

# **Reflexiones en torno a una experiencia de Evaluación Continua en el EEES**

**Francisco José PLAZA MARTÍN <sup>(1)</sup>, Esteban GÓMEZ GONZÁLEZ <sup>(1)</sup>, Ana Cristina LÓPEZ MARTÍN <sup>(1)</sup>, Andrés A. FERNÁNDEZ-FUERTES <sup>(2)</sup>.**

(1) Departamento de Matemáticas (Universidad de Salamanca)

(2) Departamento de Educación (Universidad de Cantabria)

Agradecimientos: Esta experiencia fue desarrollada en el marco del proyecto de innovación docente MD/032 de la Universidad de Salamanca.

## Reflexiones en torno a una experiencia de Evaluación Continua en el EEES

En el nuevo modelo docente en la universidad española, iniciado como consecuencia de la adaptación al Espacio Europeo de Educación Superior, se plantean numerosos cambios en las metodologías docentes. Estas deben estar centradas en el aprendizaje, y no sólo en la enseñanza: cobra una especial relevancia la adquisición, por parte del estudiante, de las competencias que se encuentran en el diseño de los nuevos títulos de Grado y configuran el perfil de egreso.

Estos cambios suponen que la participación del estudiante sea más activa y constante en su proceso de aprendizaje (González-Martín y Ramos Álvarez, 2007, 9-10); proceso que, por otro lado, el profesor debe dirigir, estimular y fomentar, pero también supervisar y valorar (Delgado y Olivier, 2009, 2-3). En consecuencia, es lógico dotar a la evaluación de una doble misión, pues ha de tener fines tanto pedagógicos como acreditativos.

La evaluación continua ha sido propuesta por diversos autores como un buen sistema para valorar la adquisición de competencias, ya que consiste en un seguimiento individualizado del proceso de aprendizaje a lo largo de todo el periodo docente. No obstante, habitualmente no se emplean sistemas “puros” de evaluación continua, sino que a estos se añaden métodos o procedimientos característicos de otros sistemas de evaluación formativa, evaluación sumativa o autoevaluación (Cabrera, 2003, 17-36; Delgado y Olivier, 2009, 3-8; López, 2001, 29-32). Estas modificaciones pretenden, en gran parte de los casos, rentabilizar al máximo las posibilidades de retroalimentación del estudiante de modo que, siendo conocedor de sus fortalezas y debilidades desde el

primer momento, pueda tomar un papel protagonista en la evolución de su propio aprendizaje a lo largo del periodo lectivo (Delgado y Olivier, 2006, 1-2).

No podemos dejar de mencionar otra faceta del sistema de evaluación, íntimamente ligada a la función acreditativa, como es la de la formación de la calificación. La prueba final goza de una fiabilidad tradicionalmente reconocida pero, bajo un sistema de evaluación continua, surgen no pocas preguntas sobre la ponderación de las distintas actividades, así como de la validez y fiabilidad de las mismas a la hora de medir la calidad del aprendizaje (Martínez Martínez y Cadenato Matía, 2010, 5-6). En principio, el proceso de formación de la calificación puede utilizarse o no como estímulo en el autoaprendizaje del estudiante, con la salvaguarda de que la respuesta positiva a todos los estímulos puede no implicar *per se* la acreditación final positiva.

Admitiendo que los procesos de enseñanza-aprendizaje y los sistemas de evaluación están estrechamente relacionados (Coll, Rochera Villach, Mayordomo Saíz y Naranjo, 2007, 786), somos de la opinión de que cabe separar el concepto de “trabajo continuado” del de “formación de la calificación”. Es decir, que aunque la interrelación de ambos conceptos es patente, cada uno puede y debe ser medido por separado.

Las conclusiones obtenidas a partir de nuestra experiencia pueden ser interpretadas desde dos ópticas distintas, pero a la vez complementarias. Tomando el símil científico de la utilización de un aparato para medir propiedades de un objeto, resulta que las mediciones nos permiten deducir ciertas propiedades tanto del aparato - en nuestro caso, el sistema de evaluación -, como del objeto, a saber, en nuestro caso, el grupo de estudiantes. En otras palabras, discutimos las posibles interpretaciones de estos datos tanto en términos de la metaevaluación del propio sistema de evaluación como, por otro lado, en términos del rendimiento del grupo de estudiantes.

En consecuencia, los objetivos específicos de este estudio consisten en la *metaevaluación de un sistema de evaluación* a partir de una experiencia concreta, así como en las correspondientes interpretaciones en términos de la adecuación de la ponderación de las distintas actividades del proceso enseñanza-aprendizaje, y de la determinación de potenciales *indicadores de rendimiento o dedicación individual y grupal*. En futuros trabajos, convendría someter este estudio a otros dos análisis consistentes en, por un lado, la percepción que el estudiante tiene del sistema empleado y, por otro, la verificación externa y/o comparada de la adquisición de competencias de modo efectivo.

