

Instructions for authors, subscriptions and further details:

<http://remie.hipatiapress.com>

¿Implicar al alumnado en la evaluación en la formación inicial del profesorado? Un estudio de caso de evaluación entre iguales de un examen

Víctor Manuel López-Pastor¹, Javier Castejón Oliva² & Ángel Pérez-Pueyo³

1) Departamento de Didáctica de la Expresión Musical, Plástica y Corporal, Universidad de Valladolid, Spain.

2) Departamento de Educación Física, Deporte y Motricidad Humana, Universidad Autónoma de Madrid, Spain.

3) Departamento de Educación Física y Deportiva, Universidad de León, Spain.

Date of publication: June 15th, 2012

To cite this article: López-Pastor, V.M., Castejón, J. & Pérez-Pueyo, Á. (2012). ¿Implicar al alumnado en la evaluación en la formación inicial del profesorado? Un estudio de caso de evaluación entre iguales de un examen. *Multidisciplinary Journal of Educational Research*, 2(2), 177-201. doi: 10.4471/remie.2012.09

To link this article: <http://dx.doi.org/10.4471/remie.2012.09>

PLEASE SCROLL DOWN FOR ARTICLE

The terms and conditions of use are related to the Open Journal System and to Creative Commons Non-Commercial and Non-Derivative License.

Involving Students in Teacher Education Assessment? A Case Study of a Peer Evaluation Experience

Víctor Manuel López-Pastor
Universidad de Valladolid

Javier Castejón
Universidad Autónoma de Madrid

Ángel Pérez-Pueyo
Universidad de León

Abstract

This study covered the grading of partial exams as a learning strategy closely linked to a range of professional competencies and the development of better learning processes. It took place in three teacher education programs. After students undertook partial exams, these were marked in-situ by peers (using an ad-hoc marking scheme) and, later on, by the tutor. Furthermore, students responded to a questionnaire in which they gave their opinions regarding that peer-grading assessment experience. The purposes of the study were: (a) to compare the grades given by the students with those given by the tutors; and (b) to gather the students' opinions on the experience. The results indicate that (a) the correlations between the grades given by the students and tutors ranged from .66 to .90, although the average grades of the student group were lower than those of the tutors; and (b) the students valued positively the experience of peer-grading as a learning activity.

Keywords: Assessment, Teacher education, Student's involvement.

¿Implicar al Alumnado en la Evaluación en la Formación Inicial del Profesorado? Un estudio de caso

Víctor Manuel López-Pastor
Universidad de Valladolid

Javier Castejón
Universidad Autónoma de Madrid

Ángel Pérez-Pueyo
Universidad de León

Abstract

Este estudio, sobre la calificación de exámenes parciales como estrategia de aprendizaje estrechamente relacionada con algunas competencias profesionales y con el desarrollo de mejores procesos de aprendizaje, se realizó en tres centros de Formación Inicial del Profesorado. Tras la realización de exámenes parciales, éstos fueron corregidos in-situ por iguales (mediante una plantilla diseñada “ad hoc”) y posteriormente por el profesor. Además, el alumnado cumplimentó un cuestionario en el que valoró esa experiencia de co-evaluación. Los objetivos del estudio fueron: (a) comparar la calificación otorgada por los iguales con la otorgada por el profesorado; y (b) obtener las valoraciones del alumnado sobre la experiencia. Los resultados indican que (a) las correlaciones de las notas entre alumnado y profesorado oscilaron entre .66 y .90, si bien la media de notas del alumnado fue más baja que la del profesorado; y (b) el alumnado valoró positivamente la experiencia de co-evaluación como actividad de aprendizaje.

Keywords: Evaluación, Formación inicial del profesorado, Participación alumnado en la evaluación.

En la actualidad, la universidad española está desarrollando el denominado “Espacio Europeo de Educación Superior” (EEES), conocido también como “el proceso Bolonia”, que armoniza las titulaciones en toda Europa, con lo que se facilita el movimiento e intercambio de alumnado y profesorado entre las universidades europeas. El EEES establece tres niveles de titulación (Grado, Master y Doctorado) y utiliza como criterio de medida el “crédito europeo” o “European Credit Transfer System” (ECTS), en el que 1 crédito ECTS suponen 25-30 horas de trabajo del alumnado, en coherencia con un nuevo paradigma educativo que centra el interés en el proceso de aprendizaje del alumnado, en vez de hacerlo en el proceso de enseñanza del profesorado. El EEES implica también otros cambios, como que la principal finalidad es el desarrollo de competencias personales y profesionales en el estudiante, la importancia de utilizar metodologías activas de aprendizaje y el uso de una evaluación más orientada al aprendizaje, lo que justifica que se hagan investigaciones sobre las estrategias de evaluación más cercanas al proceso de aprendizaje de los y las estudiantes.

En este sentido, la bibliografía especializada permite encontrar trabajos que muestran la idoneidad de las metodologías activas y los sistemas de evaluación formativa para mejorar el nivel de aprendizaje del alumnado, en la dirección establecida por el proceso de convergencia hacia el EEES (Biggs, 1999; Boud, 1995; Fraile Aranda, 2004, 2006; López-Pastor, 2008, 2009; Martínez, 2006; Pérez-Pueyo, 2008).

El nivel de autogobierno del que goza la universidad en la actualidad y que incide en la mejora de la calidad docente (LOU, 2001) permite y requiere, entre otras cosas, modificar los procesos de evaluación, los cuales demandan un complejo proceso de transformación (Algozzine et al., 2004; Emery, Kramer, & Tian, 2003). Generar cambios en la universidad es una tarea lenta (Biggs, 1999; Brown & Glasner, 1999; Knigh, 2005; Navarro, Santos, Buscá-Donet, Martínez-Mínguez, & Martínez-Muñoz, 2010), pero existe un creciente interés en la docencia

universitaria para el desarrollo de las competencias del alumnado, los procesos de evaluación y la participación de los estudiantes en dicho proceso (Arribas, Carabias, & Monreal, 2010).

