



Percepciones de alumnos, egresados y profesores sobre los sistemas de evaluación del aprendizaje*



Carolina Hamodi Galán

Profesora Asociada del Departamento de Sociología y Trabajo Social de la Universidad de Valladolid
carolinahamodi@soc.uva.es



Ana Teresa López Pastor

Profesora Titular del Departamento de Sociología y Trabajo Social de la Universidad de Valladolid
anateresa@soc.uva.es



Víctor Manuel López Pastor

Profesor Titular del Departamento de Didáctica de la Expresión Musical, Plástica y Corporal de la Universidad de Valladolid
vlopez@mpc.uva.es

| Fecha presentación: 30/09/2014 | Aceptación: 10/02/2015 | Publicación: 22/06/2015

Resumen

Este estudio pretende conocer qué tipo de evaluaciones se están llevando a cabo en un centro de Formación Inicial del Profesorado (FIP) y si se aproximan a lo que se entiende por *sistemas de evaluación formativa y compartida*. Posteriormente se analiza si existen diferencias estadísticamente significativas en las percepciones de alumnos, egresados y profesorado sobre los sistemas de evaluación del aprendizaje llevados a cabo. Para ello se aplica un cuestionario validado con una escala tipo Likert. Se realiza ANOVA, mediante SPSS.V20.0. Los resultados indican que en la mayoría de ítems del cuestionario se encuentran diferencias significativas entre la opinión del colectivo del profesorado y los colectivos de alumnado y egresados. En tan solo un ítem (referente a si se utilizaban exámenes orales) las diferencias son entre el colectivo del alumnado y los colectivos de profesorado y egresados.

Palabras clave: evaluación del aprendizaje, formación inicial del profesorado, evaluación formativa, Educación Superior, evaluación compartida

Resum

Aquest estudi pretén conèixer quin tipus d'avaluacions s'estan duent a terme en un centre de Formació Inicial del Professorat (FIP) i si s'aproximen al que s'entén per *sistemes d'avaluació formativa i compartida*. Posteriorment s'analitza si hi ha diferències estadísticament significatives en les percepcions d'alumnes, egressats i professorat sobre els sistemes d'avaluació de l'aprenentatge duts a terme. Per a això s'aplica un qüestionari validat amb una escala tipus Likert. Es realitza ANOVA, mitjançant SPSS.V20.0. Els resultats indiquen que en la majoria d'ítems del qüestionari, es troben diferències significatives entre l'opinió del col·lectiu del professorat i els col·lectius d'alumnat i titulats.

Paraules clau: avaluació de l'aprenentatge; formació inicial del professorat, avaluació formativa, Educació Superior, avaluació compartida

Abstract

This research aims to identify the types of assessments that are being carried out in a Pre-service Teacher Education (PTE) center, and if that assessment is close to what is meant by *formative and shared assessment systems*. Subsequently we analyze whether there are statistically significant differences in the students, graduates and teachers' perceptions of the learning assessment systems implemented. A validated questionnaire was used, with a Likert type scale. ANOVA was carried out by SPSS.V20.0. The results indicate that in most of the questionnaire items there are significant differences between the opinion of the group of teachers and the groups of students and graduates.

Key words: learning assessment, pre-service teacher education, formative assessment, Higher Education, shared assessment

* En base al Proyecto I+D+i: "Las competencias docentes en la formación inicial del profesorado de educación física". Convocatoria de noviembre de 2013 del Programa Estatal de Investigación, Desarrollo e Innovación Orientada a los Retos de la Sociedad, en el marco del Plan Estatal de Investigación Científica y Técnica y de Innovación 2013-2016. Referencia: EDU 2013-42024-R. Duración: 3 años (2013-2016).



1. Introducción

La evaluación es uno de los elementos claves en la Formación Inicial del Profesorado (FIP) y debe evolucionar en el contexto del EEES. Es necesario reflexionar sobre el uso de la evaluación, puesto que condiciona el qué y cómo aprende el alumnado, según afirman autores como Álvarez (2001), Biggs (2005), Bonsón y Benito (2005), Boud (1988), Casanova (2012), Dochy, Segers y Dierick (2002), Gibbs (2003), López (2009), Race (2003), Rodríguez e Ibarra (2011), Sanmartí (2007) o Santos (2003), hasta el punto de tener una “incidencia radical” (Zabalza, 2007, p. 144) en la estructura curricular, siendo “el arma más poderosa que tienen los profesores para influir en el modo en el que los estudiantes responden a las asignaturas y se comportan como alumnos” (Gibbs, 2003: 61).

Diferentes estudios señalan la evaluación formativa y compartida como aquella que incide positivamente en el logro de mayores niveles de aprendizaje en el alumnado. Entre ellos destacamos los de Biggs (2005), Black y Wiliam (2006), Boud y Falchikov (2007), Brown y Glasner (2003), Dochy et al. (2002), Falchikov (2005), Gauntlett (2007), Gibbs (2003), Knight (2005), Lapahan y Webster (2003), McDowell y Sambell (2003) o Race (2003).

Pérez, Julián y López (2009: 35) definen la evaluación formativa como “todo proceso de constatación, valoración y toma de decisiones cuya finalidad es optimizar el proceso de enseñanza-aprendizaje que tiene lugar, desde una perspectiva humanizadora y no como mero fin calificador”. Esta definición contrasta con las de otros autores. Por ejemplo, Brown y Pickford (2013) la definen como el procedimiento utilizado para reconocer y responder al aprendizaje del estudiante con el fin de reforzarlo durante el propio proceso. Un elemento característico, sin el cual no sería formativa, es la retroalimentación, que consiste en ofrecer comentarios al alumnado sobre lo que ha hecho y sugerencias sobre lo que podría mejorar en la siguiente frase del proceso de aprendizaje. Es decir, no sirve solamente con la “postinformación”, sino que la “preinformación” es crucial en ese proceso de aprendizaje del alumnado mediante la evaluación formativa. Estas autoras cuentan cómo el ex director de la Open University de Escocia describió la evaluación formativa como el motor que impulsa el aprendizaje y ellas añaden a dicha explicación “que la retroinformación [...] es el aceite que lubrica los engranajes de la comprensión” (Brown y Pickford, 2013: 23-24).

La evaluación compartida se centra en una serie de “procesos dialógicos que mantiene el profesorado con el alumnado sobre la evaluación de los procesos de enseñanza-aprendizaje que se han dado. Estos diálogos pueden ser individuales o grupales” (Pérez, Julián y López, 2009: 37). De esta manera, se trata de un sistema en el que se implica al alumnado mediante su participación en los procesos de evaluación de su aprendizaje.

Los trabajos de investigación llevados a cabo respecto a las diferencias y similitudes entre la percepción del alumnado y el profesorado en torno a la evaluación formativa del aprendizaje ofrecen datos significativos. Arribas, Carabias y Monreal (2010) llevan a cabo una investigación sobre las percepciones del alumnado, profesorado y egresados de un centro de FIP sobre la metodología y los sistemas de evaluación experimentados durante su formación. Los resultados de esta investigación parecen indicar que existen grandes diferencias en casi todos los ítems entre la percepción del profesorado y del alumnado (incluyendo a egresados), sobre todo en los que hacían referencia a la discusión del pro-

grama al comienzo del curso, la utilización de una evaluación formativa y continua, la explicación previa de los criterios de evaluación y la utilización de la retroalimentación tras los exámenes.

