



Estudio de la influencia del profesorado en la participación de los alumnos en las actividades de autoevaluación

Carlos Guerrero, Antoni Jaume-i-Capó

Departament de Ciències Matemàtiques i Informàtica

Universitat de les Illes Balears

carlos.guerrero@uib.es, antoni.jaume@uib.es

Resumen

En este trabajo se estudia la eficiencia de los métodos de motivación de los profesores para la participación de los alumnos en actividades de autoevaluación. Para ello se han propuesto actividades a distintos grupos de alumnos, y se les ha facilitado la posibilidad de autoevaluarlas. Los distintos grupos han sido motivados de distintas formas para que realizaran la autoevaluación. De los datos extraídos se ha estudiado la relación entre el porcentaje de alumnos que realizan la autoevaluación y, en primer lugar, el método utilizado para motivar a los alumnos, y en segundo lugar, el momento del curso en el que se ha realizado la actividad. Los resultados muestran que los alumnos incentivados con un incremento de la calificación final de la asignatura llevan a cabo con más frecuencia las actividades de autoevaluación y que a medida que el período de exámenes se acerca, el porcentaje de entregas disminuye.

Palabras clave: Autoevaluación, rúbricas, motivación alumnado, EEES, estudio experimental.

Recibido: 29 de noviembre de 2012; **Aceptado:** 18 de julio de 2013.

1. Introducción

El aprendizaje continuo y la retroalimentación de las actividades realizadas por los alumnos tiene un papel fundamental dentro de los nuevos planes de estudio y del nuevo espacio europeo de educación superior. Esto supone que el profesorado ha de dedicar más tiempo a la redacción y corrección de actividades, trabajos y ejercicios.

Se ha demostrado que la autoevaluación es una técnica beneficiosa para los alumnos y que puede sustituir perfectamente al trabajo de corrección del profesor incluso, en algunos casos, llegando a ser preferible [5]. Existen muchos estudios sobre los beneficios de la autoevaluación, pero es habitual dudar de cuáles son los mejores métodos para motivar al alumno a que la realice. En un gran número de estudios sobre educación se establecen métodos, tácticas de aprendizaje, etc, pero la mayoría carecen de una fase sólida de validación y sus procesos de validación suelen estar basados en la propia observación y experiencia del docente.

Las aportaciones más importantes de este estudio son: En primer lugar la definición de un método estadístico para llevar a cabo, de una forma más rigurosa, la validación de los resultados de participación en actividades. En segundo lugar un caso de estudio en el que se ha aplicado este método. En este caso de estudio se ha llevado a cabo un análisis de cómo influye la actitud del profesor en la autoevaluación del propio

alumno.

Para el caso de estudio que hemos analizado, se han tenido en cuenta tres formas con las que el profesor ha intentado motivar al alumnado de cara a autoevaluarse. Se han seleccionado distintos grupos para que los profesores motivaran de forma distinta al alumnado y los datos recolectados se han validado mediante el estudio de la diferencia significativa entre distintos grupos de experimentos. Para ello se ha realizado un contraste de hipótesis mediante el uso del test de Ji-cuadrado (χ^2) y el test de coeficiente de contingencia de Pearson. El test de χ^2 nos permite determinar si existe una relación entre dos variables aleatorias. El coeficiente de contingencia de Pearson expresa la intensidad de la relación entre dos o más variables. A partir de nuestro estudio se ha encontrado que el modo de incentivar la actividad influye en el porcentaje de alumnos que realizan autoevaluaciones. También influye el momento del curso en el que se realiza la actividad: a medida que se acercan los exámenes, el porcentaje de entregas disminuye.

2. Estado del arte

Durante los últimos años la coevaluación y la autoevaluación están recibiendo un interés creciente, seguramente al verse el profesorado obligado a trabajar en el aula competencias genéricas como el pensamiento crítico y la capacidad de aná-

lisis. Pueden parecer nuevas formas de evaluación, pero ya entre 1774 y 1826, el profesor George Jardine de la Universidad de Glasgow las incluía en su plan pedagógico [6].

Al delegar en los alumnos toda o parte de la evaluación se obtienen mejoras en su proceso de aprendizaje, con el correspondiente aumento del trabajo para el profesorado. Gracias a las herramientas en línea existentes, podemos reducir un poco este trabajo [3, 12, 14], sin olvidar otras herramientas más específicas para los estudios de informática [7, 15].

La característica más interesante de este tipo de evaluaciones es la retroalimentación, al convertir el alumnado en un elemento activo en el proceso y tener que corregir, confirmar y hacer sugerencias sobre el trabajo que evalúa. Además se añade un factor formativo en el proceso de evaluación, mucho más importante que el aditivo, al ser consciente el alumnado de sus puntos fuertes y sus puntos débiles [3, 13]. Por otra parte, está demostrado que el alumnado mejora en la gestión de su aprendizaje, dedicándose a profundizar sobre sus nuevos conocimientos, y le permite trabajar, entre otras, las competencias de reflexión y generalización de nuevas situaciones, trabajo en grupo, autoaprendizaje y gestión del tiempo [12].