A raíz del proceso de convergencia hacia el EEES, han surgido numerosas experiencias de innovación en Formación Inicial del Profesorado (FIP) basadas en el desarrollo de sistemas de evaluación formativa, claramente orientados al aprendizaje y que promueven la implicación del alumnado en los procesos de evaluación (Carless, 2007, 2010; López-Pastor, 2008, 2009; Santos, Martínez, & López, 2009). Una implicación que supone tratar y entender al estudiante como protagonista activo de su aprendizaje (Biggs, 1999; Brown & Glasner, 1999; Falchikov, 2005) y unos sistemas de evaluación claramente orientados al aprendizaje (Learning-oriented Assessment - LOA -) (Carless, 2007; Taras, 2002).

La importancia de la implicación del alumnado en la evaluación en formación inicial del profesorado

La implicación del alumnado en los procesos de evaluación que se llevan a cabo en la FIP es urgente y necesaria, por su estrecha relación con el desarrollo de competencias profesionales (Perrenoud, 2004). La evaluación en la FIP no es sólo un aspecto metodológico, sino también un contenido específico de aprendizaje y una de las principales competencias profesionales a desarrollar. En este sentido, la implicación del alumnado en la evaluación y el uso de sistemas de evaluación orientada al aprendizaje deberían ser algo habitual y extendido en las diferentes asignaturas que componen la FIP. Hay tres razones principales para ello: (a) la profesión docente tiene un fuerte componente práctico, por lo que es importante desarrollar los planteamientos de Dewey de “aprender haciendo” (“learning by doing”) (Dawkins, Ritz, & Loudon, 2009); (b) los maestros tienden a utilizar en su práctica docente la misma metodología que vivieron durante su FIP, en vez de utilizar las teorías estudiadas o las que sus profesores decían que había que utilizar (Fullan, 1991; Gimén-Sacristán, 2012); y (c) por ello, en la FIP cobra una especial relevancia la coherencia interna entre el discurso del profesorado, su metodología y la estrategia de evaluación (Biggs, 1999; Casero, 2010).

Si queremos que el futuro profesorado implique al alumnado en los procesos de evaluación y lleve a cabo sistemas de evaluación orientados al aprendizaje, es fundamental que lo experimenten durante su formación inicial (Ibernón, 2012; Struyven, Dochy, & Janssens, 2005). No es suficiente con estudiarlo de forma teórica cuando se llega al tema de “Evaluación Educativa”, en las materias de su formación inicial. Estas son las razones que hacen que en FIP sea muy importante predicar con el ejemplo (Fullan, 1991). Existen algunos trabajos que señalan las consecuencias que tiene la incoherencia de recibir discursos sobre cómo hacer la evaluación en la práctica educativa que son completamente contrarios con la evaluación que se lleva a cabo en los estudios de FIP (Aparicio, Tejedor, & Sanmartín, 1982; López-Pastor, 2009; Red IRES, 2010; Tonucci, 2010). Entendemos que esta postura es doblemente errónea en la FIP.

Se han realizado estudios sobre la falta de participación del alumnado en los procesos de evaluación y de negociación curricular sobre el sistema de evaluación de FIP (Ibarra & Rodríguez, 2010; Trillo & Porto, 2002; Wylie, Lyon, & Goe, 2009). Los resultados muestran que la participación del alumnado en el proceso de evaluación revierte positivamente en su aprendizaje (Bain, 2005; Biggs, 1999; Boud, 1995; Boud & Falchikov, 2007; Brown & Glasner, 1999; Falchikov, 2005; Fernández-Balboa, 2006; López-Pastor, 2008, 2009), aunque requieren protocolos en los que se mantenga la fiabilidad y validez (Hanrahan & Isaacs, 2001; Stefani, 1994; Topping, Smith, Swanson, & Elliot, 2000). Biggs (1999) señala que tanto la auto-evaluación como la evaluación entre iguales mejoran el aprendizaje de contenidos y dan ocasión a que los estudiantes aprendan procesos metacognitivos de supervisión útiles y aplicables a lo largo de vida académica y profesional, aunque no siempre haya percepciones similares entre profesorado y estudiantes (Andrade & Du, 2007; Fisher & Miller, 2008; McLellan, 2001; Searby & Ewers, 1997; Struyven, Dochy, & Janssens, 2003). Hay trabajos sobre la puesta en práctica de la evaluación entre iguales en la educación superior (Boud & Falchikov, 2007; Falchikov, 2005; Falchikov & Goldfinch, 2000; Gregory, Yeomans, & Powell 2003; Hanrahan & Isaacs, 2001), pero hay pocos trabajos sobre la aplicación de exámenes escritos con finalidad formativa (DiBattista, Mitterer, & Gosse, 2004; Wininger, 2005).

Examen, evaluación, metodología y proceso de enseñanza-aprendizaje

El examen ha sido el instrumento de evaluación que ha recibido más críticas, al haberse utilizado habitualmente como herramienta de valoración de carácter final y único, que suele acompañar a modelos docentes centrados de forma principal o exclusiva en la clase magistral (Bain, 2005; Carless, 2010; Williams & Wong, 2009; Winger, 2005). Este uso del examen ha sido reprochado como forma de comprobar conocimientos memorísticos, independientemente de la variabilidad en cuanto al tipo de preguntas: respuesta múltiple, respuestas cortas, respuestas amplias, incluso con la utilización de material para su resolución. Además, con estas diferentes formas de presentación en lo que se ha incidido es en la “rendición de cuentas”, es decir, en la comprobación de lo que se supone que tiene que saber el estudiante y que corrige el profesor, aplicando los criterios más o menos conocidos a las respuestas efectuadas. Esta forma de corregir las respuestas de los estudiantes, lleva a que sea el profesor el que compruebe si el estudiante sabe o no, y cuánto. Trasladado al estudiante, sería una forma pasiva de entender su propio aprendizaje, pues su actividad se limita a responder y a esperar el resultado de sus respuestas. Las sugerencias son que es necesario recibir un feedback adecuado a la actuación (Draper, 2009), con el fin de que pueda obtenerse una evaluación que incida en el desarrollo y el aprendizaje y que le ayude al control de su propio aprendizaje; dicho feedback debe realizarse de forma inmediata y adecuada para que pueda incidir en el aprendizaje (Liu & Carless, 2006; Nicol & Macfarlane-Dick, 2006).