En una línea similar, Gutiérrez et al. (2011) llevan a cabo un estudio que analiza las diferencias de percepción entre profesores y alumnos sobre la enseñanza, evaluación y desarrollo de competencias en estudios universitarios de FIP. Los resultados indican que existen claras discrepancias entre las percepciones de profesores y alumnos en lo referido a la evaluación y a la calificación. Por ejemplo, mientras el alumnado percibe que la calificación se suele reducir al examen final, el profesorado afirmaba que consideraba todo el conjunto de trabajos realizados para llegar a la calificación final. Los escasos puntos de acuerdo giran respecto a la escasa participación del alumnado en los procesos de evaluación y calificación, así como en la escasa utilización de procedimientos de evaluación formativa en FIP, como los portafolios individuales y grupales. En un estudio posterior del mismo grupo de investigación (Gutiérrez et al., 2013) se estudiaron las percepciones de profesores, estudiantes y egresados de un centro de FIP. Los resultados muestran percepciones similares entre alumnos y egresados, e importantes diferencias entre éstos y los profesores. Las mayores diferencias se centraron en: (1) coherencia de programas y sistemas de evaluación; (2) utilización de la evaluación con fines formativos; (3) presencia de capacidades cognitivas en la evaluación; (4) posibles causas de desarrollo de evaluaciones inadecuadas. Se encontraron menores diferencias en los instrumentos de evaluación y notables semejanzas en los procesos de calificación. Los autores concluyen que las percepciones del profesorado son más próximas a procesos de evaluación formativa, mientras que alumnos y egresados tienden a percibir evaluaciones más tradicionales.

En otro estudio similar, Taberero y Daniel (2012) analizaron los sistemas de evaluación aplicados durante la FIP en las cuatro universidades públicas de Castilla y León, estudiando las diferencias de percepción entre alumnado y profesorado. Los resultados apuntan a que, en general, los sistemas de evaluación son tradicionales y que el examen seguía siendo determinante en la calificación final, bien porque fuera el único medio de evaluación o bien porque, a pesar de que se complementase con otros medios, seguía teniendo un mayor peso en la calificación final. Respecto a las diferencias de percepción entre alumnado y profesorado, el profesorado tiende a considerar que los sistemas de evaluación que llevan a cabo son más negociados, formativos, continuos y variados de lo que percibe el alumnado. En un estudio similar llevado a cabo en la comunidad de Madrid, Martínez et al. (2012) encuentra también diferencias considerables entre la perspectiva del profesorado y del alumnado. A grandes rasgos, los estudiantes reclamaban una evaluación más práctica, coherente y sencilla, mientras que los profesores buscaban un mayor compromiso de los estudiantes y de la institución.

En López y Toja (2013) se analizaron las similitudes y diferencias de percepción y concepciones entre alumnado y profesorado, de otro centro de FIP de la Comunidad Gallega, con respecto a diferentes conceptos relacionados con la evaluación. Los resultados son similares a otros estudios, con diferencias estadísticamente significativas entre profesorado y alumnado en la mayoría de las cuestiones, la mayoría de ellas ya señaladas anteriormente. También aparecen algunos acuerdos entre profesorado y alumnado, en este caso

respecto a la importancia de la capacidad cognitiva en el sistema de enseñanza y la frecuencia en la utilización de algunos procedimientos de evaluación. Un estudio similar fue llevado a cabo por Ruiz *et al.* (2013) en las universidades de Castilla la Mancha, de Murcia y la UCAM. Los resultados encontrados coinciden con los de los estudios anteriores, respecto a las diferencias de percepción entre profesores y alumnos, así como en los aspectos en los que se alcanza un cierto acuerdo.

Por tanto, parece que las diferencias de percepción entre profesorado y alumnado universitario ofrecen resultados similares. En cambio, existe poca información sobre en qué medida la existencia de un grupo de innovación docente en un centro de FIP afecta a las percepciones globales que tiene el alumnado sobre los sistemas de evaluación que ha vivido durante su formación.

En ese sentido, este estudio pretende conocer qué tipo de evaluaciones se están llevando a cabo en un centro de FIP (Escuela Universitaria de Magisterio de Segovia, Universidad de Valladolid) que cuenta con un grupo de innovación docente con una considerable trayectoria y tradición en estas temáticas. El estudio también pretende conocer hasta qué punto dichos sistemas de evaluación se aproximan a lo que se entiende por *evaluación formativa y compartida*. Esto se debe a que parte del profesorado del centro sí ha participado en estos procesos de investigación e innovación docente, pero otra parte considerable no lo ha hecho; por ello, tanto el alumnado como el profesorado y los egresados cuentan con diversas experiencias que permiten el análisis comparativo que exponemos en este artículo.

Posteriormente se analiza también si existen diferencias estadísticamente significativas en las percepciones de alumnos, egresados y profesorado sobre los sistemas de evaluación del aprendizaje llevados a cabo en dicho centro.

2. Metodología

2.1 Población y muestra

El estudio se llevó a cabo en un centro de FIP (E.U. de Magisterio de Segovia, Universidad de Valladolid) que cuenta con un grupo de innovación docente con una considerable trayectoria y tradición en estas temáticas. Dicho grupo emergió en el curso 2001/02 mediante la obtención y desarrollo de diferentes proyectos de investigación e innovación docente en convocatorias competitivas, así como la organización de seminarios permanentes, en torno a las siguientes temáticas: utilización de metodologías activas en docencia universitaria, con especial énfasis en los proyectos de aprendizaje tutorados, programas de tutorización del alumnado en FIP, desarrollo de sistemas de evaluación formativa y compartida, etc. Hasta la actualidad se han desarrollado dentro del grupo diferentes proyectos de I+D+I (tanto autonómicos como nacionales) centrados en la evaluación formativa y compartida (Hamodi, 2014).

La muestra estuvo conformada por:

(a) 46 egresados de los últimos 5 años: 70% de mujeres y 30% de hombres; edad media de 26 años; especialidades cursadas de Educación Física (32,6%), Educación Infantil (28,3%), Educación Primaria (19,5%) y Educación Musical (19,5%);

(b) 205 alumnos: 58,5% mujeres y 41,5% de hombres; edad media de 22 años; especialidades cursadas de Educación Física (53,7%), Educación Infantil (22%), Educación Primaria (21%) y Educación Musical (3,4%);

(c) 14 profesores: 35,7% mujeres y 64,3% hombres; edad media de 45 años; experiencia media como docentes en la universidad de 14 años; categoría profesional como profesorado titular de Universidad o catedrático/a de Escuela Universitaria (25%), profesorado titular de Escuela Universitaria (17%), profesorado contratado doctor (17%) y profesorado contratado a tiempo parcial (41%). En torno a la mitad de esta muestra de profesorado pertenece al citado grupo de innovación.

2.2 Instrumentos de recogida de datos

Para la recogida de datos de los tres grupos (egresados, alumnos y profesores) se utilizó un cuestionario con una escala *Likert* de 5 puntos, con valores de 0 a 4. Las personas encuestadas respondían en función del grado de acuerdo con el enunciado (0=nada de acuerdo; 1= poco de acuerdo; 2=algo de acuerdo; 3=bastante de acuerdo; 4=totalmente de acuerdo). Esto permite que todas las preguntas puedan ser consideradas como variables numéricas, pues una variable ordinal puede tratarse como métrica cuando tenga cinco o más categorías (Díaz, 2002).

