2008/2009 y 2009/2010. Los cuestionarios tenían un carácter anónimo y fueron autocumplimentados por los sujetos de la muestra.

Análisis de la información

Los datos obtenidos se analizaron con el programa informático SPSS 18.0. Para el cálculo de la validez mediante Análisis Factorial Confirmatorio (AFC) se ha utilizado el programa estadístico LISREL 8.6. Dado que el objetivo del estudio es analizar las posibles diferencias entre tipologías de profesorado, para la obtención de dichos grupos homogéneos se utilizó un Análisis de Conglomerados (AC) mediante la opción de de clasificación no jerárquico (análisis de conglomerados de K-medias). Así mismo, las diferencias de percepción entre grupos se han realizados mediante ANOVAS en cada una de las preguntas de las escalas y en el resultado final de la misma.

RESULTADOS

Mediante el Análisis de Conglomerados (AC) se pretende establecer grupos homogéneos a partir de los resultados obtenidos en medida que sirva como criterios de asignación. En nuestro caso, la *variable criterio de clasificación* son los resultados de la escala EFCP-PA. En cuanto a la técnica utilizada dentro del AC se ha optado por un método de clasificación no jerárquico (análisis de conglomerados de K-medias) pues es la técnica que maximiza la varianza entre grupos y minimiza la varianza dentro de cada tipología.

De las soluciones posibles, se ha optado por aquella que establece tres conglomerados o grupos diferenciados pues es la que, además de cumplir con los criterios de máxima varianza entre conglomerados y mínima dentro de cada grupo, permite una caracterización más clara de los perfiles de los grupos. El primer clúster está formado por 61 profesores (28%), por 103 el segundo (48%) y por 52 en el tercero (24%).

Pasamos a continuación a caracterizar en cada uno de estos grupos de profesores tomando como referencia las puntuaciones de las preguntas de la escala, que recordamos posee dos factores principales correspondientes a la percepción de competencias en el uso de instrumentos de evaluación y a la participación del alumnado en dichas tareas.

Con relación a este último aspecto, encontramos diferencias significativas en los 4 ítems del factor, así como en el valor total de la escala (Tabla 4).

Tabla 4. Participación del alumnado en los procesos de evaluación (valores medios por conglomerados)

PREGUNTAS	MEDIAS POR CONGLOMERADO				ANOVA	
	1	2	3	MEDIA TOTAL	F	SIG.
Escala de Participación del alumnado (valor total)	8,96	2,67	6,31	5,32	106,82	,000
El alumno se autocalifica (parcial o totalmente)	2,57	,75	1,75	1,50	72,15	,000
Se califica de forma dialogada y consensuada (entre profesorado y alumnado) (parcial o totalmente)	2,54	,80	1,83	1,54	66,38	,000
Se califica a partir de la autoevaluación (parcial o totalmente)	2,07	,56	1,37	1,18	56,44	,000
Se califica a partir de la coevaluación (entre compañeros) (parcial o totalmente)	1,84	,57	1,38	1,12	35,75	,000

En relación al total de la escala, observamos que el primero de los conglomerados puntúa ostensiblemente por encima del segundo y del tercero; el primer grupo de profesores utiliza con mayor frecuencia que los otros dos conglomerados los sistemas de evaluación y calificación en los que el alumnado participa de manera activa, tanto a través de la autocalificación como de la calificación dialogada o de la coevaluación. El segundo conglomerado, por el contrario, apenas utiliza sistema alguno en el que participe el alumno. El tercer conglomerado utiliza tanto la autocalificación como la calificación dialogada y, en menor medida, la coevaluación, pero con valores medios menores que el primer conglomerado y claramente mayores que el segundo.

Las mayores diferencias entre las preguntas se obtienen ante la posibilidad de que el alumnado pueda autocalificarse parcial o totalmente; en el segundo conglomerado el valor medio sitúa esta posibilidad de manera inexistente o nula (por debajo del valor 1 correspondiente a casi nunca o poco). Por el contrario, el valor medio del primer conglomerado se sitúa por encima de 2,5 correspondiente a una alta utilización de dicha técnica.