Sin embargo, hay formas de plantear los exámenes que permite superar esa idea de instrumento con intención calificadora, convirtiéndole en una herramienta más dentro de un sistema global de evaluación orientado al aprendizaje del alumnado, e incluso como actividad de aprendizaje en sí misma. La forma de entender el proceso de enseñanza y aprendizaje nos lleva a apostar por una participación activa de los estudiantes, implicando al propio estudiante en la corrección de los exámenes, de modo que recibe un feedback inmediato y adecuado sobre su actuación. Es una forma diferente de entender el papel del estudiante, convirtiéndolo en un agente activo en la revisión y

corrección del examen. Los datos aportados desde diferentes materias permiten demostrar que es viable y adecuado un proceso de coevaluación con exámenes en la FIP. Tanto el examen como la evaluación entre iguales forman parte de sistemas globales de evaluación formativa orientada al aprendizaje. Estos sistemas de evaluación entre iguales y evaluación formativa llevan experimentándose y mejorándose varios cursos a través de ciclos y espirales sucesivos de investigación-acción.

En este sentido, este trabajo tiene una triple finalidad: (a) analizar los resultados obtenidos en una experiencia sobre el uso de procesos de evaluación entre iguales a través de la corrección inmediata de exámenes parciales, dentro de sistemas de evaluación orientados al aprendizaje; (b) comprobar el grado de fiabilidad de este tipo de procesos de evaluación entre iguales, contrastando la evaluación y calificación realizada por el compañero con la valoración que realiza posteriormente el profesor de la asignatura; y (c) analizar la valoración que realiza el alumnado sobre la realización de este tipo de experiencias evaluativas y formativas.

Metodología

Diseño

El diseño de esta investigación se basa en un estudio multicaso (Stake, 2006), por tanto no se pretende realizar un estudio comparativo. Se elige este diseño por tratarse de un estudio “expost-facto” y las diferencias existentes entre los tres casos. Tal y como señala Patton (1990), no hay reglas precisas por las que se eligen los estudios de casos, la amplitud de la población en cada caso depende más de lo que uno trata de conocer que del propósito de la investigación. En nuestro estudio estos casos tienen utilidad y credibilidad y se han podido hacer según los recursos y el tiempo disponible. Se analizaron los resultados obtenidos en tres centros de FIP, en una experiencia innovadora consistente en realizar un examen de la asignatura seguido de la evaluación inmediata del examen por un compañero (evaluación entre iguales), con la ayuda de una plantilla especialmente diseñada para ello. Tanto los exámenes como las plantillas utilizadas fueron revisadas posteriormente por el profesor. Tras cerrar el proceso de revisión, se analizó la correlación estadística entre la valoración otorgada por los iguales y la que realizó posteriormente el profesor, para comprobar su fiabilidad.

Participantes

Los datos han sido recogidos durante el curso 2009-10 en tres centros de formación inicial del profesorado, que aparecen detallados en la Tabla 1. En el centro A: 41 estudiantes; en el centro B: 60 estudiantes y, en el centro C: 44 estudiantes. No se realiza ningún tipo de muestreo, sino que se trabaja con el total del alumnado que realizó el examen parcial. Como se explica en la Tabla 2, la experiencia era obligatoria para los alumnos que elegían la opción de evaluación formativa y compartida. A los estudiantes se les pidió permiso para utilizar la información recogida y estuvieron de acuerdo siempre que se mantuviera la confidencialidad, como así se hizo. Los tres centros están situados en la zona centro y norte de España, con una distancia máxima de 300 kms entre los dos centros más alejados y de 90 kms entre los más cercanos. Los tres centros se encuentran en un proceso de reconversión de las antiguas diplomaturas o licenciaturas a estudios de grado y master. En la tabla 1 presentamos las características principales de cada caso y centro

Tabla 1. Datos contextuales y de población de los tres centros en que se lleva a cabo la innovación docente

| | Centro A | Centro B | Centro C |
|-------------------|--------------------------------------------------------|----------------------------------------------|------------------------------------------------------|
| Asignatura | Fundamentos de la estrategia y táctica | Didáctica EF-II | Pedagogía de la Actividad Física |
| Curso | 1º | 3º (último de la titulación) | 4º (último de la titulación) |
| Titulación | Ciencias de la Actividad Física y el Deporte (CAFyD) | Magisterio, especialidad en Educación Física | Ciencias de la Actividad Física y el Deporte (CAFyD) |
| Centro | Facultad de Formación de Profesorado y Educación (UAM) | E.U. Magisterio de Segovia (UVA) | Facultad CAFyD de León |
| Nº de estudiantes | 41 | 60 | 44 |
| Edad estudiantes | Entre 18 y 20 años | Entre 20 y 30 años | Entre 21 y 30 años |

Instrumentos

En este estudio hemos utilizado cuatro técnicas e instrumentos para la recogida de datos, que se describen a continuación:

Examen específico del temario de la asignatura. En los tres casos se realiza un examen sobre el temario de la asignatura, aunque existen ciertas variaciones según el caso. En la Tabla 2 queda explicado qué tipo de examen se realiza en cada caso.

Plantilla de corrección diseñada “ad hoc” para cada examen. Determina los criterios de calificación. Sirve para que cada estudiante corrija el examen de un compañero, así como para la corrección posterior del profesor.

Cuestionario anónimo. Utilizado en los casos A y B. Sirve para comprobar cómo estima la experiencia el alumnado. Las preguntas son las siguientes: (a) ¿qué has aprendido con esta forma de corregir el examen?; (b) ¿qué aspectos positivos le encuentras?; (c) ¿qué aspectos negativos le encuentras?; y (d) ¿qué soluciones aportas a los aspectos negativos? Son preguntas abiertas, de ahí que las respuestas sean muy variadas.

Asamblea conjunta con los grupos de prácticas. Se emplea en el caso C e incluye las mismas preguntas anteriores. Al contar con pocos alumnos, organizados en cuatro grupos prácticos, la libertad de comentar y exponer sus opiniones era generalizada.

Procedimiento

El estudio se lleva a cabo sobre los resultados obtenidos en una experiencia de innovación evaluativa. En la experiencia sólo participa el alumnado que ha elegido la vía continua de aprendizaje y evaluación. La actividad consiste en aplicar un examen parcial en el horario habitual de la asignatura. Al terminar el examen, cada estudiante corrige el examen de otro compañero con una plantilla elaborada por el profesor. El profesor determina quién evalúa a quién, respetando el anonimato y procurando que no haya lazos de afinidad importantes, buscando un mayor nivel de objetividad. En el caso B, se utiliza un sistema de códigos numéricos para que la corrección del examen sea anónima. Después de la corrección entre iguales, el profesor también corrige todos los exámenes. El cuestionario se pasa después de informar de to-

das las calificaciones. Se trata de una actividad que fue previamente acordada al inicio de curso, cuando se establecieron los procedimientos e instrumentos de evaluación y calificación de la opción continua. En la Tabla 2 aportamos explicaciones del proceso en cada caso y centro.