Un inconveniente que se suele poner a la autoevaluación es la fiabilidad, lo que sorprende un poco, pues al trabajar en investigación el profesorado está habituado a la revisión entre iguales (coevaluación) y la revisión de sus propios trabajos (autoevaluación). De hecho si no fuese así, nunca los enviaríamos a congresos y revistas, y no daríamos por buenos los resultados de la revisión recibida. Aun así, muchos artículos publicados recientemente comparan la correlación de este tipo de evaluaciones con la evaluación del profesor [4, 8, 9, 10, 11]. Para que la autoevaluación y coevaluación sean fiables es muy importante conocer qué se quiere evaluar, y dar a los alumnos los conocimientos y las herramientas necesarias para que puedan hacerlo correctamente.

Delegar la evaluación en los alumnos, para obtener un nivel aceptable de fiabilidad, implica que el profesorado debe definir con gran claridad y precisión los criterios de evaluación, de la misma forma que los editores lo hacen con los revisores [2, 5, 10, 13]. Eso, entre otras cosas, requiere mucha más preparación, sobre todo de un modelo de soluciones detallado [3], y el consiguiente seguimiento durante el proceso de evaluación [13]. La experiencia que tengan nuestros alumnos en este tipo de evaluaciones, en los instrumentos y criterios de evaluación, y la naturaleza de la materia son variables muy importantes que afectan la fiabilidad y validez. Según Sithiworachart y Joy [12], los aspectos más importantes para asegurar el éxito son:

- Explicar por qué se ha decidido aplicar la autoevaluación o coevaluación a todas las partes interesadas;
- Involucrar a todos los participantes y explicar los criterios de evaluación de forma detallada;
- Facilitar la comunicación entre los participantes;

- Proporcionar un entrenamiento de calidad, con ejemplos y prácticas;
- Proporcionar directrices y formas de comprobación;
- Especificar claramente la planificación y el calendario de actividades;
- Realizar un seguimiento del proceso y con el consiguiente apoyo;
- Evaluar las evaluaciones y retroalimentar al alumnado para que mejore.

Analizando estos aspectos, se ve que la fiabilidad de este tipo de evaluaciones depende en gran medida de la formación, la motivación, las herramientas y el soporte que se ofrezca al alumnado.

2.1. Rúbricas

Una rúbrica, o matriz de valoración, es una herramienta de calificación formada por un conjunto de criterios ligados a los objetivos de aprendizaje usados para evaluar la actuación de los alumnos en la realización de determinadas tareas que pretendamos evaluar.

La rúbrica permite detallar los criterios específicos que se tendrán en cuenta a la hora de evaluar al alumno, en definitiva, hacer más transparente y fácil el proceso de evaluación.

Algunas de sus características más destacadas son:

- Facilita la evaluación al profesor;
- Define lo que se espera de un alumno en la realización de un trabajo;
- Reduce la sensación de ambigüedad que pueda tener un alumno cuando es evaluado;
- Permite ordenar los conocimientos según un juicio de valor que se establece y, por tanto, garantiza una evaluación más equitativa.

En el proceso de elaboración de una rúbrica se realizan tareas muy diversas. Entre ellas es importante buscar las características que definen el trabajo que se ha de realizar y los contenidos sobre los que los alumnos han de evaluar su aprendizaje. Igualmente, deben quedar claramente detallados los objetivos, desempeños, competencias y comportamientos que han de demostrar los alumnos.

Finalmente, se han de establecer unos niveles de consecución de esos objetivos y competencias. Cada uno de estos niveles debe de estar claramente detallado para que el alumnado sepa establecer el nivel al que ha conseguido llegar. Toda esta información acabará mostrándose en forma de matriz donde las filas son los aspectos a evaluar y las columnas el nivel conseguido. En la Figura 1 se puede ver un ejemplo de este tipo de matriz.

El proceso de realización de rúbricas es un trabajo cíclico, ya que una vez probadas las matrices con los alumnos, se deben revisar los criterios seleccionados para la graduación de

		Necesita mejorar	Satisfactorio	Bien	Excelente
1	Sabe buscar, seleccionar y organizar la información más relevante.	No es capaz de distinguir ni seleccionar la información más relevante.	Reconoce algunos puntos relevantes pero confunde información importante con datos irrelevantes.	La información es relevante, pero no está organizada.	La información seleccionada es la más relevante. La presenta de forma muy clara y organizada
2	Es capaz de sintetizar la información y escribir correctamente textos guiados.	Hay faltas ortográficas y el texto está escrito sin coherencia.	El texto tiene cierta coherencia, pero hay muchas faltas ortográficas.	El texto es coherente y hay pocas faltas ortográficas.	El texto es coherente y no hay faltas ortográficas
3	Sabe publicar información en el blog.	Ha publicado información en el blog de manera confusa.	El blog tiene una presentación útil, pero hay un exceso de información.	El blog es visualmente atractivo y tiene una presentación útil.	El blog es visualmente atractivo, tiene una presentación muy útil y la información está bien organizada

Figura 1: Ejemplo de rúbrica.

los conocimientos. En cada uno de estos ciclos de revisión-prueba se irán refinando y mejorando las matrices de la rúbrica.