Llegados a este punto creemos adecuado describir brevemente la “atmósfera” en la que surgió esta experiencia. La implantación del Grado en Matemáticas en la Universidad de Salamanca se materializó en el curso académico 2008-2009: fue la culminación de un proceso que, basado en acciones de mejora de la calidad en la enseñanza y en proyectos de innovación docente, estaba orientado al fomento de la adquisición de competencias y a la mejora de los resultados de los estudiantes. Como profesores responsables de una de las asignaturas troncales y a la vista de nuestras experiencias previas en el Proyecto Piloto de Adaptación al EEES [1], nos propusimos adoptar nuevos métodos de evaluación adaptados al EEES, cambiando el procedimiento de “evaluación tradicional” y diseñando un modelo de evaluación basado en un sistema de evaluación continua pero enriquecido con características de otros métodos como evaluación formativa, autoevaluación y evaluación final.

## **Método**

### **Participantes**

Nuestra experiencia se realizó en la titulación “Grado en Matemáticas” de la Universidad de Salamanca durante el curso académico 2008-2009. El primer curso de dicha titulación adaptada al EEES consta de un total de 60 créditos ECTS distribuidos en 10 asignaturas semestrales, a razón de 6 créditos por asignatura e impartidos en dos semestres de 15 semanas cada uno. Tanto las metodologías de enseñanza-aprendizaje como los sistemas de evaluación de todas estas asignaturas estaban diseñados dentro del marco del EEES. No obstante, estos últimos podían variar ligeramente de unas materias a otras a elección de los docentes.

Más concretamente, se trabajó con el grupo de estudiantes matriculados en la asignatura “Álgebra Lineal I” impartida en el primer semestre de primer curso. Las actividades docentes presenciales programadas consistieron en cuatro horas semanales, repartidas en dos clases teóricas, una de problemas y un seminario de trabajo tutelado, en el que el grupo se dividía en dos mitades.

El grupo docente contaba con un total de veintinueve estudiantes (dieciocho hombres y once mujeres) que han seguido, salvo tres abandonos tempranos, el curso en su totalidad. Todos los estudiantes estaban matriculados del curso completo y todos ellos eran estudiantes de primera matrícula. La calificación media de ingreso fue de 7,43 y veintiséis de los veintinueve habían elegido estudiar Matemáticas como primera opción. El distrito de procedencia del 69% de los estudiantes era el de la Universidad de Salamanca (Ávila, Salamanca y Zamora).

## **Variables e instrumentos**

Las variables que se han utilizado en este estudio son:

- La calificación en actividades presenciales (AP), obtenida a través de dos pruebas de control (C1 y C2) y un examen final (E).
- La calificación en actividades no presenciales (ANP), medida con dos trabajos teóricos (TT1 y TT2) y tres trabajos de problemas (TP1, TP2 y TP3).
- La nota del examen (E), obtenida en un examen final.
- El nivel de dedicación del estudiante, medido por pregunta directa a través de encuestas.

## **Descripción del Sistema de Evaluación y Procedimiento**

Como se ha expuesto en la introducción, nuestro objetivo era disponer de un sistema de evaluación que tuviera en cuenta el proceso enseñanza-aprendizaje durante todo el periodo lectivo y no sólo al final, al tiempo que fuera capaz de medir el grado de adquisición de competencias, para cumplir con la función acreditativa del sistema (Coll, Rochera Villach, Mayordomo Saíz y Naranjo, 2007, 800). Nuestro sistema de evaluación se componía de actividades de distinta naturaleza (i.e., la evaluación formativa y la sumativa, la evaluación continua, final y/o la autoevaluación) y tenía distintos objetivos (e.g., motivar al estudiante, proporcionarle información sobre su proceso de aprendizaje, medir el grado de adquisición de competencias, formación de la calificación, etc.). Estas actividades se dividieron en 3 grupos: actividades presenciales, no presenciales y actividades voluntarias. La participación en cada una de ellas no era estrictamente obligatoria, aunque para poder superar la materia sí se requería una nota mínima en cada una. El desarrollo de tales actividades fue el que a continuación se describe.

El primer tipo de actividad no presencial consistió en la realización de dos trabajos teóricos (TT1 y TT2). El primer trabajo fue entregado en la cuarta semana del curso y el segundo en la decimocuarta. Los trabajos consistieron en pequeños desarrollos teóricos de la materia que se estaba tratando en la asignatura. Para su realización, el estudiante contó con la bibliografía o material de consulta necesario, las pautas de lectura, las indicaciones precisas por parte del docente y la ayuda necesaria ofrecida a través de las tutorías y de la plataforma de enseñanza virtual Moodle. Para dar lugar a una retroalimentación y, en consecuencia, dotarlas también de una función formativa (y no sólo evaluativa), los trabajos eran entregados por escrito al profesor quien, tras su corrección y si así lo consideraba pertinente, llamaba a tutoría al estudiante con dos objetivos. El primero era informar al estudiante sobre las deficiencias y/o errores de comprensión cometidos y poder, de esta forma, ayudarlo a corregirlos con prontitud, hecho que mediante el sistema de evaluación tradicional sólo se podía detectar en el examen. Y segundo, confrontar si el trabajo había sido realizado personalmente por el estudiante, garantizando de este modo la función acreditativa de la actividad. Para ello el recurso que se habilitó fue el de exposiciones orales en las que el estudiante debía contestar las preguntas que el docente consideraba oportunas. Este medio sirvió además para valorar otros aspectos del trabajo realizado y poder, en caso necesario, matizar la calificación. Los trabajos de teoría contabilizaron un 18% de la calificación definitiva.