Tabla 2. Procedimiento llevado a cabo en cada caso y centro

| | Centro A | Centro B | Centro C |
|----------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Vías de aprendizaje y evaluación que se ofrecen al alumnado a principio de curso | <p>Continua: Se les presenta una escala graduada y eligen la nota al principio del curso; según el nivel tienen una carga de trabajo mayor o menor. Tienen que hacer cuatro trabajos parciales y uno final; y el examen con apuntes.</p> <p>Final: entrega de trabajo y examen final (50% de la nota cada uno). Tienen que aprobar por separado cada actividad para poder superar la asignatura.</p> | <p>Continua: están obligados a realizar el examen parcial con evaluación entre iguales. Deben aprobar este examen para seguir en dicha vía.</p> <p>Mixta: sólo realizan el examen parcial si quieren. Les cuenta como una actividad más de aprendizaje.</p> <p>Examen final: no tienen que realizar este examen parcial.</p> | <p>Continua: están obligados a realizar el examen parcial (40% de la calificación final) con evaluación entre iguales. Deben aprobar este examen con más de un 4 para seguir en dicha vía. Además, tienen que hacer cuatro trabajos (60% de la calificación final).</p> <p>Final (100%): Realizan un examen de toda la materia y no se tienen que presentar al parcial. El tipo de examen es igual en las dos opciones, y un supuesto práctico.</p> |
| Tipo de examen | Tres preguntas tipo tema a elegir dos, con una extensión máxima de un folio por las dos caras. Aprobar el examen mantiene la vía elegida al principio de curso. El examen tiene un peso del 50% en la nota. | Siete preguntas cortas a contestar en un máximo de 50 minutos. El examen vale el 20% de la calificación final. | 20 preguntas tipo test (20 puntos), 10 preguntas cortas (50 puntos) y una a desarrollar (30 puntos). El examen equivale a un 40% en la calificación final. |
| Plantilla de corrección | Plantilla con los elementos, categorías o características que tiene que contener la respuesta. | Con tres columnas: (1) criterios de corrección y (2) criterios de calificación para cada pregunta; (3) calificación de cada pregunta y el sumatorio final. | Test de respuesta múltiple con cinco posibilidades. Preguntas cortas: las valoraciones fluctúan entre 3, 2, 1 y 0 con sus correspondientes criterios de calificación. |

| | | | |
|----------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Implicaciones de la evaluación entre iguales | El profesor asigna qué examen corrige cada alumno. Si la nota difiere con la del profesor en más de un 10%, los dos estudiantes deben realizar una revisión con el profesor. | El examen vale dos puntos. Si la diferencia entre corrector y profesor es igual o menor a 0.1, se sube una décima. Si la diferencia es igual o mayor a 0.20, se baja una décima. Si la diferencia es más grande debe ir a tutoría. | Tras la coevaluación se produce la autoevaluación, con posibilidad de que sea dialogada tanto con el coevaluador, solicitando explicaciones, como con el profesor, confirmando su valoración. |
|----------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

Respecto al caso C, cabe destacar dos cuestiones: (a) tanto las preguntas tipo test como las cortas son elegidas y diseñadas por los alumnos como parte de los trabajos que forma el 60% de la calificación final; y bajo una serie de criterios preestablecidos en un proceso constante de revisión respecto a la calidad, claridad y diversidad de las preguntas y respuestas seleccionadas de los documentos de examen; y (b) el proceso de autoevaluación y calificación dialogada final se lleva a cabo en una hora de prácticas y por grupos.

Análisis de los datos

Se han realizado dos tipos de análisis de datos: (a) la fiabilidad de las evaluaciones entre iguales; comparando la evaluación del compañero y la evaluación del profesor, utilizando el estadístico correlación de Pearson para cada caso, con un $p < .001$ bilateral, a través del programa informático SPSS 15.0. También se realizó una comparativa de las calificaciones obtenidas en cada caso según la corrección de los iguales y según la corrección del profesor; y (b) la valoración que realiza el alumnado sobre este tipo de procesos y su adecuación con la FIP, a través de los cuestionarios; para ello se llevó a cabo un análisis de contenido, diferenciando dos categorías: (1) valoraciones positivas; y (2) dificultades y aspectos a mejorar.

Limitaciones

Al tratarse de un estudio de casos, este estudio no busca la generalización de resultados, debido a que presentaría limitaciones respecto al nivel de validez externa. En cambio, por los objetivos marcados, sí podría ser transferible a otros casos y contextos similares, con las adaptaciones oportunas.

Resultados

Hemos organizado los resultados en dos grandes apartados. En primer lugar analizamos la fiabilidad estadística de las correcciones de exámenes entre iguales, comparándolas con la corrección posterior del profesor. En segundo lugar recogemos las valoraciones que realizan los estudiantes sobre la experiencia de innovación llevada a cabo, destacando en este apartado los aspectos positivos pero sin dejar de lado los inconvenientes encontrados y que nos muestran los problemas que están por resolver.

Resultados sobre la fiabilidad con la ayuda de plantillas “ad hoc”

En la Tabla 3 se presentan los resultados de la calificación entre iguales, los resultados de la calificación del profesor y la correlación entre ambas en cada uno de los tres casos estudiados. En el sistema universitario español, la equivalencia numérica de estas calificaciones es la siguiente (en una escala de 0 a 10): suspenso (0-4,9), aprobado (5-6,9), notable (7-8), sobresaliente (9-10).