Las rúbricas son el método considerado como el más adecuado para la realización de las actividades de autoevaluación. Existen otros métodos (como por ejemplo la elaboración de tests), pero no encajan tan bien, de forma general, en las actividades de este tipo ni, de forma particular, en el contexto de las asignaturas y trabajos que se han llevado a cabo en nuestros experimentos.

3. Método

En este apartado proponemos un método estadístico para validar los resultados de participación del alumnado en cualquier tipo de actividad. Esta validación es aplicable en aquellos casos en los que los objetivos de un estudio son conocer la respuesta del alumnado, ante distintos factores, en la realización de algún tipo de actividad.

A lo largo de la bibliografía de estudios de docencia, la validación de los resultados viene habitualmente dada por la propia opinión del profesorado. Este tipo de validación no es objetiva ni reproducible, y por lo tanto de poco valor. Por ello, en este artículo queremos definir una forma de proceder a la hora de validar este tipo de estudios.

Proponemos llevar a cabo la validación de los resultados mediante el estudio de la diferencia significativa entre distintos grupos de experimentos. Para ello se elegirán las variables de interés, se recogerán datos estadísticos adecuados y se hará un contraste de hipótesis mediante el uso de los tests adecua-

dos al caso. De esta forma podremos ver si dos variables son estadísticamente independientes o si hay una relación (no necesariamente causal) entre ellas.

El primer paso es elegir la variable de interés que se desea estudiar. Una vez elegida, se debe distribuir todos los grupos experimentales en bloques. Estos bloques vendrán definidos por los distintos posibles casos o valores de la variable de estudio. Una vez establecida esta clasificación en bloques, podremos recoger los datos pertinentes para calcular la métrica deseada sobre cada uno de los bloques y también sobre cada uno de los grupos experimentales de forma individual. De esta forma se puede ver cómo evoluciona en función de las condiciones de la variable a estudio. Finalmente, estos resultados han de ser validados mediante estudios estadísticos para determinar si se pueden considerar representativos.

Para el proceso de validación será necesario estudiar dos aspectos fundamentales:

- La existencia de diferencias significativas en el comportamiento de los alumnos pertenecientes a distintos bloques. La hipótesis en estudio determina que si los grupos experimentales de distintos bloques se ven sometidos a valores distintos de una misma variable independiente. Si es así, estos deben responder de distinta forma. Si responden de la misma forma, significará que la variable independiente no influye en los resultados obtenidos.
- La no existencia de diferencias significativas en el comportamiento de alumnos pertenecientes a distintos grupos experimentales de un mismo bloque. Los grupos experimentales del mismo bloque se ven sometidos a una

misma variable independiente y, dentro de variaciones aleatorias esperadas, deben responder de la misma forma. Si responden de forma distinta, significará que no se puede asegurar que la variable independiente es la que produce los cambios observados.

Dicho de otra forma, se estudia la independencia estadística de los grupos. Esta independencia debe de existir entre bloques pero no entre grupos experimentales de un mismo bloque. Para ello se realizarán los tests estadísticos adecuados. Un caso habitual en docencia es tener variables cualitativas (o factores) que pueden tomar valores discretos: cada alumno realiza o no una tarea, o indica si le gusta “poco”, “medio” o “mucho” algún aspecto del curso. En estos casos es adecuado realizar dos test: el de χ^2 y de coeficiente de contingencia de Pearson [1]. Mediante estos dos métodos se puede comprobar si las variables estudiadas son estadísticamente independientes. El test de χ^2 nos ofrece la probabilidad de concurrencia, y el test de coeficiente de contingencia de Pearson normaliza los resultados entre 0 y 1. Con esos valores se puede comprobar la asociación o no de las variables.

Para el caso de χ^2 tenemos que el valor límite, generalmente aceptado, para considerar significativo o no un resultado es $p = 0,05$. Cuando el p-valor obtenido es menor que este valor, se admite que existen diferencias entre las variables estudiadas y, por tanto, la asociación es significativa. Para el caso del test de coeficiente de contingencia, el valor que marca la asociación se encuentra en 0,33 y se consideran que las variables son diferentes cuando el valor es menor [1].

Una manera adecuada de mostrar los resultados es en tablas donde se indican, para cada uno de los bloques, el valor resultante del test de χ^2 , el valor numérico corresponde a la probabilidad con la que se acepta que la correlación entre las variables es significativa y, por último, el coeficiente de contingencia de Pearson. Así queda de forma claramente legible el cálculo y el resultado obtenido.

Veamos a continuación la aplicación de este método al caso de estudio del presente artículo: la influencia del método de motivación en la actividad de autoevaluación del alumnado.