El segundo tipo de actividad que sirvió para completar la calificación no presencial fue la entrega de trabajos de problemas. Se realizaron un total de tres trabajos de problemas (TP1, TP2 y TP3) en la segunda, sexta y novena semana del curso. En cada uno de estos trabajos al estudiante se le proponían dos o tres ejercicios o casos prácticos que debía resolver en aproximadamente cuatro días y entregar por escrito al profesor. Como en el caso de los trabajos de teoría, tras la corrección de los mismos, el docente

informaba al estudiante de los errores cometidos a fin de poder subsanarlos de manera temprana. Para esta actividad, no se habilitó ningún sistema de exposición oral o confrontación del trabajo; la razón fue que este método se concibió, más que como un sistema de evaluación, como un instrumento para potenciar el “trabajo continuo” del estudiante en cuestiones básicas de la materia. No obstante, y a fin de motivar al estudiante a participar en este tipo de actividades, a estos trabajos de problemas se les asignó un peso del 12% de la calificación final. En definitiva, el conjunto de todas estas actividades no presenciales suponía hasta un máximo del 30% de la calificación definitiva.

No queriendo olvidar la función acreditativa, se diseñaron ciertas pruebas de control de carácter presencial realizadas en horario lectivo. Cada control contenía diez cuestiones teórico/prácticas, esencialmente de tipo test, aunque también se incluyeron preguntas de respuesta corta. El primer control tuvo lugar en la séptima semana del curso y el último en la decimotercera. La fecha del segundo estaba recogida ya en la planificación docente y constaba en la guía académica de la asignatura. Para la realización del primero, la guía académica reseñaba el hecho de su existencia en el rango de dos semanas, aunque no se especificó el día concreto con el objeto de interferir lo menos posible en el desarrollo de las otras materias. Realizados en presencia del profesor, ambos tuvieron una duración de una hora. Los controles eran pues un tipo de actividad con tintes de evaluación sumativa, ya que esa era su principal finalidad. Además, al complementarse con tutorías de devolución de resultados y de corrección, estos cumplían con la función formativa permitiendo al estudiante regular su propio proceso de aprendizaje. El conjunto de estas pruebas de control contabilizaban hasta un máximo del 30% de la calificación definitiva.



Como evaluación final se realizó un examen (E en Gráficas y Tablas) que constaba también de una parte teórica y una práctica, y su peso en la calificación definitiva era del 40%. La parte teórica incluía tanto cuestiones de tipología similar a las de las pruebas de tipo test como el desarrollo de ciertas definiciones, propiedades, demostraciones o razonamientos propios de este tipo de materias. En la parte práctica el estudiante debía resolver tres o cuatro problemas propuestos. Tuvo lugar en la fecha prevista en la planificación docente de la materia, con una duración aproximada de 4 horas. El carácter acumulativo de los conocimientos y la adquisición de competencias en asignaturas de este carácter, como en general de las matemáticas, pudo valorarse a través de este protocolo de evaluación.

Además de estas, a lo largo del curso se fueron proponiendo ciertas actividades voluntarias. Su calificación fue esencialmente cualitativa y sirvió para completar el proceso evaluativo y finalmente confeccionar la calificación definitiva de cada persona.

Para aquellos estudiantes que con estas actividades no superaron la asignatura, se estableció un sistema de recuperación. En función de las calificaciones obtenidas por el estudiante en cuestión en cada una de las pruebas que se habían realizado, la recuperación varió desde el diseño de actividades extraordinarias hasta la realización de una prueba escrita.

Sobre la recogida de datos, los resultados relativos a las calificaciones corresponden a las notas obtenidas por los estudiantes en las diferentes actividades de evaluación programadas. Los datos relativos al nivel de dedicación del estudiante fueron obtenidos por una metodología distinta a la anterior, a saber, la pregunta directa en tres encuestas. La primera encuesta se realizó en la “Jornada de Acogida” el primer día del curso (15 de Septiembre de 2008) por el Decanato de la Facultad. La segunda tuvo lugar la quinta semana y fue realizada por el Coordinador de Titulación y, por último, la tercera

encuesta tuvo lugar en la octava semana de curso y fue uno de los profesores de la materia quien la llevó a cabo. Ninguna de ellas contenía preguntas de control y en los tres casos a los participantes se les preguntó de modo explícito por su nivel de trabajo tanto en general, como en la materia en cuestión o en alguna actividad específica de la misma. Previamente los estudiantes habían sido informados de que todos los datos recogidos eran anónimos y confidenciales.