Tabla 3. Resultados de la corrección de exámenes entre iguales respecto a la posterior corrección por parte del profesor

| | Centro A | Centro B | Centro C |
|-------------------------------------------|---------------------------------------------------------------|----------------------------------------------------------------|---------------------------------------------------------------|
| Nº de estudiantes | 41 | 60 | 44 |
| Resultados de la evaluación entre iguales | 5 Sobresalientes 12 Notables 24 Aprobados 0 Suspenso | 1 Sobresalientes 10 Notables 20 Aprobados 29 Suspenso | 3 Sobresalientes 22 Notables 17 Aprobados 0 Suspenso |
| Resultados de la evaluación del profesor | 2 Sobresalientes 24 Notables 15 Aprobados 0 Suspenso | 0 Sobresalientes 11 Notables 22 Aprobados 27 Suspenso | 4 Sobresalientes 16 Notables 19 Aprobados 5 Suspenso |
| Correlación | .66 | .77 | .90 |

Las correlaciones de la evaluación entre iguales con la de los profesores se pueden considerar altas (casos A y B) y muy altas (caso C), aunque presentan cierta variedad. Los valores más altos corresponden al caso C, que se diferenciaba por tener preguntas más cerradas y en cuya elaboración ha participado el alumnado en el desarrollo de cada tema a lo largo de la asignatura.

En la Tabla 3 también puede comprobarse que la variabilidad en las calificaciones asignadas es mayor, e indican que, generalmente, el alumnado es más exigente que el profesorado en la aplicación de los criterios de calificación establecidos. Sin embargo, cuando los criterios de calificación son claros, las fluctuaciones en las notas entre las valoraciones del profesor y del alumno son pequeñas. Debemos recordar que una variación de medio punto puede hacer cambiar una nota de 6.5 (Aprobado) a 7 (Notable) y viceversa; por tanto, dos valoraciones muy similares entre profesor y alumno pueden convertirse en notas finales diferentes si se encuentran en la frontera de dos calificaciones. Esta realidad, que podría parecer injusta o incoherente en la calificación final de la asignatura, queda solventada con la incorporación de los demás trabajos que forman parte de la calificación final por el proceso de evaluación continua y formativa.

Resultados encontrados en el cuestionario pasado al alumnado

Las respuestas del alumnado al cuestionario sobre la experiencia llevada a cabo muestran, en general, una valoración positiva de la experiencia de innovación, como se comprueba en los comentarios que se obtienen de los cuestionarios.

“Esta forma de evaluar me gusta. Debería hacerse igual en todas las asignaturas” (Cuestionario 28, caso C).

“Con esta forma de examinar estamos aprendiendo hasta el mismo día del examen” (Cuestionario 12, caso C).

El alumnado entiende que se trata de una competencia profesional específica de la FIP (ser capaz de corregir adecuadamente trabajos y exámenes de otras personas). Por ello, valoran positivamente el hecho de tener que enfrentarse a este tipo de actividades y comenzar a realizar algunas de las funciones docentes para las que se consideran menos formados.

“Me doy cuenta de que corregir un examen no es fácil”

(Cuestionario 14, caso A).

“Creo que esto es lo que tiene relación con las competencias profesionales” (Cuestionario 57, caso C).

Sin embargo, también surgen quejas en el sentido de que a lo largo de su formación inicial no han tenido que enfrentarse a situaciones similares y tienen miedo de no saber hacerlo correctamente.

“No sé hasta qué punto mi corrección va a ser adecuada, ¿y si me equivoco?” (Cuestionario 34, caso C).

“Pero si le doy puntuación negativa, entonces suspende. Quién soy yo para suspender a un compañero” (Cuestionario 21, caso A).

Estas quejas son más acusadas en el caso B, en el que la correcta o incorrecta evaluación del examen del compañero les puede sumar o restar nota (0.1 sobre 2 puntos; un 5% del valor del examen).

También realizan análisis interesantes sobre los motivos por los que, en algunos casos, sus correcciones son más exigentes que las del profesor. Los estudiantes tienden a justificar las calificaciones que han asignado, porque consideran que la persona que realizó el examen no había contestado adecuadamente.

“Lo que más me cuesta es dar información negativa para mi compañero, pero es que si ha hecho algo mal, tengo que decírselo” (Cuestionario 38, caso A).

“Este examen que acabo de corregir no lo entiendo nada. ¿Cómo se puede escribir tan mal?” (Cuestionario 19, caso C).

“Este examen que he corregido no tiene ni pies ni cabeza. Le pongo un suspenso porque creo que está mal” (Cuestionario 45, caso A).

En general, los alumnos se sienten satisfechos por encontrarse inmersos en procesos de enseñanza y aprendizaje en los que se pasa de ser un mero receptor y demostrador de conocimiento, a individuos comprometidos en su proceso de formación docente.

“Es una forma que yo no había visto. Me ha servido para darme cuenta de lo que había hecho en el examen” (Cuestionario 23, caso A).

“Veo lo que ha contestado mi compañero y me doy cuenta de que está mal, pero es que yo he puesto lo mismo” (Cuestionario 7, caso C).

“Es una forma más de seguir aprendiendo. Yo lo voy a hacer cuando dé clases” (Cuestionario 33, caso A).

En el caso C, al provenir los alumnos de una asignatura, durante el curso anterior, donde ya se habían llevado a cabo experiencias similares, los alumnos reconocían sentirse más cómodos y menos inseguros.

“El haber pasado el curso pasado por situaciones similares me ha permitido afrontar la coevaluación con menos miedo” (Cuestionario 38, caso B).

Los resultados aportan una información positiva de la aplicación de esta forma de corregir (Tabla 4), pero también hay inconvenientes que deben ser tenidos en cuenta. Destacan como aspectos positivos: la incidencia en el aprendizaje, las dificultades para corregir y la exigencia a la hora de corregir. Entre los negativos: su miedo y falta de preparación para hacer algo a lo que no están acostumbrados y sus dudas sobre si serán objetivos a la hora de calificar.