El objetivo del estudio es realizar un análisis completo para cada una de las variables de las que se desee estudiar la influencia sobre la participación de los alumnos. Por lo que el primer paso será seleccionar una variable que pueda influir en la participación que en este caso ha sido la motivación recibida mediante el profesor. Se han escogido tres métodos: motivación única, motivación periódica y motivación incentivada. Los bloques a diferenciar, o grupos diferentes a crear, han sido 6: dos para cada motivación. Así se ha podido estudiar si existen diferencias entre grupos con una misma motivación y entre motivaciones.

Los experimentos llevados a cabo son cada actividad de autoevaluación realizada por un grupo de alumnos. La métrica utilizada es la respuesta, por parte de los alumnos, a distintos escenarios o actividades de autoevaluación planteados. Esta respuesta se mide por el número de alumnos que realizan la autoevaluación. En particular se ha medido el número de

trabajos entregados por cada grupo experimental y el número de autoevaluaciones llevadas a cabo. Hay que remarcar que este estudio no está enfocado a analizar si el proceso de autoevaluación de los alumnos se lleva a cabo de forma correcta. Tampoco está dentro de los objetivos del estudio el estudiar si existe alguna correlación entre las notas de los alumnos y el hecho de haber realizado las autoevaluaciones. Por ello en este estudio no se ha recogido ni la calificación en cada actividad ni la calificación final del alumno, sino sólo el número de autoevaluaciones realizadas.

Esta ha sido la estructura del experimento. En la sección siguiente se detallan las condiciones seguidas.

4. Experimento

Para comprobar la viabilidad del tipo de método de validación propuesto en el apartado anterior, vamos a proponer su utilización en un estudio sobre la participación del alumnado en actividades de autoevaluación. Dicho de otra forma, ver cuáles son los factores que influyen en que los alumnos hagan o no actividades de autoevaluación.

El primer paso a llevar a cabo es definir las condiciones del experimento. La autoevaluación de los alumnos se realiza con el uso de rúbricas. Para minimizar la influencia de otros factores sobre la variable de estudio, el porcentaje de participación de los alumnos, se ha creado una rúbrica única y común para todos los grupos que formarán parte del estudio. De esta forma, se consiguen dos objetivos, que el tipo de actividad sea del mismo tipo y que la rúbrica sea siempre la misma. De esta forma se evita que la heterogeneidad sobre estos dos aspectos cree un aumento de variabilidad que enmascaren diferencias existentes en los resultados finales.

El tipo de actividad sobre el que se lleva a cabo la autoevaluación es la redacción de informes técnicos de prácticas. Las rúbricas se han centrado en evaluar aspectos comunes de la redacción de dichos informes, como por ejemplo la presencia de introducción, de referencias bibliográficas, la exposición clara de los objetivos del trabajo, etc. Los aspectos concretos de cada una de las actividades no se han recogido en la rúbrica, ya que esto supone crear una rúbrica distinta para cada grupo y actividad, lo que rompe con la intención de homogeneizar los experimentos. En el Cuadro 1 se muestra la rúbrica.

La rúbrica creada se facilita al profesorado que forma parte del estudio para que sea distribuida entre sus alumnos y estos se autoevalúen en una o más actividades llevadas a cabo durante el curso y en las que se deben de redactar algún tipo de informe técnico. Posteriormente se solicita al profesorado información sobre cada uno de los experimentos. La información que se solicita es la siguiente:

- Número de alumnos matriculados en la asignatura.
- Número de alumnos que han entregado la actividad a autoevaluar.
- Número de alumnos que han llevado a cabo la autoevaluación.