## Resultados

Los resultados que se exponen a continuación pertenecen a los 26 estudiantes que siguieron la materia en su totalidad. Para cada uno de ellos se tienen diversos datos que consisten en la puntuación en cada actividad prevista en el sistema de evaluación anteriormente explicado. En la Tabla 1 se recogen los resultados obtenidos individualmente por cada estudiante. Cada casilla representa una puntuación sobre 10, cada fila de la tabla corresponde a un estudiante y cada columna a una determinada actividad. Las filas están ordenadas por orden creciente de la calificación final y las columnas según la semana de realización de la actividad.

TABLA 1: Puntuación de los estudiantes a cada una de las actividades

TP1	TT1	TP2	C1	TP3	C2	TT2	E
4,0		7,0	3,7	0,0	0,0	2,0	0,5
6,0	4,0	8,5	2,3	10,0	0,7		0,0
8,5	1,0	5,5	0,0	0,0	1,0	5,0	1,5
8,5		6,5		10,0	1,7	4,0	0,8

3,5		4,0	2,3	10,0	1,7	2,0	1,6
9,5	3,0	8,0	2,0	0,0	2,7	7,0	1,3
9,0	3,0	6,0	3,7	8,0	2,3	4,0	3,5
9,0	4,0	5,5	4,0	7,0	2,7	4,0	3,5
9,5	5,0	8,5	4,7	10,0	6,0	3,0	2,0
9,0	8,5	8,0	5,0	0,0	2,7	4,5	3,2
10,0	2,0	7,5	6,0	3,0	1,3	3,0	5,0
10,0	9,5	7,5	5,7	8,5	4,3	4,0	1,5
7,5	3,0	6,0	6,3	10,0	2,0	3,5	5,6
10,0	9,5	8,0	1,0	8,5	4,7	5,0	4,8
9,0	7,0	6,5	3,7	10,0	3,7	4,5	5,0
10,0	7,5	5,5	4,7	10,0	3,0	5,5	4,8
9,5	7,5	8,5	3,0	9,5	4,3	9,5	4,5
10,0	6,0	5,0	5,0	6,5	5,0	8,5	5,6
8,5	2,0	4,0	7,7	8,5	6,3	4,0	6,2
9,5	8,0	9,0	7,3	8,5	6,3	9,0	4,9
9,5	5,0	9,5	5,0	9,5	7,7	9,5	6,8
10,0	7,0	7,0	8,7	10,0	6,7	6,0	6,9
9,0	7,0	10,0	4,7	9,0	7,7	8,5	9,0
10,0	9,5	7,0	9,0	8,0	4,7	10,0	8,5
9,5	10,0	10,0	9,0	10,0	6,3	9,5	9,3
10,0	10,0	9,5	10,0	10,0	10,0	9,5	9,5

Para analizar los niveles de exigencia en cada actividad y la adecuación de los pesos asignados a cada una de ellas, calculamos para cada sujeto la media ponderada de las

notas obtenidas en las actividades de tipo presencial, en las de tipo no presencial y en las de control. Además calculamos la desviación típica ponderada (de las notas de cada sujeto respecto de su media ponderada) en estas tres situaciones (Tabla 2). Hay que tener en cuenta que los pesos se normalizaron para que la suma de los datos involucrados en cada caso volviera a ser 10.

TABLA 2: Calificaciones, medias ponderadas y desviaciones típicas ponderadas de las actividades

TODAS LAS ACTIVIDADES (Calificación Definitiva)				ACTIV. PRESENCIALES						ACTIV. NO PRESENCIALES		
				Controles			Examen	Todas				
Calificación Definitiva	Desv. Típica	Pond.	Media ponderada	Desv. Típica	Pond.	Examen (E)	Media ponderada	Desv. Típica	Pond.	Media ponderada	Desv. Típica	Pond.
1,37	1,79		1,83	1,83		0,5	1,07	1,37		2,07	2,47	
1,79	2,73		1,50	0,84		0,0	0,64	0,92		4,47	4,45	
1,85	2,02		0,50	0,50		1,5	1,07	0,59		3,67	3,36	
1,91	2,65		0,83	0,83		0,8	0,78	0,54		4,53	4,44	
2,12	1,80		2,00	0,34		1,6	1,77	0,29		2,93	3,22	
2,80	2,42		2,33	0,33		1,3	1,71	0,58		5,33	3,99	
3,85	1,55		3,00	0,67		3,5	3,28	0,50		5,17	2,58	
3,98	1,35		3,33	0,67		3,5	3,43	0,45		5,27	2,21	
4,24	2,42		5,33	0,67		2,0	3,43	1,71		6,13	3,34	