Para la construcción de dicha escala y otorgarle validez se tuvieron en cuenta instrumentos ya validados utilizados para investigar sobre temáticas similares a la nuestra (Gutiérrez *et al.*, 2011, p. 503). Primeramente, se recogió una amplia lista de posibles preguntas que fueron evaluadas por distintos expertos, con larga experiencia en sistemas de evaluación formativa y compartida en docencia universitaria y con experiencia en el desarrollo de proyectos I+D+I sobre estas temáticas. También acumulan una larga experiencia como profesionales de la FIP. Posteriormente se seleccionaron las más pertinentes por su relevancia (los ítems deberían estar claramente relacionados con el objeto de estudio) y claridad (fácilmente comprensibles, con afirmaciones simples). Una vez seleccionadas las preguntas, dicha escala fue sometida a una comprobación previa o prueba piloto. Dicha prueba fue aplicada a un número reducido de personas seleccionadas (profesores/as y alumnos/as) para comprobar que su administración no provocaba dificultades imprevistas. La escala también alcanzó niveles elevados de fiabilidad, obteniendo en la prueba de *Alfa de Cronbach* los siguientes resultados: alumnado=0,836; egresados=0,827; profesorado=0,681. El cuestionario definitivo tenía 25 preguntas, agrupadas en 4 bloques: (1) sobre el sistema de evaluación en general; (2) sobre el tipo de examen; (3) sobre la calificación; (4) sobre la participación de los alumnos en el proceso de evaluación y calificación.

Además, se calcularon las siguientes variables: MB1 y MB4. La variable MB1 fue hallada realizando la media aritmética de todas las respuestas a las preguntas incluidas en el *bloque 1*. Cuanta más puntuación obtiene un sujeto en MB1, mayor percepción tiene éste de haber vivenciado sistemas de evaluación formativa. De igual manera se calculó la variable MB4, donde cuanto mayor es su puntuación, mayor es la percepción de que los sistemas de evaluación fueron compartidos, implicando al alumnado en los mismos.

2.3. Análisis de datos

Para el análisis de datos se utilizó el programa estadístico *SPSS.V20.0* Con el objetivo de conocer qué discrepancias entre unos grupos y otros eran estadísticamente significativas, se realizó un ANOVA. Posteriormente, se realizó un análisis *post hoc* con los ítems que habían obtenido niveles de significación iguales o inferiores a 0,1 (Sig. \leq 0,01), aplicándoles la pruebas de *Tukey* y *Scheffé*.

Ítem	Egresados/as		Alumnado		Profesorado		Total	
	\bar{x}	σ	\bar{x}	σ	\bar{x}	σ	\bar{x}	σ
P.1.1. Se realizaban procesos de evaluación continua (realización y entrega de documentos durante el desarrollo de la asignatura)	2,00	,816	2,05	,938	3,00	1,000	2,09	,941
P.1.2. Se utilizaban procesos de evaluación formativa (el profesor devolvía los documentos corregidos y daba información sobre cómo hacerlos mejor y corregir errores)	1,87	,909	1,38	,887	3,00	,707	1,55	,958
P.1.3. Las observaciones realizadas sobre los trabajos demandados podían derivar en la repetición para mejorarlos	1,50	,960	1,38	,809	2,50	1,314	1,45	,893
P.1.4. Se explicaban previamente los criterios de calidad de los diferentes trabajos o exámenes	1,70	,916	1,75	,921	3,31	,630	1,82	,967
P.1.5. Existían similitudes entre los sistemas de evaluación de las diferentes asignaturas	1,76	,981	1,36	,922	2,85	,899	1,50	,988
MB1	1,77	,660	1,58	,634	3,00	,512	1,68	,699
P.2.1. Se utilizaban portafolios o carpeta individual para entregar todos los materiales	1,53	1,100	1,19	,901	1,77	1,691	1,28	,998
P.2.2. Se utilizaban portafolios grupales o carpetas colaborativas para	1,13	,919	1,10	,866	1,08	1,441	1,11	,906
P.2.3. Se utilizaban exámenes orales	,67	,920	,14	,385	,54	1,330	,25	,621
P.2.4. Se usaban exámenes escritos de desarrollo (con preguntas largas, desarrollo de un tema, etc.)	2,96	,893	2,77	1,041	1,57	1,399	2,74	1,073
P.2.5. Se realizaban exámenes con preguntas cortas	2,43	,925	2,36	,883	2,31	1,251	2,37	,907
P.2.6. Se utilizaban exámenes tipo test	1,33	,953	1,35	,668	1,00	1,354	1,33	,769
P.2.7. Se utilizaban exámenes prácticos (resolución de problemas, planteamiento de actividades, etc.)	1,16	,903	,94	,790	2,08	1,656	1,03	,901
P.3.1. La calificación provenía únicamente del examen final	2,35	1,079	2,33	,979	,08	,277	2,22	1,089
P.3.2. La nota final provenía del examen final y algún trabajo	2,48	,836	2,71	,836	1,36	1,216	2,60	,908
P.3.3. La calificación provenía del examen final y otros instrumentos (cuadernos, recensiones, trabajos en grupo, etc.)	2,00	,977	2,19	,939	3,31	,751	2,21	,970
P.3.4. La calificación final era la suma de los exámenes parciales realizados	1,63	1,103	1,22	,857	,38	1,121	1,25	,947
P.3.5. La calificación se obtenía sin exámenes	,76	,848	1,01	,798	,77	1,301	,95	,840
P.3.6. El sistema y los criterios de calificación se explicaban al principio de la asignatura	2,50	,960	2,25	1,050	3,77	,439	2,37	1,064
P.3.7. La nota que ponía el profesor es justificada	2,40	1,009	1,83	,934	3,31	,855	2,00	1,010
P.3.8. A pesar de realizar diferentes trabajos teórico-prácticos, el examen tenía un peso definitivo en la nota final	2,98	,783	2,94	,968	1,85	1,214	2,89	,979
P.4.1. Se discutió el sistema de evaluación al comienzo del curso	1,63	1,123	1,71	1,034	2,71	1,267	1,75	1,083
P.4.3. Los alumnos realizaban una valoración del trabajo realizado por ellos mismos (autoevaluación)	1,11	,875	1,27	,859	2,23	1,166	1,29	,903
P.4.4. Los alumnos evaluaban los trabajos de sus compañeros (evaluación entre iguales)	,80	,778	,86	,742	1,38	1,193	,88	,780
P.4.5. La calificación final era consensuada entre el profesor y el alumno (calificación dialogada)	,78	,867	,80	,684	1,08	1,256	,81	,753
P.4.6. Eran los propios alumnos los que emitían una calificación de manera justificada (autocalificación)	,54	,887	,60	,717	1,43	1,284	,64	,805
MB4	1,00	,624	1,05	,583	1,74	,921	1,08	,626

Tabla 1. Media y desviación típica de todos los ítems del cuestionario y las variables MB1 y MB4 en función del colectivo (egresados/as, alumnado o profesorado).

3. Resultados y discusión

En la Tabla 1 presentamos las medias y las desviaciones típicas obtenidas en los tres colectivos, en cada uno de los ítems del cuestionario (incluyendo las variables MB1 y MB4).