Las diferencias entre grupos que acabamos de comentar vuelve a aparecer al respecto de si es el profesor el único que decide la calificación a partir de su evaluación (Tabla 5). La práctica totalidad del profesores de segundo grupo se consideran a sí mismos como determinantes exclusivos de la calificación; valor que sigue siendo alto o en el tercero de los grupo y bastante menor en el primero.

Tabla 5. Papel del profesorado en la calificación por conglomerados

PREGUNTAS	MEDIAS POR CONGLOMERADO				ANOVA	
	1	2	3	MEDIA TOTAL	F	SIG.
La calificación la decide el profesorado a partir de su evaluación	2,20	3,51	3,12	3,05	48,548	,000

En relación a la *Escala de Percepción de Competencias* y concretamente al uso de instrumentos y procedimientos de evaluación, el valor medio obtenido permite concluir que el profesorado en términos generales se percibe como bastante competente (media de 33 puntos sobre una escala de 0 a 50). El mayor valor y, por tanto la mayor percepción de competencia, se obtiene en la evaluación de los informes o trabajos escritos y de los exámenes de preguntas cortas. En el extremo

contrario, entre los que se perciben como menos competentes están los exámenes tipo test, las pruebas prácticas de carácter físico y el uso de los portafolios.

Las diferencias vuelven a ser significativas al comparar los diferentes conglomerados de profesores en relación a la percepción de competencias en el uso de instrumentos de evaluación (Tabla 6). Concretamente, la percepción de competencia en el desarrollo y elaboración de exámenes tipo test es significativamente mayor en el segundo conglomerado que en el primero; situación que se vuelve a repetir con los exámenes de preguntas abiertas y de preguntas cortas. Por el contrario, el conglomerado primero puntúa significativamente en mayor medida en los cuadernos de campo y, sobre todo, en el uso de portafolios. El tercer conglomerado destaca significativamente sobre los demás en las pruebas prácticas de carácter físico.

Tabla 6. Percepción de competencias en el uso de instrumentos de evaluación (por conglomerados)

PREGUNTAS DEL CUESTIONARIO	MEDIAS POR CONGLOMERADO				ANOVA	
	1	2	3	TOTAL	F	SIG.
Total escala de percepción de competencias	33,48	31,90	35,30	33,16	2,46	,080
Observación del profesor en clase	3,19	2,21	2,88	2,65	2,118	,123
Control de la participación en el aula	2,85	2,66	2,92	2,78	1,969	,142
<i>Examen tipo test</i>	1,82	2,58	2,35	2,32	5,281	,006
<i>Examen de preguntas abiertas</i>	2,69	3,09	2,90	2,93	3,474	,033
<i>Examen de preguntas cortas</i>	2,75	3,13	3,10	3,02	3,449	,034
P7f. Examen de preguntas cerradas	2,28	2,64	2,76	2,57	2,096	,126
P7g. Exámenes escritos con documentos	2,60	2,41	2,57	2,50	,490	,613
P7h. Exámenes orales	2,66	2,26	2,53	2,43	2,354	,098
<i>Pruebas prácticas de carácter físico</i>	2,16	2,18	2,96	2,37	6,457	,002
<i>Portafolios</i>	2,95	1,73	2,67	2,31	21,103	,000
<i>Cuadernos de campo</i>	2,69	2,23	2,52	2,43	2,939	,045
Informes o trabajos escritos	3,25	3,15	3,13	3,17	,380	,684
Ensayos a partir de textos o materiales audiovisuales	2,77	2,76	2,83	2,78	,071	,931

El mayor número de diferencias estadísticas significativas entre los conglomerados aparece en la *frecuencia real de uso* de instrumentos de evaluación (Tabla 7).