	1	2	3	4
<i>Presentación del trabajo</i>	Presenta título sin relación al tema a tratar. Falta: Tema, nombre de autores, curso y carrera.	Presenta título que informa y concuerda con el tema a tratar. Falta: Nombre de autores, curso y/o carrera.	Presenta título que informa y concuerda con el tema. Adjunta nombre de autores sin los apellidos y falta curso y/o carrera.	Presenta título que informa y concuerda con el tema. Adjunta nombre y apellidos de autores, curso y carrera.
<i>Organización</i>	El formato de la presentación y la organización del material son confusos.	La información está bien organizada.	La información está bien organizada y se ajusta a los parámetros de presentación: Tamaño y tipo de letra	La información está bien organizada y se ajusta a los parámetros de presentación: Tamaño y tipo de letra así como respeta el interlineado y los márgenes.
<i>Introducción</i>	La introducción está débilmente redactada y no ofrece información relacionada al tema, o no presenta introducción.	La introducción plantea el tema, pero no argumenta el motivo de su elección ni ofrece una presentación de cómo se tratará el tema.	La introducción plantea el tema y argumenta la elección del mismo. No ofrece una presentación de cómo se tratará el tema.	La introducción plantea el tema, argumenta la elección del mismo y presenta cómo se desarrollará el tema.
<i>Desarrollo</i>	Las ideas no se entienden y muestra ausencia de uso de los tecnicismos propios de la temática a tratar.	Las ideas son difíciles de entender ya que no se exponen claramente o se usan vagamente los tecnicismos propios de la temática.	Las ideas se entienden pero no se hace un uso correcto de los tecnicismos propios de la temática, ya sea por defecto o exceso del uso de los mismos.	Las ideas son claras, se respetan los signos de puntuación y se hace un buen uso de los tecnicismos propios de la temática.
<i>Conclusión</i>	Inexistencia de conclusiones o este apartado se confunde con el resto del tema abordado sin llegar a ser claro.	Las conclusiones están débilmente presentes, es difícil para el lector identificar la opinión personal del autor del tema abordado.	Las conclusiones dan muestra de estudio del tema, son elaboradas de forma personal pero muestran débil congruencia al tema abordado.	Las conclusiones dan muestra de apropiación del tema, son elaboradas de forma personal y en absoluta congruencia al tema abordado.
<i>Ortografía</i>	Hay varios errores (más de 10) gramaticales en la presentación y desarrollo del documento.	Hay varios errores (entre 5 y 10) gramaticales en la presentación y desarrollo del documento.	Hay varios errores (menos de 5) gramaticales en la presentación y desarrollo del documento.	No hay errores gramaticales en la presentación y desarrollo del documento.
<i>Claridad</i>	Cada sección carece de una estructura definida de presentación, tratamiento y conclusión del tema.	La mayor parte de las secciones carecen de, al menos, dos de estas estructuras: Presentación, tratamiento y conclusión del tema.	Casi todas las secciones tienen una presentación, tratamiento y conclusión del tema.	Todas las secciones están bien estructuradas y presentan una presentación, tratamiento y conclusión del tema.
<i>Contenido</i>	Menos de la mitad del trabajo presenta información no correcta, no contrastada y de fuentes poco fiables.	La información del trabajo es adecuada pero no está contrastada o no se exponen las fuentes consultadas.	La mayor parte de la información presentada es correcta y contrastada, a pesar de que algunas fuentes pueden ser poco fiables.	Toda la información presentada es correcta, contrastada y de fuentes fiables y variadas.
<i>Esfuerzo</i>	El trabajo final demuestra que los estudiantes no pusieron ningún esfuerzo.	El trabajo final demuestra que con un poco de esfuerzo, el resultado final, podría ser bastante mejor.	El trabajo final demuestra que los estudiantes no alcanzaron a poner todo su esfuerzo.	El trabajo final demuestra que los estudiantes se esforzaron al máximo
<i>Presentación de referencias bibliográficas</i>	Las referencias bibliográficas no se citan en ningún momento del trabajo	La presentación de las referencias bibliográficas se presenta al final del documento sin respetar la normativa.	La presentación de las referencias bibliográficas se presentan al final del documento respetando la normativa.	La presentación de las referencias bibliográficas se presenta a lo largo del documento así como al final, respetando la normativa.

Cuadro 1: Rúbrica de evaluación para informes técnicos de grado de informática.

- Número de actividades de autoevaluación llevadas a cabo en dicha asignatura anteriormente.
- Fecha en las que se llevaron a cabo la actividad y la autoevaluación.
- Curso en la que se imparte la asignatura.
- Estudios en los que se imparte la asignatura.
- Número de profesores que dan clase en dicha asignatura.
- Utilización de técnicas de aprendizaje activo.
- Forma en la que se ha motivado a los alumnos.

De los puntos anteriores, el único que necesita establecer una homogeneidad entre los distintos grupos de experimentación es el último. Para ello se han definido tres formas diferenciadas de motivar al alumno. Se han redactado unas guías en las que el profesor tendrá los detalles de cómo llevar a cabo la tarea de motivación de los alumnos.

La primera forma de motivar a los alumnos, a la que llamaremos *motivación única*, consiste en comentar una única vez durante las clases la importancia, justificación y bondades de la autoevaluación. Se explicarán todos los aspectos importantes comentados previamente en la Sección 2. Esto se hará únicamente en una ocasión, a la hora de presentar la actividad. Tras esta primera sesión, no se volverá a recordar ni a animar al alumnado a realizar la autoevaluación.

Un segundo tipo de motivación, *motivación periódica*, está basada en los mismos términos que la anterior, pero se irá recordando de forma periódica, entre el momento de presentar la actividad y la fecha de entrega de la misma, la importancia de los procesos de autoevaluación y las ventajas que los alumnos pueden recibir de estos procesos.

Finalmente se utilizará una *motivación incentivada*, en la que el hecho de realizar la actividad de autoevaluación se ve recompensada con un porcentaje de la nota final. Para estos casos, también se solicitará al profesorado el valor de este porcentaje de nota bonificada al realizar la autoevaluación.

En los experimentos han participado cinco profesores del Departamento de Matemáticas e Informática de la Universitat de les Illes Balears. Estos profesores han llevado a cabo las tareas de autoevaluación en seis grupos distintos de alumnos, todos ellos enmarcados en estudios técnicos, principalmente en la rama de informática (cinco de los seis grupos). El tamaño de los grupos es bastante variable, cubriendo casos de grupos muy reducidos (9 alumnos) hasta grupos de un número considerable (casi 50 alumnos). El total de alumnos que se han visto involucrados en los experimentos ha sido de 227. En el Cuadro 2 se puede ver más detalladamente los grupos que han intervenido.