Tabla 4. Ventajas e inconvenientes de la aplicación del examen parcial con corrección inmediata entre iguales

| Casos | Ventajas | Inconvenientes |
|------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Aspectos comunes | Valoran positivamente que el examen les provoque aprendizajes y puedan disponer de feedback inmediato sobre sus propias respuestas. Valoran la novedad del proceso, el experimentar otras formas de evaluación. Es una forma diferente de estudiar, de prepararse y de hacer exámenes más valiosos educativamente. | Miedo a no corregir de manera adecuada y que este hecho pueda perjudicar al compañero corregido. Temor a que los errores puedan influir negativamente en el corrector, en el sentido de bajarles la nota. La falta de preparación en este tipo de sistemas, no haber tenido formación previa. Negación de la responsabilidad de calificar a compañeros. |

Diferencias entre casos

| | | |
|---------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>Caso A</p> | <p>Tomar conciencia de la dificultad de corregir exámenes con preguntas abiertas, incluso con la utilización de apuntes que tienen en el dossier. Poder utilizar los apuntes durante el examen, aunque suponga un examen más complejo que los puramente memorísticos. La vivencia del examen es diferente; según van corrigiendo el examen de otro compañero, se van dando cuenta de si sus respuestas son correctas y en qué han fallado. Este proceso rompe su rutina habitual, de limitarse a contestar el examen y luego esperar la nota del profesor.</p> | <p>Consideran que es difícil precisar una puntuación en preguntas abiertas entre 1 y 4. Piden que se puedan poner decimales. Solicitan más claridad para corregir los exámenes, pues no saben si están haciendo bien esa actividad. Para ello piden una plantilla de corrección más precisa y más completa.</p> |
| <p>Caso C</p> | <p>Poder ver el examen nada más ser evaluado por el compañero, teniendo así la oportunidad de preguntarle y pedirle explicaciones. Valor formativo de la propuesta y su utilidad en el proceso de formación docente, tanto por tener que diseñar preguntas que formarían parte del examen parcial como por realizar el proceso de corregir el examen de un compañero.</p> | <p>Prefieren no discutir con los compañeros que han evaluado su examen. La dificultad de corregir cuando la letra del examen es poco legible o tiene muchas faltas de ortografía y redacción. Reconocen que les genera una predisposición negativa a una calificación alta. Preocupación por no ser justos en la valoración, aunque les tranquilizaba bastante disponer de criterios de corrección claros.</p> |

Discusión

Las correlaciones entre la evaluación de los estudiantes y la del profesor se pueden considerar altas y muy altas, aunque presentan cierta variedad en función del caso. Una posible explicación es que la correlación ha sido más alta cuanto más cerradas son las preguntas y más ha participado el alumnado en la elaboración de las mismas. Los principales meta-análisis sobre evaluación entre iguales en educación superior indican resultados similares a los obtenidos por nosotros, así como las variaciones que pueden encontrarse en función del grado de entrenamiento del alumnado, las herramientas facilitadas, el tipo de estudios y el curso (Boud, 1995; Falchikov, 2005; Falchikov & Goldfinch, 2000). Por ejemplo, Searby & Ewers (1997) señalan este tipo de evaluación como una ventaja, pero en contra señalan la variabilidad que se produce entre cursos iniciales y superiores, igual que ocurre en nuestro caso. El trabajo de Orsmond, Merry & Reiling (2002) sobre exposición oral muestra unos resultados similares (entre 51% y 65% de acuerdos entre estudiantes) y es parecido también al de Davies (2000) sobre ingeniería de computación, con acuerdos en el 61%, o al de Stefani (1994), con diferencias mínimas entre profesorado y estudiantes, y cercano a los resultados que ofrece Topping et al. (2000), con acuerdos positivos entre estudiantes y profesorado en un 60%. En nuestro caso, los resultados son más altos que el estudio de Magin & Helmore (2001) que obtiene correlaciones de .40 a .53.

En cuanto a la información proporcionada por los estudiantes, en nuestro caso hemos detectado que es necesaria una correcta estrategia para la utilización del proceso de evaluación entre iguales y la utilización del feedback, algo que ya ha sido señalado por Liu & Carless (2006). Entre los resultados, encontramos que supone una incidencia en la autorregulación del aprendizaje, también señalado por Draper (2009) y Nicol & Macfarlane-Dick (2006). Este tipo de procesos parecen ayudar a generar nuevas estrategias de aprendizaje en el alumnado, sobre todo cuanto mayor sea la participación del alumno en el proceso de aprendizaje a lo largo de la asignatura. Por otra parte, los resultados parecen indicar que cuando los procesos de coevaluación se realizan con rigor, las valoraciones del alumnado tienden a ser más exigentes que las del propio profesor (López-Pastor, 2008, 2009), aunque también pueden

encontrarse otros estudios con resultados contrarios (Brown & Glasner, 1999; Fernández, González, López, & Manso, 2011), que insisten en la necesidad de utilizar diferentes técnicas para mejorar la concordancia entre estudiantes y profesorado.

El alumnado valora positivamente la realización de este tipo de actividades, al entender que se trata de procesos de aprendizaje y evaluación auténticos y estrechamente relacionados con algunas competencias profesionales. McLellan (2001) señala las diferencias en la percepción de estudiantes, que se centran más en agradecer la forma de evaluación, mientras que los profesores se centran más en el conocimiento y la aplicación del conocimiento; en todo caso, este tipo de propuestas mejoran las expectativas hacia la evaluación y el aprendizaje por parte del profesorado y de los estudiantes (Fisher & Miller, 2008; Stefani, 1998). Por otro lado, existen trabajos que defienden esta necesaria coherencia entre la evaluación que se practica en la FIP y las competencias profesionales que se pretenden desarrollar (Fullan, 1991; López-Pastor, 2009; Red IRES, 2010; Tonucci, 2010).

Por otro lado, las desventajas encontradas por el alumnado en relación al miedo a no corregir de manera adecuada y que este hecho pueda perjudicar al compañero corregido, puede guardar relación con la sensación que muchas veces tiene el profesorado novel sobre si estará valorando y calificando adecuadamente al alumnado, dentro del concepto de autoeficacia docente (Prieto, 2007).

Conclusiones

En este trabajo hemos presentado una experiencia innovadora sobre el uso de procesos de evaluación entre iguales a través de la corrección inmediata de exámenes parciales, dentro de sistemas de evaluación orientados al aprendizaje. Los resultados encontrados en estos tres casos muestran una alta correlación entre la evaluación realizada por el compañero y la valoración que realiza posteriormente el profesorado, aunque con cierta variabilidad entre los casos, que parecen indicar que cuanto más claros son los criterios de calificación utilizados mayor es la correlación. El estudio también muestra que el profesorado en formación inicial valora positivamente este tipo de experiencias evaluativas, tanto por su mayor relación con el desarrollo de competen-

cias profesionales, como por su coherencia con los planteamientos que defienden el desarrollo de sistemas de evaluación formativos, auténticos y orientados al aprendizaje.