Tabla 7. Frecuencia de uso de instrumentos de evaluación (por conglomerados)

PREGUNTAS DEL CUESTIONARIO	MEDIAS POR CONGLOMERADO				ANOVA	
	1	2	3	TOTAL	F	SIG.
<i>Observación del profesor en clase</i>	2,13	1,58	2,71	2,01	15,931	,000
<i>Control de la participación en el aula</i>	2,77	2,37	2,96	2,63	6,831	,001
<i>Examen tipo test</i>	,52	1,83	1,69	1,43	18,108	,000
Examen de preguntas abiertas	1,85	2,27	2,22	2,14	1,930	,148
<i>Examen de preguntas cortas</i>	1,48	2,35	2,42	2,12	11,512	,000
<i>Examen de preguntas cerradas</i>	,28	,90	1,34	,82	16,692	,000
<i>Exámenes escritos dejando disponer de documentos</i>	1,16	,86	1,43	1,08	3,755	,025
Exámenes orales	1,20	,84	,92	,96	1,836	,162
<i>Pruebas de carácter físico</i>	,90	1,10	2,68	1,42	27,714	,000
<i>Portafolios</i>	2,74	,84	2,39	1,75	48,910	,000
<i>Cuadernos de campo</i>	2,18	1,01	2,00	1,59	17,214	,000
Informes o trabajos escritos	3,20	2,91	2,77	2,96	2,734	,067
Ensayos a partir de textos escritos o materiales audiovisuales	2,36	1,92	2,21	2,11	2,092	,126

Ahora las diferencias significativas están presentes en nueve de las trece preguntas relacionadas con el uso de instrumentos de evaluación. El primero de los conglomerados presenta mayores frecuencias de uso en los cuadernos de campo y, sobre todo, en el uso de portafolios que, recordamos, era el instrumento con mayor percepción de competencia en este grupo. El conglomerado segundo destaca significativamente en el uso de los exámenes tipo test (siendo también el grupo que se siente más competente en dicho instrumento) y presenta valores muy bajos en el uso de los portafolios, los cuadernos de campo y los exámenes escritos con documentación. El tercero de los grupos se presenta como el más ecléctico, al destacar sig-

nificativamente en seis de los nueve ítems con diferencias significativas. Ordenados de mayor a menor puntuación son los siguientes: 1) control de participación en el aula, 2) observación del profesor en clase, 3) pruebas de carácter físico, 4) examen de preguntas cortas, 5) exámenes escritos dejando disponer de documentos y 6) examen de preguntas cerradas.

Cabe resaltar que el instrumento más usado por todos los docentes es el informe o trabajo escrito, seguido del control de la participación en el aula y los exámenes de preguntas abiertas. Entre los instrumentos menos usados destacan los exámenes de preguntas cerradas y los exámenes orales. Los ítems en los que no aparecen diferencias significativas entre los tres grupos son los siguientes: los informes o trabajos escritos, exámenes de preguntas abiertas, ensayos a partir de textos escritos o materiales audiovisuales y exámenes orales.

Terminamos la descripción de los valores medios de cada conglomerado analizando datos relativos a la cantidad de docencia semanal y a sus características así como a la edad y el sexo de los entrevistados.

Con relación al número de horas que se imparte a la semana, encontramos diferencias significativas estadísticas en el valor medio en cada conglomerado. Concretamente, el segundo presenta una carga semanal de horas significativamente mayor que los otros dos. No existen diferencias significativas en la media de las horas semanales de clases teóricas pero sí las encontramos con la carga de clases prácticas, significativamente mayores en el segundo y tercer conglomerado (Tabla 8).