4,28	2,17	3,83	1,17	3,2	3,47	0,83	6,17	3,55
4,35	2,08	3,67	2,34	5,0	4,40	1,66	4,23	2,83
4,35	2,88	5,00	0,67	1,5	3,00	1,78	7,52	3,98
5,02	1,90	4,17	2,17	5,6	5,00	1,59	5,08	2,48
5,11	2,47	2,83	1,83	4,8	3,93	1,53	7,88	3,41
5,15	1,62	3,66	0,00	5,0	4,43	0,66	6,85	2,55
5,24	1,78	3,83	0,83	4,8	4,36	0,71	7,30	2,76
5,53	2,23	3,67	0,67	4,5	4,14	0,60	8,77	3,36
5,92	1,27	5,00	0,00	5,6	5,36	0,31	7,22	2,10
5,96	1,70	7,00	0,67	6,2	6,54	0,59	4,60	2,84
6,61	1,64	6,83	0,50	4,9	5,71	1,02	8,70	2,16
7,04	1,54	6,33	1,33	6,8	6,57	0,90	8,15	2,34
7,30	1,08	7,66	1,00	6,9	7,21	0,76	7,50	1,58
7,96	1,56	6,16	1,50	9,0	7,78	1,71	8,38	1,10
8,20	1,63	6,83	2,17	8,5	7,78	1,64	9,18	1,45
8,93	1,14	7,67	1,34	9,3	8,57	1,17	9,78	0,88
9,74	0,25	10,00	0,00	9,5	9,71	0,25	9,78	0,25

Finalmente, enfocando ahora el estudio a todo el grupo en lugar de a cada individuo, se han calculado algunas correlaciones entre los datos de varias parejas de columnas de la Tabla 2 (Tabla 3).

TABLA 3: Correlaciones entre las variables objeto de estudio

		CORRELACIÓN
--	--	-------------

Calificación Definitiva (CD)	Nota del Examen (E)	0,96
Calificación Definitiva (CD)	Media ponderada de los Controles (MC).	0,93
Calificación Definitiva (CD)	Media ponderada de las Actividades Presenciales (MAP))	0,98
Calificación Definitiva (CD)	Media ponderada de las Actividades No Presenciales (MANP).	0,87
Media ponderada de las Actividades Presenciales (MAP)	Media ponderada de las Actividades No Presenciales (MANP).	0,77

Los datos así obtenidos se han de interpretar como indicadores del propio sistema de evaluación (metaevaluación) o bien del comportamiento o rendimiento del grupo, y no de modo individual.

Para estudiar la incidencia en la calificación definitiva del rendimiento de los estudiantes en las pruebas presenciales y no presenciales, presentamos unas nuevas tablas o gráficos, dónde utilizamos la desviación típica ponderada y dividimos los estudiantes en dos grupos, aprobados y suspensos. Estos resultados nos permitirán obtener nuevas conclusiones sobre el sistema de evaluación e intuir algunas consideraciones más específicas sobre el rendimiento grupal de los estudiantes.

TABLA 4: Variables más relevantes en función cualitativa de la calificación

(aprobado/suspense)

		Valor Medio de la Desviación	Correlación de la Desviación con la Calificación Definitiva (CD)
Estudiante s aprobados	Desviación Típica Ponderada de las Actividades Presenciales (DAP)	1,56	0,02
Estudiante s aprobados	Desviación Típica Ponderada de las Actividades No Presenciales (DANP)	2,09	-0,92
Estudiante s suspensos	Desviación Típica Ponderada de las Actividades Presenciales (DAP)	2,15	0,40
Estudiante s suspensos	Desviación Típica Ponderada de las Actividades No Presenciales (DANP)	3,37	-0,25

Los datos obtenidos a través de preguntas directas en las tres encuestas realizadas son los siguientes: en la primera de las encuestas, realizada por el Decanato, una de las preguntas era: “¿Cuánto piensas que habrá que estudiar?”, a lo que el 67% opinan “A diario, un par de horas”, un 19% opinan “Todos los días, cuatro horas o más”, y un 14% dan otras respuestas.

En la encuesta realizada por el Coordinador de la Titulación, preguntados específicamente por la materia “Álgebra Lineal”, tenemos la pregunta: “Tu trabajo personal, ¿se corresponde con la dedicación prevista?”, a lo que contestaron “Sí” un 69% y “No” un 31%.

Finalmente, en la encuesta realizada por uno de los profesores de la asignatura que se interesó por el trabajo dedicado a la preparación del seminario semanal, se preguntó por las horas dedicadas a la semana, así como el número de ejercicios intentados y/o resueltos. Los resultados obtenidos mostraron claramente una alta proporcionalidad entre tiempo y cantidad de ejercicios abordados, tal y como se muestra en la Tabla 5.