Todos los experimentos se han llevado a cabo durante el segundo cuatrimestre del curso 2010/2011. La autoevaluación se ha llevado a cabo sobre actividades que ya estaban

contempladas y planificadas anteriormente en el desarrollo de las asignaturas, añadiendo la realización de la rúbrica como una actividad adicional. Los métodos que el profesorado ha utilizado para motivar a los alumnos son los definidos anteriormente.

5. Resultados y Discusión

El análisis de los resultados se ha realizado dividiendo los 6 grupos de alumnos estudiados en tres bloques. Los grupos que forman parte de cada bloque se han determinado por el tipo de motivación llevada a cabo (única, periódica o incentivada). Cada bloque cuenta con dos grupos experimentales, o grupos de alumnos, llamados B y C¹, para controlar que las diferencias de resultados se deban únicamente al método de incentivación utilizado. Los grupos vendrán identificados por el mismo nombre utilizado para el tipo de motivación. Por lo que tendremos el bloque de motivación única (grupos B-único y C-único), el bloque de motivación periódica (grupos B-periódico y C-periódico) y el bloque de motivación incentivada (grupos B-incentivado y C-incentivado).

Para realizar el análisis de los datos se tuvieron en cuenta el número de trabajos entregados por cada grupo experimental y el número de autoevaluaciones llevadas a cabo. En el Cuadro 3 se pueden ver los resultados obtenidos por la métrica. Se ha realizado un análisis de la información mediante la comparación de los porcentajes de los trabajos entregados por grupos experimentales y bloques. El Cuadro 4 muestra información, resumiendo el porcentaje de entregas para los distintos grupos experimentales, por un lado, y para los bloques, por otro.

Estudiemos primero si hay diferencias entre los dos grupos de cada uno de los tres bloques. Realizamos los test de χ^2 y el coeficiente de contingencia de Pearson. El Cuadro 5 muestra los resultados. La primera columna de esta tabla nos muestra, para cada uno de los bloques, el valor resultante del test de χ^2 . El segundo valor numérico corresponde a la probabilidad con la que se acepta que la correlación entre las variables es significativa, y por último, se muestra el coeficiente de contingencia de Pearson.

Únicamente en el caso del bloque incentivado se tiene un número suficiente de valores experimentales para que estos test se puedan realizar de forma fiable. El elevado p-valor ($p = 0,82 > 0,05$) y el bajo valor del coeficiente de contingencia ($c = 0,012 < 0,33$) indica que los resultados son homogéneos y por lo tanto las diferencias encontradas con otros bloques son útiles para el estudio.

Desgraciadamente, para los otros dos bloques el bajo valor de la muestra impide que los test sean fiables. Mostramos los datos por completitud.

Si estudiamos la diferencia entre los bloques obtenemos unos valores que parece indicar que diferencias entre los distintos bloques parece que puede ser explicada por el tipo de incentivo que han recibido los alumnos: $\chi^2 = 106,03$; $p <$

¹No se utilizan las vocales a la hora de dar nombre a los grupos y por eso no existe un grupo A

Id. grupo	Núm. alumnos	Curso	Estudios	Tipo asignatura	Id. profesor
B-único	19	3	Informática	Obligatoria	W
B-periódico	9	4	Informática	Troncal	W
B-incentivado	43	5	Informática	Optativa	W
C-único	23	3	Informática y Industriales	Optativa	X
C-periódico	22	3	Informática	Obligatoria	Y
C-incentivado	111	2	Informática	Troncal	Z

Cuadro 2: Grupos en los que se ha llevado a cabo el experimento.

Grupo experimental	Cantidad	Bloque único	Bloque periódico	Bloque incentivado
B	Entregas	10	9	39
B	Autoevaluaciones	4	8	31
C	Entregas	15	12	333
C	Autoevaluaciones	0	0	289

Cuadro 3: Número de trabajos entregados y autoevaluaciones realizadas en cada uno de los grupos experimentales.

0,01; $c < 0,001$. De esta forma quedaría claro que los alumnos, por sí solos, no realizarían las autoevaluaciones y que éstas sólo las llevarían a cabo cuando se les reitera su importancia o cuando se les premia por su realización. Esto dejaría de manifiesto que los alumnos no tienen interiorizados los beneficios de la autoevaluación, y que necesitan de una motivación externa para llevarla a cabo. De esta forma, la primera de nuestras hipótesis estaría comprobada.

Aunque el pequeño tamaño de la muestra impide obtener test fiables, es interesante estudiar por qué parece que no hay homogeneidad entre grupos. La primera explicación posible para este hecho podría venir dada por los profesores que han realizado los experimentos y el tipo de actividades que llevan a cabo a lo largo de la asignatura. Parece lógico pensar que en asignaturas donde se lleva a cabo una evaluación constante de las actividades y trabajos del alumno, estos están más acostumbrados a este tipo de actividades y están más acostumbrados a una docencia basada en el aprendizaje. Este hecho podría afectar a la tasa de autoevaluaciones realizadas. Los grupos experimentales B han sido evaluados por el mismo profesor, y todos ellos muestran unas tasas de autoevaluación mayores. En cualquier caso, la segunda de nuestras hipótesis no quedaría validada por los resultados.