Tabla 8. Carga horaria semanal por conglomerado

PREGUNTAS DEL CUESTIONARIO	MEDIAS POR CONGLOMERADO				ANOVA	
	1	2	3	TOTAL	F	SIG.
<i>Número de horas que impartes a la semana</i>	6,87	8,47	7,06	7,68	3,97	,020
Número de horas de clases teóricas	3,36	3,87	2,96	3,51	2,046	,132
<i>Número de horas de clases prácticas</i>	3,59	4,91	4,04	4,31	3,47	,033

Tampoco encontramos valores medios significativos en la edad y la pertenencia a uno u otro conglomerado (Tabla 9).

Tabla 9. Edad por conglomerados

	MEDIAS POR CONGLOMERADO				ANOVA	
	1	2	3	TOTAL	F	SIG.
Edad	44,56	44,33	44,67	44,48	,027	,973

No obstante, la distribución de los intervalos de edad por grupos presenta algunos datos interesantes. Así, el primer conglomerado tienen en la franja de los 41 a los 50 los mayores efectivos, llegando a presentar este intervalo de edad el 50% del profesorado de este grupo. En el conglomerado segundo, apenas hay profesores de más de 60 años, siendo llamativo el alto porcentaje (1 de cada 3) en la franja de los 31 a los 40. El conglomerado tercero es el que presenta mayores porcentajes de los más jóvenes así como los de más edad.

En relación al sexo, los porcentajes de profesores y profesoras serían muy parecidos en el primer conglomerado, una mayor presencia de hombres en el segundo y de mujeres en el tercero (Tabla 10).

Tabla 10. Distribución del sexo por conglomerados

			GRUPO DE PERTENENCIA			
			1	2	3	TOTAL
Sexo	Hombre	Recuento	39	73	24	136
		% dentro de Sexo	28,7%	53,7%	17,6%	100,0%
	Mujer	Recuento	22	30	28	80
		% dentro de Sexo	27,5%	37,5%	35,0%	100,0%
Total		Recuento	61	103	52	216
		% dentro de Sexo	28,2%	47,7%	24,1%	100,0%

Para facilitar una visión de conjunto presentamos a continuación una tabla resumen con los valores más interesantes comentados en los párrafos anteriores, sobre las características básicas de cada conglomerado (Tabla 11).

Tabla 11. Características asociadas a cada uno de los conglomerados

CONGLOMERADO 1	CONGLOMERADO 2	CONGLOMERADO 3
Igual número de profesoras que de profesores	Mayor presencia de profesores	Mayor presencia de profesoras (el doble)
Grupo de edad de los “maduros”	Profesorado más joven	Profesorado de mayor edad
Menor carga lectiva semanal	Mayor carga lectiva semanal	Carga lectiva semanal media
Menor porcentaje de horas de clases prácticas	Mayor porcentaje de horas de clases prácticas	Porcentaje medio de horas de clases prácticas
Elevada participación del alumnado en la evaluación y calificación	Escasa o nula participación del alumnado en tareas de evaluación	Participación media del alumnado en tareas de evaluación
El profesor no es el único que decide la calificación	El profesor es el único que decide la calificación	El profesor es el más importante pero no el único que decide la calificación
El profesorado muestra diferencias significativamente mayores en la percepción de competencia sobre el uso de portafolios y cuadernos de campo	El profesorado muestra diferencias significativamente mayores en la percepción de competencia sobre el uso de exámenes de preguntas abiertas y de preguntas cortas así como en los exámenes tipo test	El profesorado muestra diferencias significativamente mayores en la percepción de competencia de pruebas prácticas de carácter físico.
Los instrumentos de evaluación en que obtienen resultados significativamente mayores que los otros grupos son portafolios y cuadernos de campo	El único instrumentos de evaluación en que obtiene resultados significativamente mayores es el examen tipo test.	Obtiene resultados significativamente mayores que los otros grupos en seis instrumentos: control de participación en el aula, observación del profesor en clase, pruebas de carácter físico, examen de preguntas cortas, exámenes escritos dejando disponer de documentos y examen de preguntas cerradas

CONGLOMERADO 1	CONGLOMERADO 2	CONGLOMERADO 3
Docentes que favorecer la implicación del alumnado en la evaluación y utilizan instrumentos más cercanos a una evaluación orientada al aprendizaje	Docentes que utilizan instrumentos de evaluación más cercanos a los modelos tradicionales, basados en la realización de exámenes finales y test.	Docentes eclécticos, con una gran diversidad de instrumentos de evaluación y que favorecen en parte la participación del alumnado en la evaluación. Los instrumentos utilizados tienen un fuerte carácter práctico.