En el bloque 1 (B1) se trataba de analizar el carácter formativo de los sistemas de evaluación aplicados. Como bien sabemos, para que un sistema de evaluación sea formativo, también ha de ser continuo en el tiempo. Pero la continuidad no implica que lleve consigo el que sea formativo. En este sentido, los resultados indican que la percepción media de los egresados, alumnado y profesorado, es que los sistemas fueron continuos en algunas ocasiones (\bar{x} P.1.1= 2,09) pero formativos lo fueron en pocas (\bar{x} P.1.2= 1,55), puesto que en escasas ocasiones el profesorado hacía observaciones que posteriormente pudiesen derivar en que el alumnado mejorase sus trabajos teniendo; éstas en cuenta, aprendiendo y autorregulando sus propios errores (\bar{x} P.1.3= 1,45). Recordemos que, “la nota ofrecida a un alumno o alumna sin más no tiene ningún valor didáctico” (Fernández, 2006, p. 93). Asimismo, la media de las percepciones de los tres colectivos indica que algunas veces se explicaban los criterios de calidad y de evaluación de los trabajos y exámenes (\bar{x} P.1.4= 1,82), pero que pocas veces se comentaban los resultados y las dudas surgidas en los trabajos y exámenes (\bar{x} P.1.5= 1,50). La media de los tres colectivos en este bloque, indica que los egresados, el alumnado y el profesorado perciben que los sistemas de evaluación formativos se llevan a cabo en algunas ocasiones (MB1= 1,45). Esto es una percepción media, pues como podemos comprobar en la tabla, la percepción del profesorado es más favorable que la del alumnado (que obtiene las puntuaciones más bajas en el bloque), como sucede en trabajos similares (Arribas et al., 2012; López y Toja, 2013; Taberero y Daniel, 2012; Ruiz et al., 2013). Esta idea también se pone de manifiesto en la investigación de Gargallo (2009) donde concluye que aunque existe un escaso uso, se observa un cierto avance hacia la evaluación formativa que permite que la evaluación se lleve a cabo a lo largo del proceso de enseñanza-aprendizaje, permitiendo al alumno conocer sus avances sin esperar al final del mismo. Álvarez (2009: 371) es más crítico en este sentido, pues los datos de su investigación le llevan a concluir que “el poder formativo de la evaluación, [está] tan permanentemente presente en la narrativa didáctica como ausente en las prácticas de aula”.

Sobre los medios de evaluación se preguntó en el segundo bloque por los portafolios y los exámenes. A pesar de que la literatura especializada considera que los portafolios son un recurso básico en la evaluación orientada al aprendizaje (Castejón et al., 2009: 76; Klenowski, 2005: 39; Mokhtari y Yellin, 1996: 246; Palomares, 2011: 602), los datos muestran consenso entre los tres colectivos sobre que su utilización es escasa, aunque con un mayor uso de los portafolios individuales (\bar{x} P.2.1= 1,28) que de los grupales (\bar{x} P.2.2= 1,11). Respecto a los exámenes, también hay consenso entre los tres grupos respecto al escaso uso de los exámenes orales (\bar{x} P.2.3= 0,25), lo cual coincide con los estudios de Arribas et al. (2010: 33), Gutiérrez et al. (2013: 141) y Taberero y Daniel (2012: 142). En este sentido Díaz (1993: 113) explicaba que la cultura en la universidad fomentaba los medios de evaluación escritos frente a los “hablados”. Y a día de hoy su argumento sigue siendo válido: “es debido al número de alumnos que tenemos en las aulas. Y en otras [ocasiones], porque es un método que nos da mayor

seguridad, dado que por tradición, cumple los requisitos tradicionales”. El disenso aparece al analizar qué tipo de examen se percibe como el más utilizado: mientras el profesorado considera que son los exámenes de preguntas cortas (\bar{x} P.2.5= 2,31), los egresados y los alumnos consideran que son los de preguntas de desarrollo (\bar{x} P.2.4 egresados= 2,96 y \bar{x} P.2.4 alumnos= 2,77; frente a una profesorado= 1,57). Quizá las diferencias en estos ítems (\bar{x} P.2.4- utilización de exámenes escritos de desarrollo y P.2.5- utilización de exámenes con preguntas cortas) podrían deberse, como apuntan Arribas et al. (2010, p. 33), a “problemas de carácter nominalista”, al asignar diferentes significados a un mismo concepto, como podrían serlo los términos “preguntas cortas” y “preguntas de desarrollo”.

Sobre la obtención de la calificación final también aparece disenso: el profesorado considera que la calificación proviene bastantes veces del examen final y otros medios (\bar{x} P.3.3= 3,31, frente a un 2,12 de alumnos y 2,19 de egresados). Los egresados y el alumnado consideran que aunque es así, el examen tiene un peso definitivo en la nota (\bar{x} P.3.8 egresados= 2,98 y \bar{x} P.3.8 alumnos= 2,94, frente a 1,85 de los profesores). En este sentido, no coincidimos con Arribas et al. (2010: 33) cuando afirman que en los tres colectivos estudiados existe consenso al señalar el “peso definitivo que tiene el examen final al margen de cuantas otras actividades se hayan realizado”, pues en el caso del profesorado, esta premisa no se corrobora (pues no consideran que tenga el peso definitivo). Pero respecto a la forma de alcanzar la calificación final, sí estamos de acuerdo con Arribas et al. (2010, p. 33) cuando dicen que “casi nunca [existe] la posibilidad de calificar una asignatura sin un examen que lo acredite”, pues en nuestro caso hay bastante consenso en considerar que casi nunca se obtiene sin exámenes (\bar{x} P.3.5 egresados= 0,76, \bar{x} P.3.5 alumnos= 1,01 y P.3.5 profesores= 0,77). Estos datos difieren de los de Marrahí (2012, p. 59) donde se afirma que los trabajos prácticos son el medio de evaluación más utilizado.

Finalmente, a pesar de los estudios que muestran evidencias sobre el efecto positivo de la participación del alumnado en el proceso de evaluación (Gessa, 2011: 760; Ibarra et al., 2012: 213-215), los tres colectivos consideran que las técnicas de autoevaluación, evaluación entre iguales, calificación compartida y autocalificación son poco utilizadas.

ÍTEM	Sig.
P.2.1. Se utilizaban portafolios o carpeta individual para entregar todos los materiales	,021
P.2.2. Se utilizaban portafolios grupales o carpetas colaborativas	,972
P.2.5. Se realizaban exámenes con preguntas cortas	,858
P.2.6. Se utilizaban exámenes tipo test	,290
P.3.5. La calificación se obtenía sin exámenes	,138
P.4.4. Los alumnos evaluaban los trabajos de sus compañeros (evaluación entre iguales).	,050
P.4.5. La calificación final era consensuada entre el profesor y el alumno (calificación dialogada)	,414

Tabla 2. ANOVA con variable independiente Colectivo y todos los ítems del cuestionario como variable dependiente, con valores Sig. > 0,01.

Variable dependiente	Colectivo	(J) Colectivo	Diferencia de medias (I-J) HSD de Tukey y Scheffé
	Egresados	Alumnos	,18680
		Profesores	-1,23111 (*)
	Alumnos	Egresados	-,18680
		Profesores	-1,41791 (*)
	Egresados	Alumnos	-,04570
		Profesores	-,73105 (*)
	Alumnos	Egresados	,04570
		Profesores	-,68536 (*)
	Egresados	Alumnos	-,049
		Profesores	-1,000 (*)
	Alumnos	Egresados	,049
		Profesores	-,951 (*)
	Egresados	Alumnos	,124
		Profesores	-1,000 (*)
	Alumnos	Egresados	-,124
		Profesores	-1,124 (*)
	Egresados	Alumnos	-,054
		Profesores	-1,612 (*)
	Alumnos	Egresados	,054
		Profesores	-1,558 (*)
	Egresados	Alumnos	,187
		Profesores	1,385 (*)
	Alumnos	Egresados	-,187
		Profesores	1,198 (*)
	Egresados	Alumnos	,214
		Profesores	-,921 (*)
	Alumnos	Egresados	-,214
		Profesores	-1,135 (*)
	Egresados	Alumnos	,016
		Profesores	2,271 (*)
	Alumnos	Egresados	-,016
		Profesores	2,255 (*)
	Egresados	Alumnos	-,229
		Profesores	1,121 (*)
	Alumnos	Egresados	,229
		Profesores	1,350 (*)
	Egresados	Alumnos	-,186
		Profesores	-1,308 (*)
	Alumnos	Egresados	,186
		Profesores	-1,121 (*)
	Egresados	Alumnos	,246
		Profesores	-1,269 (*)
	Alumnos	Egresados	-,246
		Profesores	-1,516 (*)
	Egresados	Alumnos	,036
		Profesores	1,132 (*)
	Alumnos	Egresados	-,036
		Profesores	1,095 (*)
	Egresados	Alumnos	-,082
		Profesores	-1,084 (*)
	Alumnos	Egresados	,082
		Profesores	-1,002 (*)
	Egresados	Alumnos	-,164
		Profesores	-1,122 (*)
	Alumnos	Egresados	,164
		Profesores	-,958 (*)
	Egresados	Alumnos	-,061
		Profesores	-,885 (*)
	Alumnos	Egresados	,061
		Profesores	-,824 (*)

Tabla 3. Resultados del análisis post hoc donde se encuentran diferencias significativas entre el colectivo del profesorado y los colectivos del alumnado y los egresados [P≠(A=E)]. * La diferencia es significativa al nivel .05.