Creemos que este trabajo puede ser de interés para el profesorado encargado de la FIP. También para los docentes universitarios que estén interesados en el uso de procesos y técnicas de coevaluación entre el alumnado y en potenciar su implicación en los procesos de evaluación. Puede resultar especialmente interesante para el profesorado europeo que esté implicado en los estudios de FIP establecidos por el proceso de Convergencia hacia el EEES. Por último, puede ser de interés para las personas que estén llevando a cabo investigaciones sobre docencia universitaria y evaluación del aprendizaje del alumnado en la FIP, especialmente en lo relativo a la implicación del alumnado en los procesos de evaluación.

En cuanto a la perspectiva, nuestra intención es continuar investigando los resultados de este tipo de procesos de evaluación orientada al aprendizaje y al desarrollo de competencia profesionales en FIP. Para futuras investigaciones se podrían intentar trabajos cuasi-experimentales que permitan confrontar los datos una vez se opere con variables que relacionen diferentes formas de utilizar los exámenes y el rendimiento académico. Futuros objetos de investigación pueden ser: (a) analizar la influencia de este tipo de actividades formativas en su desempeño profesional posterior como maestros en primaria y secundaria, y (b) promover la extensión de este tipo de procesos de evaluación a través de seminarios y dinámicas de investigación-acción con el profesorado implicado en la FIP.

Referencias

- Algozzine, B., Beattie, J., Bray, M., Flowers, C., Gretes, J., Howley, L. ... Spooner, F. (2004). Student evaluation of college teaching: a practice in search of principles. *College teaching*, 52(4), 134-141.
- Andrade, H. & Du, Y. (2007). Student responses to criteria-referenced self-assessment. *Assessment & Evaluation in Higher Education*, 32(2), 159-181.
- Aparicio, J. J., Tejedor, F. J., & Sanmartín, R. (1982). *La enseñanza universitaria vista por sus alumnos: Un estudio para la evaluación de los cursos de la enseñanza superior*. Madrid: ICE/ Universidad Autónoma de Madrid.
- Arribas, J. M., Carabias, D., & Monreal, I. (2010). La docencia universitaria en la formación inicial del profesorado. El caso de la escuela de magisterio de Segovia. *Revista Electrónica Interuniversitaria de Formación del Profesorado*, 13(3), 27-35. Disponible en:
http://aufop.com/aufop/uploaded_files/articulos/1285861727.pdf
- Bain, K. (2005). *Lo que hacen los mejores profesores universitarios*. Valencia: Universidad de Valencia.
- Biggs, J.B. (1999). *Teaching for Quality Learning at University*. Buckingham: Open University Press.
- Boud, D. (Ed.). (1995). *Enhancing learning through self-assessment*. London: Kogan Page.
- Boud, D. & Falchikov, N. (Eds.). (2007). *Rethinking Assessment in Higher Education. Learning for the long term*. London: Routledge.
- Brown, S. & Glasner, A. (Eds.). (1999). *Assessment Matters in Higher Education*. Buckingham: Open University Press.
- Carless, D. (2007). Learning-oriented assessment: conceptual bases and practical implications. *Innovations in Education and Teaching International*, 44(1), 57-66.
- Carless, D. (2010). *From testing to productive student learning*. London: Routledge.
- Casero, A. (2010). Factores moduladores de la percepción de la calidad docente. *RELIEVE*, 16(2), 1-17. Disponible en:
http://www.uv.es/RELIEVE/v16n2/RELIEVEv16n2_3.htm

- Davies, P. (2000). Computerized Peer Assessment. *Innovations in Education and Training International*, 37(4), 346-355.
- Dawkins, S., Ritz, M-E., & Loudon, W. (2009). Learning by Doing: Preservice Teachers as Reading Tutors. *Australian Journal of Teacher Education*, 34(2), 40-49.
- DiBattista, D., Mitterer, J. O., & Gosse, L. (2004). Acceptance by undergraduates of the immediate feedback assessment technique for multiple-choice testing. *Teaching in Higher Education*, 9(1), 17-28.
- Draper, S.W. (2009). What are learners actually regulating when given feedback? *British Journal of Educational Technology*, 40(2), 306-315.
- Emery, C. R., Kramer, T. R., & Tian, R. G. (2003). Return to academic standards: a critique of student evaluations of teaching effectiveness. *Quality Assurance in Education*, 11(1), 37-46.
- Falchikov, N. (2005). *Improving Assessment Through Student Involvement. Practical solutions for aiding learning in higher and further education*. London: Routledge.
- Falchikov, N. & Goldfinch, J. (2000). Student peer assessment in Higher Education: A meta-analysis comparing peer and teacher marks. *Review of Educational Research*, 70(3), 287-332.
- Fernández, M. A., González, J. L., López, I., & Manso, M. (2011). Evaluación participativa en habilidades para comunicar en 3º de grado de enfermería en el curso 2009/10, la escenificación como método docente y de evaluación. *Revista de Docencia Universitaria*, 8(2), 73-93.
- Fernández-Balboa, J. M. (2006). Dignity and democracy in the college classroom: The practice of self-evaluation. En R. A. Goldstein (Ed.), *Useful Theory: Making Critical Education Practical* (pp. 105-128). New York: Peter Lang Publishing.
- Fisher, R. & Miller, D. (2008). Responding to student expectations: a partnership approach to course evaluation. *Assessment & Evaluation in Higher Education*, 33(2), 191-202.
- Fraile, A. (2004) Un cambio democrático en las aulas universitarias: Una experiencia en la formación del profesorado de Educación Física. *Contextos educativos: Revista de Educación*, 6-7, 213-234.
- Fraile, A. (2006). Cambios en el aula universitaria ante los nuevos retos europeos. *Tandem*, 20, 57-72.
- Fullan, M. (1991). *The New Meaning of Educational Change*. London: Cassell.