DISCUSIÓN

Mediante las técnicas de análisis multivariantes utilizadas se han obtenido tres perfiles de docentes universitarios en los estudios de FIP, diferenciados según la participación del alumnado en la evaluación y las técnicas e instrumentos de evaluación más utilizados en sus asignaturas.

El profesorado del primer conglomerado muestra una actitud positiva hacia el desarrollo de modelos de evaluación alternativa, cuyas características principales son:

- Favorecer la participación e implicación del alumnado en los procesos de evaluación y calificación.
- Utilizar técnicas de evaluación variadas y orientadas al aprendizaje, recopiladas normalmente a través de portafolios, así como cuadernos de campo.
- También son los que se sienten más competentes en el uso de estos instrumentos de evaluación.

Por tanto, al primer conglomerado podríamos denominarlo: “*Profesorado Innovador*”. Uno de cada cuatro docentes se ajusta a este perfil (28%), con un porcentaje similar de hombres y mujeres respecto al total por cada sexo (28,7 y 25,5%, respectivamente). La mayor parte del profesorado de este conglomerado está en el intervalo de 41-50 años y muestra la carga lectiva media más baja (6.87 horas), significativamente menor que los otros dos grupos, y también la menor carga de horas prácticas. Estas características del *Profesorado innovador* corresponden en gran medida con las encontradas en algunos trabajos recientes sobre el desarrollo de sistemas de evaluación formativa y compartida en FIP (López, 2009; López, Manrique y Vallés, 2011; Navarro y González, 2010; Santos, Martínez y López, 2009), espe-

cialmente en lo relativo a la participación del alumnado en la evaluación, las dinámicas colaborativas de aprendizaje y en el uso del portafolios.

El segundo conglomerado está compuesto mayoritariamente por profesorado que obtiene puntuaciones muy bajas en los aspectos relacionados con la implicación del alumnado en los procesos de evaluación y calificación, y bastante altos en los que hacen referencia a la realización de sistemas de evaluación tradicionales (exámenes tipo test o de preguntas cortas). Por todo ello, al segundo conglomerado podríamos denominarlo *Profesorado orientado a la calificación*. La mitad del profesorado se ajusta a este perfil (48%) con bastante más porcentaje de hombres que de mujeres (53,7 frente a 37,5%). La mayoría del profesorado está en el intervalo de 31-50 años, aunque también hay un porcentaje importante en el intervalo 51-60, y presenta una carga semanal de horas significativamente mayor que los otros dos grupos (8.47). Las características del *Profesorado orientado a la calificación* corresponden en gran medida con las encontradas en algunos trabajos recientes sobre los sistemas e instrumentos de evaluación y calificación que predominan actualmente en la formación inicial del profesorado (Gutiérrez *et al*, 2011; Palacios y López, 2012; Trillo y Porto, 1999).