Como puede comprobarse, es en el bloque que mayor consenso hay, pues los tres colectivos perciben como muy baja la participación del alumnado en el proceso de evaluación y calificación (\bar{X} B4= 1,08). En 1993, Díaz también llegaba a esta conclusión, encontrando como explicación que la utilización de técnicas que implican al alumnado conlleva romper con el sistema tradicional existente (y mayoritario) en la universidad (Díaz, 1993: 111). Podemos afirmar que dicho argumento sigue vigente en nuestros días y coincide con otros estudios de características similares (Arribas *et al.*, 2010: 33; Gutiérrez *et al.*, 2011: 510; Gutiérrez *et al.*, 2013: 141; López y Palacios, 2012: 334; Marrahí, 2012: 60; Palacios y López, 2013: 191; Ruiz *et al.*, 2013: 27; Taberner y Daniel, 2012: 138).

Respecto a las desviaciones típicas, cabe resaltar que son más elevadas en el cuestionario del profesorado (σ media del profesorado= 1,091). Estas cifras no son tan elevadas en el del alumnado y los egresados (σ media de los egresados= 0,911 y σ media del alumnado= 0,838). Este planteamiento difiere del planteado por Arribas *et al.* (2010, p. 32), donde tras analizar cuestionarios de egresados, alumnos y profesores resaltan “la marcada divergencia en cuanto a las respuestas emitidas por los grupos de alumnos y egresados en numerosos ítems, lo que podríamos denominar como incongruencia intra-grupal”. En nuestro caso, como acabamos de indicar, encontramos más divergencia en las respuestas del profesorado. Esto se debe fundamentalmente a dos cuestiones: la primera de ellas, es que la muestra de este colectivo es la menor de las tres, lo que provoca que cuando un sujeto se desvía de la respuesta media, provoca una variación mayor que si la muestra fuese más grande. La segunda cuestión es que en el mismo centro co-existen prácticas docentes muy diversas entre sí, correspondiéndose con las tres tipologías de profesorado (“innovador”, “eclectico” y “tradicional”) establecidas por los estudios de López y Palacios (2012) y Palacios y López (2013). Estas realidades tan diferentes generan una considerable disparidad en las respuestas. Por el contrario, alumnado y egresado muestran una desviación típica menor, debido a que la realidad que perciben es más parecida entre unos y otros.

Como acabamos de describir, existen discrepancias en las percepciones en función de los colectivos. Con el objetivo de conocer cuáles eran estadísticamente significativas, se realizó un ANOVA. Los resultados arrojan ideas muy claras: de los 27 ítems del cuestionario, 20 obtienen niveles de significación iguales o inferiores a 0,1 (Sig. \leq 0,01), lo que querría decir que en esos 20 ítems (incluidos en las tablas 1, 2 y 3) existen diferencias significativas en las respuestas en función del colectivo de pertenencia. En la Tabla 2 presentamos los ítems cuyos niveles son superiores y en los que por lo tanto no existen diferencias significativas entre los colectivos.

Posteriormente, se realizó un análisis *post hoc* con los 20 ítems que habían obtenido niveles de significación iguales o inferiores a 0,1 (Sig. \leq 0,01), aplicándoles la pruebas de Tukey y Scheffé con el objetivo de comprobar entre que colectivos existía un disenso estadísticamente significativo. Los valores de las dos pruebas (Tukey y Scheffé) dan valores idénticos y los resultados muestran que de los 20 ítems a los que se les aplicó las pruebas: (a) en 15 de ellos la discrepancia se encuentra entre el colectivo de profesores por un lado, y el de egresados y alumnos por otro (tabla 3); (b) en un ítem la diferencia se encuentra entre el colectivo del alumnado por un lado, y el profesorado y los egresados por otro (tabla 4); (c) en 4 de ellos existe una diferencia significativa entre todos los colectivos (Tabla 5).

Gutiérrez *et al.*, (2011), utilizando un cuestionario muy similar al nuestro y con una población más amplia (egresados, alumnado y profesorado pertenecientes a las Universidades de Castilla y León) no obtienen este mismo resultado en dos ítems referentes a la participación del alumnado (P.4.3 y P.4.6). En el citado estudio se considera que “las diferencias o no existen (no son significativas) o lo son pero de escasa magnitud” (Gutiérrez *et al.*, 2011: 507), pero en nuestro caso son significativamente diferentes entre el profesorado y los egresados y alumnos [$P \neq (A=E)$].

No obstante, en muchas investigaciones centradas en analizar las diferencias de las percepciones entre los egresados, los estudiantes y el profesorado sobre la evaluación mediante cuestionarios se concluye que la diferencia entre las percepciones de un grupo y otro son elevadas (Marrahí, 2012: 60; Martínez *et al.*, 2012: 69; Taberner y Daniel, 2012: 138-142) e incluso estadísticamente significativas (Gutiérrez *et al.*, 2013: 142; López y Toja, 2013; Ruiz *et al.*, 2013: 22). Es lo que Arribas *et al.* (2010: 32) denominan “incongruencia inter-grupal” para referirse a las “discrepancias muy significativas en cuanto a las respuestas emitidas por los distintos grupos respecto de un mismo ítem, especialmente cuando este no supone una valoración subjetiva sino la constatación de un hecho”, y la encuentran sobre todo “entre el grupo de profesores y los grupos de alumnos y egresados: $P \neq (A=E)$. La incoherencia inter-grupal entre alumnos y egresados es mucho menor, coincidiendo entre sí en la mayoría de las respuestas”. En el estudio citado (en el que se utilizó un cuestionario de la misma naturaleza que el nuestro) los resultados fueron obtenidos únicamente mediante porcentajes, sin un análisis donde se indicase si la discrepancia era estadísticamente significativa o no.

Analizando los 15 ítems incluidos en la tabla 3 se observa que una elevada puntuación media en los ítems MB1, MB4, P.1.1, P.1.3, P.1.4, P.2.7, P.3.3, P.3.6, P.4.1, P.4.3 y P.4.6 indica que el sistema de evaluación tiende a ser formativo y compartido. Por el contrario, una elevada puntuación en los ítems P.2.4, P.3.1, P.3.2 y P.3.8 indica todo lo contrario (utilización de exámenes escritos de desarrollo, obtención de la nota final mediante examen final y algún trabajo, concesión de un peso definitivo al examen en la nota final a pesar de realizar diferentes trabajos teórico-prácticos). Teniendo esta premisa en consideración, y deteniéndonos en el análisis concreto de los ítems de la tabla anterior y de los valores medios indicados en la Tabla 4, se observa que la puntuación media del profesorado es significativamente más elevada que la del alumnado y los egresados en todos los ítems que hacen alusión a prácticas favorables a la evaluación formativa y compartida.