- Gimeno-Sacristán, J. (2012). ¿Por qué habría de renovarse la enseñanza en la universidad? En J.B. Martínez (Coord.), *Innovación en la universidad. Prácticas, políticas y retóricas* (pp. 27-51). Barcelona: Graó.
- Gregory, A., Yeomans, E., & Powell, J. (2003). Peer assessment and enhancing student learning. En R. Kaye & D. Hawkrigde (Eds.), *Learning and teaching for business: Case studies in successful innovations* (pp. 108-123). London: Kogan Page.
- Hanrahan, S. J. & Isaacs, G. (2001). Assessing self- and Peer-assessment: the students' view. *Higher Education Research & Development*, 20(1), 53-70.
- Ibarra, M. S. & Rodríguez, G. (2010). Aproximación al discurso dominante sobre la evaluación del aprendizaje en la universidad. *Revista de Educación*, 351, 385-407.
- Ibernón, F. (2012). La formación del profesorado universitario: orientaciones y desorientaciones. Las prácticas de formación. En J.B. Martínez (Coord.), *Innovación en la universidad. Prácticas, políticas y retóricas* (pp. 85-103). Barcelona: Graó.
- Knight, P. (2005). *Calidad del aprendizaje universitario*. Narcea: Madrid.
- Liu, N.F. & Carless, D. (2006). Peer feedback: the learning element of peer assessment. *Teaching in Higher Education*, 11(3), 279-290.
- López-Pastor, V. M. (2008). Implementing a Formative and Shared Assessment System in Higher Education Teaching. *European Journal of Teacher Education*, 31(3), 293-311.
- López-Pastor, V. M. (Coord.). (2009). *La Evaluación Formativa y Compartida en Docencia Universitaria: propuestas, técnicas, instrumentos y experiencias*. Madrid: Narcea.
- LOU. *Ley Orgánica 6/2001, de 21 de diciembre, de Universidades* (BOE nº 307. Lunes 24 de diciembre de 2001).
- Magin, D. J. & Helmore, P. (2001). Peer and Teacher Assessments of Oral Presentation Skills: How Reliable Are They. *Studies in Higher Education*, 26(3), 287-298.
- Martínez, M.A. (Coord.) (2006). *Investigar en docencia universitaria: redes de colaboración para el aprendizaje*. Alcoy: Marfil.
- McLellan, E. (2001). Assessment for learning: the differing perception of tutors and students. *Assessment & Evaluation in Higher Education*, 26(4), 307-318.

- Navarro, V., Santos, M.L., Buscá-Donet, F., Martínez-Mínguez, L., & Martínez-Muñoz, L.F. (2010). La experiencia de la red universitaria española de evaluación formativa y compartida: proceso y abordaje. *Revista Iberoamericana de Educación*, 52(7). Disponible en: <http://www.rioei.org/expe/3428Navarro.pdf>
- Nicol, D.J. & Macfarlane-Dick, D. (2006). Formative assessment and self-regulate learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199-218.
- Orsmond, P., Merry, S., & Reiling, K. (2002). The Use of Exemplars and Formative Feedback when Using Student Derived Marking Criteria in Peer and Self-assessment. *Assessment & Evaluation in Higher Education*, 27(4), 309-323.
- Patton, M.Q. (1990). *Qualitative Evaluation and Research Methods*. Newbury Park: Sage.
- Pérez-Pueyo, A., Taberero, B., López-Pastor, V.M., Ureña Ortín, N., Ruiz-Lara, E., Capllonch, M., ... Castejón Oliva, F. J. (2008). Evaluación formativa y compartida en la docencia universitaria y el Espacio Europeo de Educación Superior: Cuestiones clave para su puesta en práctica. *Revista de Educación*, 347, 435-451. Disponible en: http://www.revistaeducacion.mec.es/re347_20.html
- Perrenoud, Ph. (2004). *Diez nuevas competencias para enseñar*. Barcelona: Graó.
- Prieto, L. (2007). *Autoeficacia del profesor universitario: eficacia percibida y práctica docente*. Madrid: Narcea.
- Red IRES. (2010). No es verdad. Manifiesto pedagógico. *Educarnos*, 49, 5-8.
- Santos, M. L., Martínez, L. F., & López-Pastor, V. M. (Coords.). (2009). *La innovación docente en el EEES*. Almería: Universidad de Almería.
- Searby, M. & Ewers, T. (1997). An evaluation of the use of peer assessment in higher education: a case study in the school of music, Kingston University. *Assessment & Evaluation in Higher Education*, 22(4), 371-384.
- Stake, R. E. (2006). *Multiple Case Study Analysis*. New York: Guilford Press.
- Stefani, L. (1994). Peer, self and tutor assessment: Relative reliabilities. *Studies in Higher Education*, 19(1), 69-76.

- Stefani, L. (1998). Assessment in partnership with learners. *Assessment & Evaluation in Higher Education*, 23(4), 339-350.
- Struyven, K., Dochy, F., & Janssens, S. (2003). Students' perceptions about new modes of assessment in Higher Education. En M. Siegers, F. Dochy & E. Cascallar (Eds.), *Optimising new modes of assessment: In search of qualities and standards* (pp. 171-223). London: Kluwer.
- Struyven, K., Dochy, F., & Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: a review. *Assessment & Evaluation in Higher Education*, 30(4), 325-341.
- Taras, M. (2002). Using assessment for learning and Learning from Assessment. *Assessment and Evaluation in Higher Education*, 27(6), 501-510.
- Tonucci, F. (2010). Apuntes sobre el manifiesto. *Educar(nos)*, 49, 9-11.
- Topping, K. J., Smith, E. F., Swanson, I., & Elliot, A. (2000). Formative peer assessment of academic writing between postgraduate students. *Assessment & Evaluation in Higher Education*, 25(2), 149-169.
- Trillo, F. & 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.
- Wylie, E. C., Lyon, C. J., & Goe, L. (2009). *Teacher Professional Development Focused on Formative Assessment: Changing Teachers, Changing Schools*. Princeton: Educational Testing Service. Disponible en:
<http://www.ets.org/Media/Research/pdf/RR-09-10.pdf>
- Williams, J. B. & Wong, A. (2009). The efficacy of final examinations: A comparative study of closed-book, invigilated exams and open-book, open-web exams. *British Journal of Educational Technology*, 40(2), 227-236.
- Wininger, S. R. (2005). Using Your Tests to Teach: Formative Summative Assessment. *Teaching of Psychology*, 32(3), 164-166.

Víctor Manuel López-Pastor is Associate Professor at Teachers College-Segovia, Universidad de Valladolid.

Javier Castejón is Associate Professor at Teachers College, Universidad Autónoma de Madrid.

Ángel Pérez-Pueyo is Assistant Professor at the Faculty of Sports and Physical Activity, Universidad de León.

Contact Address: EU. Magisterio Segovia. Plaza de Colmenares, 1. (40001) Segovia, Spain. Email: vlopez@mpc.uva.es