El profesorado del tercer conglomerado tiene unas características muy especiales, que dificultan su clasificación y denominación. En muchas variables se sitúa entre los dos conglomerados anteriores, pues si bien muestra una actitud favorable hacia la participación del alumnado en los procesos de evaluación y calificación, sigue utilizando en gran medida los diferentes tipos de examen final (tipo test, preguntas cortas, con material, etc.), que combina con otros procedimientos para llegar a la calificación definitiva. Existen algunas características curiosas de este conglomerado que nos hacen pensar que puede estar agrupando en un gran porcentaje al profesorado de asignaturas con un gran componente de práctica motriz. Por ejemplo, el elevado peso que se da a la realización de pruebas de carácter físico, a la participación del alumnado en el aula y a la utilización de portafolios y cuaderno de campo. Por otra parte, se trata del grupo que muestra valores mayores que los demás en seis de los nueve ítems con diferencias significativas: control de participación en el aula, observación del profesor en clase, pruebas de carácter físico, examen de preguntas cortas, exámenes escritos dejando disponer de documentos y examen de preguntas cerradas. Otra característica del grupo va en la misma línea, dado que este grupo tiene una carga lectiva intermedia, aunque la carga de horas prácticas es considerablemente mayor que la de horas teóricas. Por tanto, una posible denominación sería la de *Profesorado ecléctico de asignaturas de práctica motriz*. Poco menos de uno de cada cuatro docentes se ajusta a este perfil (24%), siendo el único grupo en el que predominan las mujeres sobre los hombres (17,6%

frente al 35%). La mayoría del profesorado de este conglomerado está en el intervalo de 31-50 años, aunque también hay un porcentaje importante en el intervalo 51-60. No hemos encontrado ningún trabajo que guarde similitud con las características globales del profesorado de este conglomerado, aunque sí con algunas de sus características por separado. Por ejemplo, en lo relativo a la realización de pruebas prácticas de carácter físico como forma de llegar a la calificación en algunas asignaturas de FIP, puede encontrarse algún trabajo que critican su realización, tanto por sus efectos negativos en el aprendizaje del alumnado, como por la grave incoherencia pedagógica que supone (López, Pérez y García, 2001); así como en relación a la importancia dada a fomentar la implicación del alumnado en la evaluación en FIP (Fernández-Balboa, 2006; López, 2009; Navarro y González, 2010).

En términos generales, puede observarse una predominancia de perfiles de docentes orientados a la calificación y ecléctico en la FIP, aunque se observa un crecimiento notable del perfil innovador respecto a estudios genéricos realizados en la última década en la formación inicial del profesorado (Trillo y Porto, 1999), en línea con lo detectado en los últimos años (Gutiérrez *et al*, 2011; López, Manrique y Vallés, 2011; Palacios y López, 2012; Ibarra y Rodríguez, 2010), quizás influido por el proceso de Convergencia hacia el EEES, que da una mayor importancia a los procesos de evaluación orientados al aprendizaje.

La edad no parece tener demasiada relación con las actitudes hacia la evaluación formativa. El sexo en cambio sí parece tener influencia. En el conglomerado de profesorado innovador se da la misma proporción entre hombres y mujeres, mientras que en el orientado a la calificación predominan los hombres y en el ecléctico predominan las mujeres.

Es interesante señalar la baja utilización de la calificación a partir de la evaluación entre iguales (coevaluación) en todos los grupos, incluido el primer conglomerado de docentes; resultado que coincide con lo que puede encontrarse en la literatura especializada, que aconseja utilizar la coevaluación como una técnica muy potente de evaluación orientada al aprendizaje, pero que resulta ser una técnica con fuertes inconvenientes cuando se la utiliza como sistema de calificación (López, 2009; McMahon, 2010). Todo el profesorado tiene una percepción de competencia muy alta en la utilización de informes o trabajos escritos, que coincide también con una alta utilización de este instrumento en los tres conglomerados. Estos aspectos específicos pueden ser una explicación de porqué se da una utilización variada de técnicas e instrumentos de evaluación en los tres conglomerados, así como algunas de las coincidencias encontradas en los tres grupos (utilización muy alta de informes o trabajos escritos, exámenes de preguntas abiertas y realización de ensayos, y muy escasa de exámenes orales). La similitud de edad en los tres conglome-

rados indica que la edad no parece ser un factor que afecte a la forma de organizar los sistemas de evaluación y calificación.