Es decir, la puntuación media del profesorado es significativamente mayor que la del alumnado y los egresados en los siguientes ítems: “Media del Bloque 1 sobre sistemas formativos” (MB1); “Media del Bloque 4 sobre sistemas compartidos” (MB4); “Realización de procesos de evaluación continua” (P.1.1); “Repetición de los trabajos para mejorarlos tras las observaciones del profesor” (P.1.3); “Explicación previa de los criterios de calidad de los trabajos” (P.1.4); “Utilización de exámenes prácticos” (P.2.7); “Obtención de la calificación mediante examen final y otros instrumentos” (P.3.3); “Explicación inicial del sistema y los criterios de calificación” (P.3.6); “Discusión del sistema de evaluación al comienzo del curso” (P.4.1); “Autoevaluación” (P.4.3); “Autocalificación” (P.4.6).

Variable dependiente	Colectivo	(J) Colectivo	Diferencia de medias (I-J) HSD de Tukey y Scheffé
	Egresados	Alumnos	,537 (*)
		Profesores	,135
	Alumnos	Egresados	-,537 (*)
		Profesores	-,402 (*)

Tabla 4. Resultados del análisis post hoc donde se encuentran diferencias significativas entre el colectivo del alumnado y los colectivos del profesorado y los egresados [A≠(P=E)]

Por el contrario, en los ítems “Uso de exámenes escritos de desarrollo” (P.2.4); “Calificación proveniente únicamente del examen final” (P.3.1); “Calificación proveniente del examen final y otros instrumentos” (P.3.2); “Peso definitivo del examen en la nota final a pesar de realizar otros trabajos” (P.3.8), que hacen referencia a prácticas más habituales en sistemas de evaluación tradicionales y sumativos, el profesorado es el colectivo que ha obtenido unas medias significativamente menores que los egresados y el alumnado.

Analizando las Tablas 1 y 3, se puede afirmar que el alumnado y los egresados consideran que los sistemas de evaluación formativa y compartida todavía son escasos, aunque en algunas ocasiones sí que han experimentado sistemas de este tipo. Conclusión a la que también se llega en otros estudios de la misma naturaleza, afirmando que “continúan predominando los modelos tradicionales de docencia universitaria [...] frente a modelos innovadores de evaluación formativa” (López y Palacios, 2012: 334-336).

En Gutiérrez et al. (2011: 506) también se concluye que las valoraciones de los profesores en el uso de sistemas de evaluación formativa y compartida son más positivas que las de los alumnos y egresados. En nuestro caso, podemos corroborar esta idea gracias al ANOVA y al análisis *post hoc*, donde se observa esta relación en 15 de los 20 ítems con diferencias significativas. En los otros 5 se observan otro tipo de diferencias: en un ítem la diferencia se encuentra entre el colectivo del alumnado por un lado, y el profesorado y los egresados por otro (Tabla 4) y en 4 de ellos existe una diferencia significativa entre todos los colectivos (Tabla 5).

La percepción sobre la utilización de los exámenes orales divide significativamente la opinión de los colectivos en dos grupos: mientras el alumnado considera que no se utilizan (\bar{X} P.2.3= 0,14), los egresados y el profesorado consideran que se utilizan un poco (\bar{X} P.2.3= 0,67 y \bar{X} P.2.3= 0,54 respectivamente). Es posible que la percepción del alumnado quede en esta ocasión alejada de la del profesorado y los egresados debido a una mínima transformación: que en promociones anteriores (egresados) se hiciesen exámenes orales en alguna ocasión más, probablemente debido a que el número de alumnos lo permitiese. Otra posible explicación puede ser que este tipo de exámenes otorgan al profesorado menos seguridad que los exámenes escritos, que cumplen con los requisitos tradicionales (Díaz, 1993, p. 113). De cualquier manera, el dato sobre la frecuencia en los tres colectivos es claro: indica un escaso uso.

Los 4 ítems que se presentan en la tabla 5 son aquellos en los que existe una diferencia significativa entre todos los colectivos.

Estos resultados muestran que en los ítems P.1.2 (“Se utilizaban procesos de evaluación formativa”), P.1.5 (“Se comentaban los resultados y las dudas de los exámenes y actividades de evaluación”) y P.3.7 (“La nota que ponía el profesor es justificada”), se acentúan significativamente las diferencias entre los colectivos, de manera que, en los tres casos, la percepción del profesorado es la más favorable, la

Variable dependiente	Colectivo	(J) Colectivo	Diferencia de medias (I-J) HSD de Tukey y Scheffé
	Egresados	Alumnos	,489 (*)
		Profesores	-1,130 (*)
	Alumnos	Egresados	-,489 (*)
		Profesores	-1,620 (*)
	Egresados	Alumnos	,395 (*)
		Profesores	-1,091 (*)
	Alumnos	Egresados	-,395 (*)
		Profesores	-1,485 (*)
	Egresados	Alumnos	,406 (*)
		Profesores	1,246 (*)
	Alumnos	Egresados	-,406 (*)
		Profesores	,840 (*)
	Egresados	Alumnos	,572 (*)
		Profesores	-,908 (*)
	Alumnos	Egresados	-,572 (*)
		Profesores	-1,479 (*)

Tabla 5. Resultados del análisis post hoc donde se encuentran diferencias significativas entre todos los colectivos (P≠A≠E), y los egresados [A≠(P=E)]. * La diferencia es significativa al nivel .05.

de los egresados la intermedia y la del alumnado la menos favorable. A diferencia de los ítems anteriores, el P.3.4 (“La calificación final era la suma de las calificaciones de los exámenes parciales realizados”) arroja el siguiente resultado: el profesorado obtiene la menor puntuación, considerando que nunca se obtenía la calificación final mediante la suma de exámenes parciales (\bar{x} P.3.4= 0,38). En medio se encuentra el alumnado, que considera que este sistema de obtener la calificación final se daba pocas veces (\bar{x} P.3.4= 1,22) y los egresados son los que mayor puntuación obtienen, considerando que se daba algunas veces (\bar{x} P.3.4= 1,63).

De todo este análisis se deduce que los alumnos y los egresados perciben la realidad sobre los sistemas de evaluación vividos de manera parecida, mientras que el profesorado percibe la realidad de una manera significativamente diferente. Aunque utilizando ítems diferentes a los nuestros, también en otros trabajos se han resaltado las diferencias entre las percepciones del alumnado y del profesorado. Por ejemplo, en Barnett, Matthews y Jackson (2003: 5) o en Maclellan (2001). Además, en Levy, Wubbels, den Brok y Brekelmans (2003: 5), se utilizaron observadores externos para analizar quienes estaban más cerca de la realidad, llegando a la conclusión de que las percepciones de los alumnos parecían estar más cerca de la realidad que las de los profesores. Por el contrario, en la investigación de López y Toja (2013: 9), “no se han encontrado demasiadas diferencias entre la opinión del alumnado y el profesorado”.