TABLA 5: Horas de dedicación y la cantidad de problemas intentados o resueltos

%	Horas de dedicación	Cantidad de problemas intentados o resueltos
23%	No contesta.	9
41%	Menos de 1h 30m.	3
27%	Entre 1h 30m y 2h 30m	8
9%	Más de 2h 30m	14

Para comprender estos resultados, tengamos en cuenta que de acuerdo a la dedicación prevista en términos de trabajo autónomo (dentro del ECTS y según la guía de esta asignatura), la dedicación de un estudiante medio debía haber sido 4 horas diarias para las 5 asignaturas y, en particular, 4 horas a la semana para la materia en cuestión, de las cuales deberían ser 2 horas semanales para la parte práctica de la asignatura. Por tanto, de los resultados anteriores, tenemos que al principio de curso, un 67% pensaba trabajar de forma autónoma menos tiempo del previsto (primera encuesta), pero posteriormente



en la semana quinta este porcentaje se redujo al 31% (segunda encuesta) en el trabajo en nuestra asignatura. Sin embargo en la semana octava, pasado el ecuador del semestre, ese porcentaje aumentó al 41% en la parte práctica de la materia.

## **CONCLUSIONES Y DISCUSIÓN**

Las conclusiones que presentamos se articulan en torno a tres ejes: la metaevaluación del sistema implantado, la búsqueda de indicadores de rendimiento del estudiante y la relación entre rendimiento y dedicación.

Comencemos por las conclusiones sobre la metaevaluación del propio sistema. En los resultados de la Tabla 3 se aprecia una fuerte correlación entre los diferentes datos comparados. Por un lado, la alta correlación entre la calificación definitiva y la nota del examen indica que podíamos haber considerado la nota del examen como “patrón” para realizar comparaciones entre las distintas actividades. El valor de esta nota es que el examen se ha realizado de modo presencial y bajo la supervisión/vigilancia del profesor.

Por otro lado, la alta correlación entre la calificación definitiva y las diferentes actividades previstas en el método de evaluación, con sus ponderaciones en la calificación definitiva, nos permite concluir que existen niveles homogéneos de exigencia en las diferentes pruebas o, si queremos, niveles homogéneos de rendimiento del estudiante en las distintas pruebas. Por tanto, podemos asegurar que el modelo de evaluación diseñado da muestras de ser homogéneo y equilibrado.

El segundo bloque de conclusiones analiza si alguno de los indicadores presentados en el capítulo anterior puede servir para medir el rendimiento de los estudiantes. Aunque las correlaciones siguen siendo altas, observamos que descienden con respecto a las anteriores cuando aparece en alguna comparación la media de las actividades no

presenciales. Esto puede ser debido a que estas últimas actividades no se han realizado bajo la supervisión/vigilancia del profesor y los estudiantes han obtenido calificaciones más altas que en otro tipo de actividades. Por tanto, deducimos que el rendimiento del estudiante en estas pruebas no ha sido tan homogéneo con respecto a las pruebas no presenciales, como sí ha ocurrido en las comparaciones de la Tabla 3.

Para estudiar con más profundidad la diferencia en el rendimiento de los estudiantes en las pruebas presenciales y en las no presenciales, utilizamos la Tabla 4, en la que hemos dividido el grupo de los estudiantes en dos subgrupos, aprobados y suspensos. Entre los estudiantes suspensos la correlación entre la desviación típica ponderada de las actividades no presenciales y la calificación definitiva es baja y, además, es negativa, en otras palabras, a mayor calificación definitiva, menor desviación. Sin embargo, entre los estudiantes aprobados la correlación es muy alta y, como en el caso anterior, es negativa. Si observamos el valor medio de la desviación típica (DANP) se observa que entre los estudiantes suspensos este valor medio es más alto, que entre los estudiantes aprobados (la diferencia es mayor que un punto).

De estos resultados se concluye que los estudiantes suspensos tuvieron una gran fluctuación entre sus notas en las actividades no presenciales y no guardaron un patrón claro, en clara contraposición a los estudiantes aprobados.

Podemos aventurar alguna interpretación o explicación de estos hechos: un valor bajo en DANP significa que el estudiante ha tenido poca variabilidad en sus puntuaciones a lo largo del curso debido, por ejemplo, a:

- Se ha organizado de manera sistemática desde el comienzo, ya que las actividades no presenciales se organizaron desde comienzo del curso y con una distribución temporal homogénea a lo largo del periodo docente.

- No ha abandonado una parte del temario (por ejemplo, por tener actividades de evaluación de otras asignaturas o por razones personales), pues ello provocaría tener notas más bajas que su media en algunas actividades.
- No ha plagiado tareas de otros compañeros mejores, pues ello conllevaría notas más altas que su media.

Desde la perspectiva de la metaevaluación, demostraría que el nivel de presión o exigencia de las actividades diseñadas en el sistema de evaluación ha sido sostenido y similar a lo largo del curso.