Nuestros datos y conclusiones permiten vislumbrar ciertos caminos a seguir para mejorar la situación de la evaluación formativa en la universidad en general y en la formación inicial del profesorado en particular. Tema este último de enorme interés, pues en dicha formación inicial la evaluación formativa y compartida es un contenido de aprendizaje, pero también una competencia profesional a adquirir, así como un elemento clave en los procesos de innovación educativa en los niveles no universitarios.

Independientemente de la importancia que para el tema podría tener la modificación de la normativa reguladora de la evaluación universitaria, es a través de la formación como se perfila una de las vías de mejora y de acercamiento a modelos evaluativos centrados en los procesos y no tanto en los productos. Nuevas formas de aprender deben precisar nuevas maneras de evaluar, que han de ser conocidas, comprendidas y trabajadas por el profesor universitario. No obstante, la formación por sí sola no asegura un cambio verdadero si no está incardinada en el trabajo del día a día. Por ello, parece importante que se generen proyectos y grupos de innovación docente en la FIP (a través, por ejemplo, de ciclos de investigación-acción), dado que parece ser la dinámica de formación permanente que está siendo más eficaz a la hora de desarrollar sistemas de evaluación más formativos y participativos en la FIP.

Por último, quisiéramos resaltar algunas limitaciones del presente trabajo y las líneas futuras de investigación. Entre las primeras, cabe mencionar lo limitado de la muestra y la imposibilidad de hacer comparaciones con otros colectivos docentes no dedicados a la FIP. Respecto a las propuestas de mejora, parece importante que se generen proyectos y grupos de innovación docente en la FIP, dado que parece ser la dinámica de formación permanente que está siendo más eficaz a la hora de desarrollar sistemas de evaluación más formativos y participativos en la FIP. Entre las líneas futuras de investigación sería recomendable profundizar en las posibles relaciones que estos perfiles pudieran tener con otros aspectos de la docencia, como la metodología o el tipo de competencias profesionales.

Fecha de recepción del original: 19 de septiembre de 2012

Fecha de aceptación de la versión definitiva: 10 de diciembre de 2012

REFERENCIAS

- Arribas, J. M. (2012). El rendimiento académico en función del sistema de evaluación empleado. *Revista Electrónica de Investigación y Evaluación Educativa*, 18(1), 1-15. Extraído el 22 de marzo de 2013 de http://www.uv.es/RELIEVE/v18n1/RELIEVEv18n1_3.pdf
- Black, P. y Williams, D. (1998). Assessment and classroom learning. *Assessment in Education*, 4(1), 7-71.
- Bonsón, M. y Benito, A. (2005). Evaluación y Aprendizaje. En A. Benito y A. Cruz, (Eds.), *Nuevas claves para la docencia universitaria en el Espacio Europeo de Educación Superior* (pp. 87-100). Madrid: Narcea.
- Boud, D. y Falchikov, N. (2007). *Rethinking Assessment in Higher Education. Learning for the long term*. Oxon: Routledge.
- Brown, S. y Glasner, A. (2003). *Evaluar en la universidad. Problemas y nuevos enfoques*. Madrid: Narcea.
- Carless, D., Joughin, G. y Mok, M. M. C. (2006). Learning-oriented assessment: principles and practice. *Assessment & Evaluation in Higher Education*, 31(4), 395-398.
- Dochy, F., Segers, M. y Dierick, S. (2002). Nuevas vías de aprendizaje y enseñanza y sus consecuencias: una era de evaluación. *Revista de Docencia Universitaria*, 2(2), 13-30. Extraído el 22 de marzo de 2013 de <http://revistas.um.es/redu/article/view/20051/19411>
- Fernández-Balboa, J. M. (2006). Dignity and democracy in the college classroom: The practice of self-evaluation. En A. Goldstein (Ed), *Useful Theory: Making Critical Education Practical* (pp. 105-130). New York: Peter Lang Publishing.
- Gutiérrez, C., Pérez, A., Pérez, M. y Palacios, A. (2011). Percepciones de profesores y alumnos sobre la enseñanza, evaluación y desarrollo de competencias en estudios universitarios de formación del profesorado. *Cultura y Educación*, 23(4), 499-514.
- Hamodi, C. y López, A. T. (2012). La evaluación formativa y compartida en la Formación Inicial del Profesorado desde la perspectiva del alumnado y de los egresados. *Psychology, Society, & Education*, 4(1), 103-106.
- Ibarra, M. S. y Rodríguez, G. (2010). Aproximación al discurso dominante sobre la evaluación del aprendizaje en la universidad. *Revista de Educación*, 351, 385-407.
- López, V. M. (Coord.) (2009). *La Evaluación Formativa y Compartida en Docencia Universitaria: propuestas, técnicas, instrumentos y experiencias*. Madrid: Narcea.
- López, V. M., Manrique, J. C. y Vallés, C. (2011). La evaluación y la calificación en los nuevos estudios de Grado. Especial incidencia en la Formación Inicial del