Una segunda posibilidad es que los periodos en los que se han llevado a cabo la entrega de las autoevaluaciones influyan en su tasa de éxito. La recogida de las autoevaluaciones se ha llevado a cabo en diferentes fechas durante el segundo cuatrimestre del año, y no han coincidido entre ellas. De ahí viene la justificación del estudio realizado para contrastar nuestra tercera hipótesis.

Los datos del grupo experimental C-incentivado corresponden a la suma de tres actividades de autoevaluación sobre el mismo grupo de alumnos. De ahí los números tan eleva-

dos que se muestran (y que no corresponden con el número de alumnos de la asignatura). Para comprobar la validez de esta última hipótesis vamos a estudiar la evolución de la participación a lo largo de las tres autoevaluaciones realizadas. El porcentaje de participación de los alumnos fue bajando al irse acercando el final de curso. En la primera entrega participó un 92,8 %, en la segunda un 86,5 % y en la última un 81,1 %. El tamaño del grupo C permite estudiar si estas diferencias son estadísticamente significativas. Realizamos un test de χ^2 y de contingencia y los resultados obtenidos son $\chi^2 = 6,652$ para un p-valor de $0,036 < 0,05$ y un coeficiente de contingencia de $0,038 < 0,33$.

De nuevo existen diferencias en la entrega de las autoevaluaciones por parte de un mismo grupo, por lo que se puede decir que la asociación en la entrega de las autoevaluaciones por parte del mismo grupo es significativa. Los resultados parecen mostrar que a medida que el período de exámenes se acerca, la tasa de éxito de entregas disminuye. Esto podría estar asociado, de nuevo, a que los alumnos no tienen interiorizados los beneficios de la autoevaluación, y se dedican a afrontar los exámenes de una forma más tradicional basada en el estudio. De esta forma, nuestra hipótesis sobre la existencia de una relación directa entre la época en que está programada la actividad de autoevaluación y la proximidad de los exámenes quedaría contrastada.

6. Conclusiones

En este artículo se muestran los resultados de un estudio sobre la influencia de los métodos de motivación utilizados por los profesores sobre las tasas de éxito de participación de los alumnos. Para ello se han propuesto actividades a distin-

Bloque único		Bloque periódico		Bloque incentivado	
Grupo B	Grupo C	Grupo B	Grupo C	Grupo B	Grupo C
40 %	0 %	88,9 %	0 %	97,5 %	86,8 %
16 %		38,1 %		87,1 %	

Cuadro 4: Porcentajes de autoevaluaciones entregadas.

Bloque	χ^2	Correlación significativa	Coefficiente de contingencia
Único	2,85 (p = 0,091)	p > 0,05	0,513
Periódico	5,62 (p = 0,02)	p < 0,05	0,642
Incentivado	3,863 (p = 0,82)	p > 0,05	0,012

Cuadro 5: Resultados de los test χ^2 y de coeficiente de contingencia de Pearson para los bloques.

tos grupos de alumnos, y se les ha facilitado la posibilidad de autoevaluarlas. Los distintos grupos han sido motivados de distintas formas para que realizaran la autoevaluación. En algunos casos esta motivación ha sido mínima y única en el tiempo, en otros se ha realizado de forma periódica, y en otros grupos se ha incentivado añadiendo esta actividad a las evaluables dentro de la asignatura.

De los datos extraídos se han estudiado fundamentalmente dos aspectos: la relación entre el método utilizado para motivar a los alumnos, y el éxito de la actividad de autoevaluación; y la influencia del período en el que se han realizado las actividades sobre la tasa de éxito de la misma.

Las conclusiones extraídas son que los alumnos incentivados con un incremento de la calificación final de la asignatura llevan a cabo con más frecuencia las actividades de autoevaluación. También se muestra más probables en aquellos grupos a los que se les motiva periódicamente, pero con menor eficacia que en el caso de la utilización de incentivos. Del mismo modo se ha visto que la proximidad del período de exámenes influye negativamente en la realización de actividades de autoevaluación.

Agradecimientos

Este trabajo ha sido financiado parcialmente por el Departamento de Matemáticas e Informática de la Universitat de les Illes Balears y por el proyecto "Estudi d'autoevaluació i coevaluació de l'EEES" financiado por la convocatoria de "Projectes d'innovació i millora de la qualitat docent 2010-2011 y 2011-2012" del Vicerrectorado de Profesorado e Innovación Docente de la Universitat de les Illes Balears. Igualmente agradecer el trabajo de elaboración de rúbricas y recolección de datos a los profesores Isaac Lera y Carmen Erice.

Referencias

[1] Lluís Ballester: *Bases metodológicas de la investigación educativa*, 2ª edición. Edicions Universitat de les Illes Ba-

lears. España. 2004.