Teniendo en cuenta los valores de las medias de cada ítem y el análisis *post hoc*, podemos afirmar que los profesores perciben una mayor utilización de los sistemas de evaluación formativa y compartida que los estudiantes y egresados. En esta ocasión, nuestros resultados si coinciden con los de López y Toja (2013, p. 9), cuando afirman que “los profesores suelen ser más optimistas en las valoraciones que los estudiantes”. Además, por lo general, cuando hay diferencias significativas entre estudiantes y egresados (como en los ítems P.1.2, P.1.5, P.3.4 y P.3.7), son los primeros (los estudiantes) los que perciben que la utilización de estos sistemas se daba en menor medida. Incluso cuando las diferencias no son significativas, la percepción de los egresados suele ser más benévola. Una posible explicación a este hecho podría encontrarse en la idea de que los estudiantes, con el paso del tiempo, tras su salida de la universidad, podrían diluir las experiencias concebidas como más negativas, reafirmando las más positivas. Se trata pues de un proceso psicológico que apunta a que las huellas mnémicas que hacen posible el recuerdo se desvanecen con el paso del tiempo y, por lo tanto, el sujeto no puede recuperar la información pertinente (“teoría del decaimiento o desvanecimiento automático de la huella”, según Peterson y Peterson, 1959). Aunque esto no quiere decir que los egresados hayan olvidado experiencias negativas, sí podría decirse que con el paso del tiempo las han difuminado, otorgando una valoración más intermedia que las que hubiesen manifestado durante su etapa de estudiantes. Por todo esto no podemos verificar la idea presentada por Gutiérrez et al. (2013, p. 139) que indican que “las percepciones específicas sobre la evaluación recibida [...] tienden a mantenerse al menos durante algunos años con bastante solidez”.

A pesar de esta posible idea, no podemos pasar por alto que, valoraciones diferentes ante el uso de determinados sistemas de evaluación (medios, técnicas, instrumentos, participación del alumnado, etc.) entre el alumnado y los egresados pueden deberse a cambios reales en el uso de los mismos.

4. Conclusiones

Los resultados de la investigación muestran varios consensos entre los tres colectivos. La percepción media de los egresados, alumnado y profesorado respecto a los sistemas de evaluación es que fueron continuos en algunas ocasiones, pero formativos en pocas. Respecto a los medios de evaluación utilizados, claramente aparecen el examen y en ocasiones el portafolio, con mucha más frecuencia el individual que el grupal. Existe también consenso al considerar los tres colectivos que la calificación casi nunca se obtiene sin exámenes, siendo bastante escaso el uso de exámenes orales. Por último, también hay bastante acuerdo en constatar la baja participación del alumnado en el proceso de evaluación y calificación.

En cambio, respecto a otras prácticas los resultados señalan una clara disparidad inter-grupal, pero con matices particulares. Cuando las respuestas han sido significativamente diferentes entre unos colectivos y otros, se han dado de la siguiente manera: (a) en la mayoría de casos el profesorado por un lado y el alumnado y los egresados por otro [$P \neq (A=E)$]; (b) en algunos casos las percepciones de los grupos han sido diferentes entre los tres grupos ($P \neq A \neq E$); (c) en un solo caso las diferencias han separado al grupo de alumnos del de egresados y profesores [$A \neq (P=E)$]. Estas diferencias podrían explicarse también con el argumento anterior: por un lado los profesores puntúan más alto en los ítems que se relacionan con la evaluación formativa y compartida porque son conscientes de que son sistemas más acordes a las exigencias del EEES y, por lo tanto, hacia los que se deberían estar acercando. Por otro lado, para acercar las percepciones del profesorado a las de los estudiantes y egresados, los primeros deberían asumir una posición crítica ante su propia práctica docente, realizando un mayor esfuerzo de reflexión y autocrítica. Cuando se comparan las similitudes y diferencias de los resultados entre egresados y alumnos, suelen ser, por lo general, más favorable la de los primeros que la de los segundos.

Esperamos que este estudio sea de interés para los investigadores en temáticas de evaluación formativa en educación superior, formación inicial del profesorado, así como para los profesionales implicados en procesos de innovación docente en la universidad interesados en la temática de la evaluación. Como prospectiva de futuro, estamos interesados en profundizar en la temática de las percepciones de alumnado y egresados sobre los sistemas de evaluación vividos durante su formación inicial como profesores, a través de metodologías de investigación más cualitativas.

5. Bibliografía

- Álvarez, Juan M. (2001). *Evaluar para conocer, examinar para excluir*. Madrid: Morata.
- Álvarez, Juan M. (2009). La evaluación en la práctica de aula. Estudio de campo. *Revista de Educación*, 350, 351-374.
http://www.revistaeducacion.mec.es/re350/re350_15.pdf
- Arribas, José M^a., Carabias, David; Monreal, Inés (2010). La docencia universitaria en la Formación Inicial del Profesorado. El caso de la Escuela de Magisterio de Segovia. *REIFOP. Revista Electrónica Interuniversitaria de Formación del Profesorado*, 13(3), 27-35.
http://www.aufop.com/aufop/uploaded_files/articulos/1285861727.pdf
- Barnett, Candace W.; Matthews, Hewitt W.; Jackson, Richard A. (2003). A comparison between students