- Profesorado. *Revista Electrónica Interuniversitaria de Formación del Profesorado*, 14(4), 57-72. Extraído el 22 de marzo de 2013 http://www.aufop.com/aufop/uploaded_files/articulos/1324595888.pdf
- López, V. M., Pérez, D. y García, A. (2001). La capacidad físico-deportiva ¿Reflejo directo de la capacidad profesional?: Algunas reflexiones sobre las ilógicas prácticas evaluativas en la formación y selección del profesorado de Educación Física. *Tandem*, 5, 77-89.
- López-Pastor, V., Pintor, P., Muros, B. y Webb, G. (2012). Formative assessment strategies and their effect on student performance and on student and tutor workload: the results of research projects undertaken in preparation for greater convergence of universities in Spain within the European Higher Education Area (EHEA). *Journal of Further and Higher Education*, 37(2), 163-180.
- McMahon, T. (2010). Peer feedback in a undergraduate programme: using action research to overcome students' reluctance to criticise. *Educational Action Research*, 18(2), 273-288.
- Navarro, I. y González, C. (2010). La autoevaluación y la evaluación entre iguales como estrategia para el desarrollo de competencias profesionales. Una experiencia docente en el grado de maestro. *Revista de Docencia Universitaria*, 8(1), 181-193.
- Nieto, S. (2000). El discurso del profesorado universitario sobre la evaluación del aprendizaje de los alumnos como estrategia de innovación y cambio profesional: exposición y análisis de una experiencia. *Revista de Educación*, 322, 305-324.
- Palacios, A. y López, V. M. (2013). Haz lo que yo digo pero no lo que yo hago: sistemas de evaluación del alumnado en la formación inicial del profesorado. *Revista de Educación*, 361, 1-16. Extraído el 22 de marzo de http://www.revistaeducacion.mec.es/doi/361_143.pdf
- Prieto, L. (2005). *Las creencias de autoeficacia docente del profesorado universitario*. Tesis Doctoral inédita. Madrid: Universidad Pontificia Comillas de Madrid.
- Ruiz, M. A. (2000). *Introducción a los modelos de ecuaciones estructurales*. Madrid: UNED.
- Santos, M., Martínez, L. F. y López, V. M. (Coords.) (2009). *La Innovación docente en el Espacio Europeo de Educación Superior*. Almería: Universidad de Almería.
- Trillo, F. y Porto, N. (1999). La percepción de los estudiantes sobre su evaluación en la Universidad. Un estudio de caso en la Facultad de Ciencias de la Educación. *Revista de Innovación Educativa*, 9, 55-75.

- Trillo, F. y Porto, N. (2002). La evaluación de los estudiantes en el marco de la evaluación de la calidad de las universidades. *Revista de Educación*, 328, 283-301.
- Zabalza, M. (2003). *Diseño curricular en la universidad. Competencias del docente universitario*. Madrid: Narcea.