- [2] C. Bauer, K. Figl, M. Derntl, P.P. Beran y S. Kabicher: *The student view on online peer reviews*. Actas de 14th Annual ACM SIGCSE Conference on innovation and Technology in Computer Science Education, ITiCSE '09, pp. 26 – 30. Paris, Julio de 2009.
- [3] David J. Boud y W. Harvey Holmes: *Self and Peer Marking in an Undergraduate Engineering Course*. IEEE Transactions on Education, vol. 24, núm. 4, pp. 267 – 274, noviembre de 1981.
- [4] Nancy Falchikov y Judy Goldfinch: *Student Peer Assessment in Higher Education: A Meta-Analysis Comparing Peer and Teacher Marks*. Review of Educational Research vol. 70, núm. 3, pp. 287–322, septiembre de 2000.
- [5] S. Fallows, B. Chandramohan: *Multiple Approaches to Assessment: reflections on use of tutor, peer and self-assessment*, Teaching in Higher Education, vol. 6, núm. 2, pp. 229–246, 2001.
- [6] L.I. Gaillet: *A foreshadowing of modern theories and practices of collaborative learning: The work of the Scottish rhetorician George Jardine*. Ponencia presentada en 43rd Annual Meeting of the Conference on College Composition and Communication, Cincinnati (EE.UU), marzo de 1992.
- [7] E.F. Gehringer: *Electronic peer review and peer grading in computer-science courses*. SIGCSE Bulletin, vol. 33, núm. 1, pp. 139 – 143, marzo de 2001.
- [8] E.J. Mirielli: *Using peer-evaluation in a website design course*. Journal of Computing in Small Colleges vol. 22, núm. 4, pp. 14 –21, abril de 2007.
- [9] Javier Oliver Bernal y Verónica Canivell Castillo: *Evaluación entre compañeros: estudio de su correlación con la evaluación del profesor*. Actas de las Jornadas de Enseñanza Universitaria de la Informática, JENUI 2009, pp. 241 – 246. Barcelona, julio de 2009.

[10] Miguel Riesco Albizu y Marián Díaz Fondón: *La revisión entre iguales como herramienta de aprendizaje y evaluación en la asignatura de sistemas operativos*. Actas de las Jornadas de Enseñanza Universitaria de la Informática, JENUI 2007, pp. 277 – 284. Teruel, julio de 2007.

[11] J. Sitthiworachart y M. Joy: *Effective peer assessment for learning computer programming*. SIGCSE Bulletin, vol. 36, núm. 3, pp. 122–126, septiembre de 2004.

[12] Jirarat Sitthiworachart y Mike Joy: *Self and Peer Assessment in School and University: Reliability, Validity and Utility*. En el libro de M. Segers, F. Dochy y E. Cascallar (editores): *Optimising New Modes of Assessment: In Search of Qualities and Standards*. pp. 55–87. Kluwer Academic Publishers, 2003.

[13] J. Sitthiworachart y M. Joy: *Web-based peer assessment in learning computer programming*, Actas de 3rd IEEE International Conference on Advanced Learning Technologies, ICALT 2003. pp. 180–184, Atenas, julio de 2003.

[14] H. Sondergaard: *Learning from and with peers: the different roles of student peer reviewing*. Actas de 14th Annual ACM SIGCSE Conference on innovation and Technology in Computer Science Education, ITiCSE '09, pp. 31 – 35. Paris, Julio de 2009.

[15] W.J. Wolfe: *Online student peer reviews*. Actas de 5th Conference on information Technology Education CITC 2004, octubre de 2004.



Carlos Guerrero obtuvo el grado de Doctor en Informática en la Universitat de les Illes Balears en el año 2012. Actualmente ocupa una plaza de profesor colaborador en el área de Arquitectura y Tecnologías de los Computadores en el Departamento de Matemáticas e Informática de la misma universidad. Sus temas de investigación están relacionados con el rendimiento web, ingeniería web, aplicaciones web y móviles, minería de datos, innovación en la educación y utilización de las tecnologías de la información en la educación. Es autor de 12 artículos publicados en congresos internacionales y en revistas. Igualmente ha formado parte de un gran número de comités de programa y de revisión. Ha participado en proyectos de investigación nacionales e internacionales y ha liderado proyectos locales de innovación docente durante los últimos cinco años.



Antoni Jaume-i-Capó obtuvo el grado de Doctor en Informática en la Universitat de les Illes Balears en el año 2009. Actualmente ocupa una plaza de profesor contratado doctor en el área de Ciencias de la Computación e Inteligencia Artificial en el Departamento de Matemáticas e Informática de la misma universidad. Sus temas de investigación están relacionados con la visión por computador, interfaces, serious games, rehabilitación, innovación en la educación y utilización de las tecnologías de la información en la educación. Es autor de 11 publicaciones en revistas y más de 15 aportaciones en congresos internacionales.

©2014 C. Guerrero, A. Jaume-i-Capó. Este artículo es de acceso libre, distribuido bajo los términos de la Licencia Creative Commons de Atribución, que permite copiar, distribuir y comunicar públicamente la obra en cualquier medio, sólido o electrónico, siempre que se acrediten a los autores y fuentes originales