- ratings and Faculty Self-ratings of Instructional Effectiveness. *American Journal of Pharmaceutical Education*, 67(4), 1-6. Recuperado de <http://archive.ajpe.org/view.asp?art=aj6704117&pdf=yes>
- Biggs, John (2005). *Calidad del aprendizaje universitario*. Madrid: Narcea.
- Black, P, y Wiliam, D. (2006). Assessment for Learning in the Classroom. En Gardner, J. (Ed.), *Assessment and Learning*. London: SAGE, pp.9-25
- Bonsón, M., y Benito, Á. (2005). Evaluación y aprendizaje. En Benito, Á. y Cruz, A. (Coords.), *Nuevas claves para la docencia universitaria en el Espacio Europeo de Educación Superior*. Madrid: Narcea, pp.87-100.
- Boud, David (Ed.) (1988). *Developing Student Autonomy in Learning*. New York: Kogan.
- Boud, David; Falchikov, Nancy (Eds.) (2007). *Rethinking Assessment in Higher Education. Learning for the long term*. Oxon: Routledge.
- Brown, Sally; Glasner, Angela (Eds.) (2003). *Evaluar en la Universidad. Problemas y nuevos enfoques*. Madrid: Narcea.
- Brown, Sally; Pickforf, Ruth (2013). *Evaluación de habilidades y competencias en Educación Superior*. Madrid: Narcea.
- Casanova, María A. (2012). El diseño curricular como factor de calidad educativa. *REICE. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 10(4), 7-20. Recuperado de <http://www.rinace.net/reice/numeros/arts/vol10num4/art1.pdf>
- Castejón, F. J., Capllonch, M., González, N., y López, V. M. (2009). Técnicas e instrumentos de evaluación formativa y compartida para la docencia universitaria. En López, V. M. (coord.), *Evaluación formativa y Compartida en Educación Superior*. Madrid: Narcea, pp.65-91.
- Díaz, M. D. (1993). Evaluación participativa en la Universidad: una aventura formativa. *Revista de Enseñanza Universitaria*, 6, pp.111-121.
- Díaz, Vidal (2002). *Técnicas de Análisis Multivariante para Investigación Social y Comercial*. Madrid: Rama.
- Dochy, Filip; Segers, Mien; Dierick, Sabine (2002). Nuevas vías de aprendizaje y enseñanza y sus consecuencias: Una nueva era de evaluación. *Revista de Docencia Universitaria*, 2(2), 13-31. <http://revistas.um.es/redu/article/view/20051>
- Falchikov, Nancy (2005). *Improving Assessment Through Student Involvement. Practical solutions for aiding learning in higher and further education*. Oxon: Routledge.
- Fernández, J. (2006). ¿Evaluación? No gracias, calificación. *Cuadernos de Pedagogía*, 243, 92-97.
- Gargallo, Ana (2009). Aproximación al estudio de la evaluación en la universidad: un análisis exploratorio con alumnos semipresenciales. *Contextos educativos*, 12, 77-93. <https://publicaciones.unirioja.es/ojs/index.php/contextos/article/view/612/575>
- Gauntlett, N. (2007). *Literature review on formative assessment in higher education*. London: Middlesex University.
- Gessa, Ana (2011). La coevaluación como metodología complementaria de la evaluación del aprendizaje: análisis y reflexión en las aulas universitarias. *Revista de Educación*, 354, 749-764. Recuperado de http://www.mecd.gob.es/revista-de-educacion/numeros-revista-educacion/numeros-antiores/2011/re354/re354_30.html
- DOI: 10.4438/1988-592X-RE-2011-354-019
- Gibbs, G. (2003). Uso estratégico de la evaluación en el aprendizaje. En Brown, S., y Glasner, A. (Eds.), *Evaluar en la Universidad: Problemas y nuevos enfoque*. Madrid: Narcea, pp.61-75.
- Gutiérrez, C., Pérez, Á., y Pérez, M. (2013). Percepciones de profesores, alumnos y egresados sobre los sistemas de evaluación en estudios universitarios de formación del profesorado de Educación Física. *Revista Ágora para la Educación Física y el deporte*, 15(2), 130-151.
- Gutiérrez, C., Pérez, Á., Pérez, M., y Palacios, A. (2011). Percepciones de profesores y alumnos sobre la enseñanza, evaluación y desarrollo de competencias en estudios universitarios de formación de profesorado. *Cultura y Educación*, 23(4), 499-514.
- Hamodi, Carolina (2014). *La evaluación formativa y compartida en Educación Superior: un estudio de caso*. Tesis doctoral, Universidad de Valladolid. <http://uvadoc.uva.es/bitstream/10324/5668/1/TESIS562-140728.pdf>
- Ibarra, María S.; Rodríguez, Gregorio; Gómez, Miguel Á. (2012). La evaluación entre iguales: beneficios y estrategias para su práctica en la universidad. *Revista de Educación*, 359, 206-231. http://www.mecd.gob.es/revista-de-educacion/numeros-revista-educacion/numeros-antiores/2012/re359/re359_11.html DOI: 10.4438/1988-592X-RE-2011-359-092
- Klenowski, Val (2005). *Desarrollo del portafolios para el aprendizaje y la evaluación*. Madrid: Narcea.
- Knight, Peter (Ed.) (2005). *El profesorado de Educación Superior. Formación para la excelencia*. Madrid: Narcea.
- Lapahan, A. y Webster, R. (2003). Evaluación realizada por los compañeros: motivaciones, reflexión y perspectivas de futuro. En Brown, S., y Glasner, A. (Eds.), *Evaluar en la universidad. Problemas y nuevos enfoques*. Madrid: Narcea, pp. 203-210
- Levy, Jack; Wubbels, Theo; den Brok, Perry; Brekelmans, Mieke (2003). Students perceptions of interpersonal aspects of the learning environment. *Learning Environments Research*, 6, 1, pp.5-36. <http://link.springer.com/article/10.1023/A%3A1022967927037#page-1>
- López, C., y Toja, B. (2013). Percepciones y valoraciones del profesorado y del alumnado universitario sobre la evaluación: similitudes y diferencias. *Revista de Educación Física*, 29 (3), pp.1-13.
- López, Víctor M. (Coord.) (2009). *La Evaluación Formativa y Compartida en Educación Superior: propuestas, técnicas, instrumentos y experiencias*. Madrid: Narcea.
- López, Víctor M.; Palacios, Andrés (2012). Percepción de los futuros docentes sobre los sistemas de evaluación de sus aprendizajes. *Revista Teoría de la Educación: Educación y Cultura en la Sociedad de la Información*. 13 (3), pp.317-340. http://campus.usal.es/~revistas_trabajo/index.php/revistatesi/article/view/9143/9376
- Maclellan, Effie (2001). Assessment for learning: the differing perceptions of tutors and students. *Assessment*

- and *Evaluation in Higher Education*, 26 (4), pp.307-318. <http://strathprints.strath.ac.uk/2427/>
- Marrahí, María (2012). *Evaluación Formativa en la Formación Inicial del Profesorado De Educación Física de la Escuela universitaria de Magisterio de Segovia*. Trabajo Fin de Máster, Universidad de Valladolid. <http://uvadoc.uva.es/handle/10324/2467>
- Martínez, Luis F.; Castejón, Francisco J.; Santos, María L. (2012). Diferentes percepciones sobre evaluación formativa entre profesorado y alumnado en formación inicial en Educación Física. *REIFOP. Revista Electrónica Interuniversitaria de Formación del Profesorado*, 15(4), 57-67. Recuperado de <http://www.aufop.com/aufop/revistas/arta/digital/1-71/1747>
- McDowell, L., y Sambell, K. (2003). La experiencia en la evaluación innovadora. En Brown, S., y Glasner, A. (Eds.), *Evaluar en la Universidad. Problemas y nuevos enfoques*. Madrid: Narcea, pp.91-103.
- Mokhtari, D.; Yellin, D. (1996). Portfolio Assessment in Teacher Education: Impact on Preservice Teachers' Knowledge and Attitudes. *Journal of Teacher Education*, 47, 245-248.
- Palacios, Andrés; López, Víctor 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, 279-305. http://www.revistaeducacion.educacion.es/doi/361_143.pdf DOI: 10-4438/1988-592X-RE-2011-361-143
- Palomares, Ascensión (2011). El modelo docente universitario y el uso de nuevas metodologías en la enseñanza, aprendizaje y evaluación. *Revista de Educación*, 355, 591-604. Recuperado de http://www.mecd.gob.es/revista-de-educacion/numeros-revista-educacion/numeros-antiores/2011/re355/re355_25.html DOI: 10.4438/1988-592X-RE-2011-355-038
- Pérez, A., Julián, J. A., y López, V. M. (2009). Evaluación formativa y compartida en el Espacio Europeo de Educación Superior (EEES). En López, V. M. (Coord.). *Evaluación formativa y Compartida en Educación Superior*. Madrid: Narcea, pp.19-43.
- Peterson, L. R. y Peterson, M. J. (1959). Short term retention of individual verbal items. *Journal of Experimental Psychology*, 58, 193-198.
- Race, P. (2003). ¿Por qué evaluar de un modo innovador? En Brown, S., y Glasner, A. (Eds.), *Evaluar en la Universidad. Problemas y nuevos enfoques*. Madrid: Narcea, pp.77-90.
- Rodríguez, G.; Ibarra, M. S. (2011). *e-Evaluación orientada al e-Aprendizaje estratégico en Educación Superior*. Madrid: Narcea.
- Ruiz, J. R.; Ruiz, E.; Ureña, N. (2013). La evaluación en la formación inicial del profesorado: qué creemos hacer y qué perciben los alumnos. *Cultura, Ciencia y Deporte*, 8(22), 17-29. <http://dx.doi.org/10.12800/ccd.v8i22.220>
- Sanmartí, Neus (2007). *10 ideas clave: Evaluar para aprender*. Madrid: Graó.
- Santos, Miguel Á. (2003). *Una flecha en la diana: La evaluación como aprendizaje*. Madrid: Narcea.
- Tabernero, Belén; Daniel, María José (2012). Análisis de los sistemas de evaluación del alumnado en la formación del profesorado: estudio comparativo entre lo que opinan profesores y egresados. *REIFOP. Revista Electrónica Interuniversitaria de Formación del Profesorado*, 15(3), 133-144. <http://www.redalyc.org/pdf/2170/217026228014.pdf>
- Zabalza, Miguel Á. (2007). *Competencias docentes del profesorado universitario: Calidad y desarrollo profesional*. Madrid: Narcea.

| Cita recomendada de este artículo

Hamodi Galán, Carolina; López Pastor, Ana Teresa; López Pastor, Víctor Manuel (2015). Percepciones de alumnos, egresados y profesores sobre los sistemas de evaluación del aprendizaje. *@tic. revista d'innovació educativa*. (nº 14). URL. Fecha de consulta, dd/mm/aaaa.