En los estudiantes aprobados se observa una alta correlación con signo negativo entre la calificación definitiva (CD) y la desviación típica de las actividades no presenciales (DANP), es decir, el índice DANP es menor en aquellos estudiantes con CD mayor. Este hecho muestra que efectivamente el índice DANP constituye un indicador del rendimiento autónomo del estudiante.

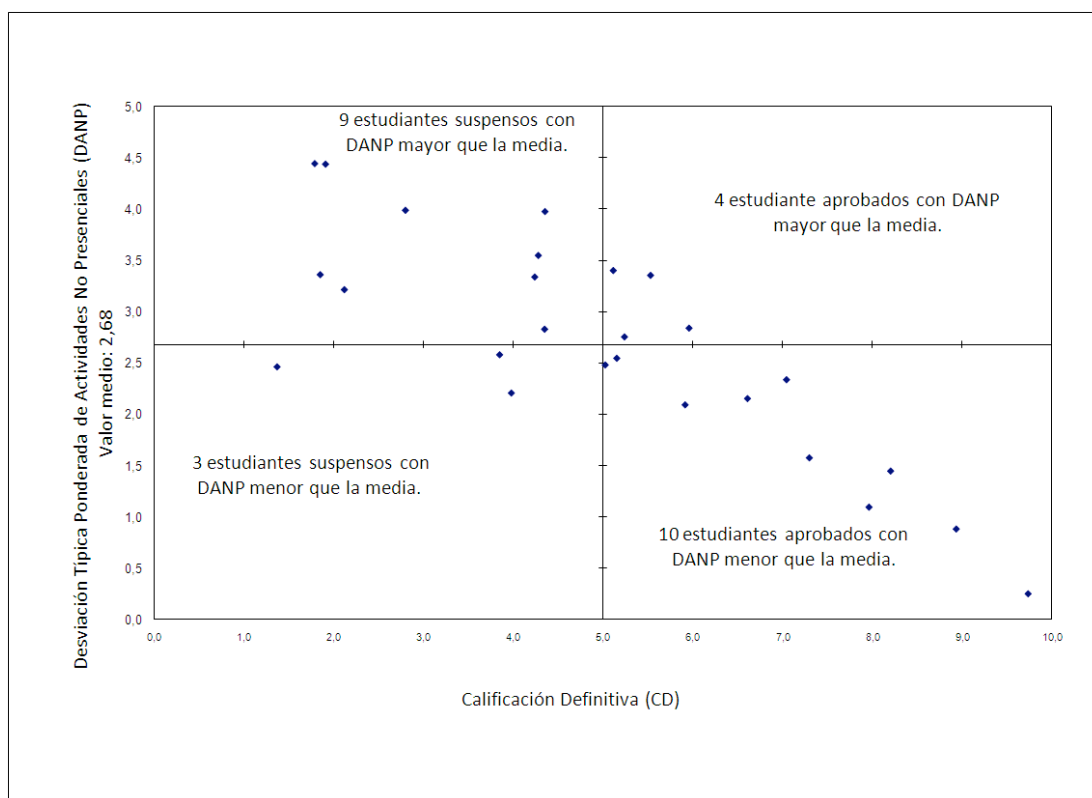
Sin embargo, si realizamos el estudio similar con la desviación típica de las actividades presenciales, observamos que los valores medios de dicha desviación en cada grupo de estudiantes son significativamente más bajos que en el caso de las actividades no presenciales. La explicación de esto se debe a que en este tipo de actividades se descartan alguno de los factores anteriormente citados, más concretamente el tercero, ya que estas actividades se realizan bajo la supervisión/vigilancia del profesor.

En definitiva, podemos afirmar que los estudiantes suspensos fueron más irregulares en el desarrollo de las actividades de evaluación, especialmente en las actividades no presenciales, y podemos intuir que ésta pudo ser la causa de su calificación definitiva, ya que no fueron constantes en su trabajo a lo largo del periodo lectivo. Para fundamentar un poco más esta última intuición presentamos el siguiente diagrama

dónde apreciamos más claramente la incidencia en la calificación definitiva de la irregularidad en las notas de las actividades no presenciales.

En el Gráfico 1 se observa claramente dos grupos significativos: los estudiantes suspensos con dicha desviación superior al valor medio y los estudiantes aprobados con desviación menor que el valor medio. Incluso de los 7 estudiantes que no quedan encuadrados en estos dos grupos, 6 de ellos están próximos a alguno de los grupos. Nuevamente, reforzamos la propuesta del índice DANP como un indicador del rendimiento autónomo y del rendimiento global del estudiante.

GRÁFICO 1: Distribución de DNAP en función de la nota cualitativa (aprobado/suspenso)



Como tercer y último bloque de conclusiones, planteamos una posible relación entre el rendimiento de los estudiantes y su dedicación. Para ello incorporamos a la discusión

otras mediciones del trabajo del estudiante obtenidos por una metodología distinta a la anterior, a saber, la pregunta directa a través de encuestas. No podemos garantizar una fiabilidad máxima en estas encuestas, ya que no tenían preguntas de control de la deseabilidad social. De los resultados obtenidos, podemos concluir que entre un 31% y un 41% de los estudiantes no dedicaban el tiempo previsto para su trabajo autónomo en la materia, una vez superada la mitad del curso.

En particular, si recordamos el Gráfico 6, podemos colegir que esencialmente el 41% de la tercera encuesta, es decir, los 11 estudiantes que no cumple ni por tiempo ni por ejercicios resueltos (parte de trabajo autónomo) con las previsiones realizadas en la guía docente, son los mismos que finalmente no superaron la materia y que en su mayor parte coinciden con aquellos con valor de DANP mayor que el valor medio.

Direcciones para correspondencia:

- Francisco José Plaza Martín. Departamento de Matemáticas. Plaza de la Merced, 1.

37008 - Salamanca

- [fplaza@usal.es](mailto:fplaza@usal.es)

Fecha de recepción de la versión definitiva: 09/03/2010

## NOTAS

[1] Proyecto titulado “*Implantación de nuevas Metodologías de Aprendizaje y Evolución en el primer curso de la Licenciatura de Matemáticas*” realizado en el curso 2005-06 y financiado por el MEC (orden CI/924/2005).

## BIBLIOGRAFÍA

CABRERA, A. F. (2003) *Evaluación de la formación* (Madrid, Síntesis).

COLL, C., ROCHERA VILLACH, M. J., MAYORDOMO SAÍZ, R. M. y NARANJO, M. (2007) Evaluación continua y ayuda al aprendizaje. Análisis de una experiencia de innovación en educación superior con apoyo de las TIC, *Electronic Journal Of Research in Educational Psychology*, 5: 3, 783-804.

DELGADO, A. M. Y OLIVIER, R. (2006) *La evaluación continua en un nuevo escenario docente*. Ver [http://www.uoc.edu/rusc/3/1/dt/esp/delgado\\_oliver.pdf](http://www.uoc.edu/rusc/3/1/dt/esp/delgado_oliver.pdf) (consultado el 21.III.2009).

DELGADO, A. M. Y OLIVIER, R. (2009) Interacción entre la evaluación continua y la autoevaluación formativa: la potenciación del aprendizaje autónomo. Ver [http://www.um.es/ead/Red\\_U/4/delgado.pdf](http://www.um.es/ead/Red_U/4/delgado.pdf) (consultado el 21.III.2009).

GONZÁLEZ-MARTÍN, C. y RAMOS ÁLVAREZ, M. P. (2007). Aplicación de un sistema de evaluación continua a alumnos de 5º de Farmacia, en el proceso de adaptación al Espacio Europeo de Educación Superior, *Revista d'educació superior en Farmàcia*, 1:1, 1-10.

LÓPEZ, M. (2001) *La evaluación del aprendizaje en el aula* (Madrid, Edelvives).

MARTÍNEZ MARTÍNEZ, M. y CADENATO MATÍA, M. (2010) Correlación entre la evaluación continua y el rendimiento académico. Ver <http://upcommons.upc.edu/e-prints/bitstream/2117/6334/1/Madrid%20092104enviado.pdf> (consultado el 12.I.2010)

## **Resumen**

### **Reflexiones en torno a una experiencia de Evaluación Continua en el EEES**

En el proceso de adaptación al Espacio Europeo de Educación Superior, se plantean numerosos cambios en las metodologías docentes. Así, se ha planteado la evaluación continua como un buen sistema para valorar la adquisición de competencias, al tiempo que este sistema habitualmente es enriquecido y complementado con características de sistemas de evaluación sumativa, formativa, etc.

En la adopción de un nuevo sistema de evaluación, lógicamente los docentes se plantean diversas cuestiones como la metaevaluación del sistema y, en este caso concreto, la eventual determinación de indicadores de rendimiento o dedicación individual y grupal.

En este artículo abordamos, partiendo de una experiencia concreta en la Universidad de Salamanca, distintos resultados y reflexiones en torno a estas dos cuestiones.

**DESCRIPTORES:** Sistemas de evaluación, Espacio Europeo de Enseñanza Superior, European Credit Transfer System, Evaluación Continua, Metaevaluación.



## **Abstract**

### **Critical reflections about an experience of continuous assessment according to ESHE**

The process of adaptation to the European Space for Higher Education requires important changes in teaching methods. Thus continuous assessment is a valuable approach to understanding a student's thinking and skills; nevertheless this approach can be complemented with other assessment systems as formative or summative evaluation.

In the adoption of new assessment systems, different matters must be taken into account such as evaluating assessment activities (i.e., meta-assessment), and setting indexes of academic performance and progress, both individual and group nature.

This study arises from a pilot scheme carried out in the University of Salamanca: it attempts to reflect on these topics.

**KEY WORDS:** Assessment systems; European Space for Higher Education; European Credit Transfer System, Continuous assessment, Meta-assessment.