

Relación entre el método de evaluación del trabajo y el nivel de aprendizaje de los estudiantes *

José I. Aliaga Estellés, Gregorio Quintana Ortí
Departamento de Ingeniería y Ciencia de los Computadores
Universitat Jaume I
12071 - Castellón de la Plana
{aliaga, gquintan}@uji.es

Resumen

El objetivo del presente trabajo es presentar dos métodos para evaluar el trabajo realizado por los estudiantes fuera del aula y comparar el nivel de aprendizaje adquirido en cada uno de ellos. El primero se fundamenta en la evaluación entre compañeros, mientras que el segundo combina la autoevaluación y la realización de una prueba objetiva. En ambos casos, el objetivo fundamental es aportar una rápida retroalimentación a los alumnos. La comparación de las calificaciones de los estudiantes permite concluir que el uso de pruebas objetivas mejora el nivel de aprendizaje de los alumnos. La segunda opción ha requerido el desarrollo de una herramienta informática que evalúa las respuestas de los estudiantes a la vez que detecta posibles problemas en los enunciados de las pruebas objetivas.

Abstract

The main goal of this paper is to present two methods to evaluate the students' homework, and to compare their learning level when these methods are used. The first one is based on peer-assessment, while the second one includes a self-assessment and a test. In both cases, the main objective is to provide a fast feedback to the students. Analyzing the students' grades, we conclude that the use of tests improves the learning level of the students. The second method has required the development of an application which computes the assessment of the students and, at the same time, detects any problem in the formulation of the objective tests.

Palabras clave

Evaluación entre compañeros, autoevaluación, corrección y validación simultánea de pruebas objetivas.

*Financiado por la Unitat de Suport Educatiu de la Universitat Jaume I en el marco de un PIE (curso 2013/14).

1. Motivación

La reforma de los planes de estudios, asociada a la construcción del Espacio Europeo de Educación Superior (EEES), ha propiciado una renovación meteorológica [3], en la que el estudiante pasa a ser el protagonista del proceso de enseñanza-aprendizaje. Como consecuencia, la planificación de las materias no sólo ha de especificar qué debe hacer un alumno sino que también debe indicar cómo debe hacerlo y los recursos que debe utilizar. En esta renovación, la evaluación pasa a ser la herramienta que utiliza el docente para dirigir al estudiante a través de todo el proceso de enseñanza-aprendizaje [4], y debe considerar todas las tareas que éste realice, tanto presenciales como no presenciales.

Desde el año 2011, un grupo de profesores de los departamentos de Informática de la Universitat Jaume I constituyeron el Seminario Permanente de Innovación Educativa (SPIE) "Eficacia Docente en Informática", con el objeto de debatir sobre todos estos aspectos. En este foro abierto los profesores presentan sus experiencias de autoevaluación y de evaluación por compañeros, debaten sobre su implementación y proponen posibles acciones de mejora [6]. Algunas de las conclusiones alcanzadas en este seminario aparecen reflejadas en este artículo.

A continuación se describe la estructura del resto del artículo. En primer lugar se presenta el contexto en el que se enmarca el trabajo, describiendo el nombre, competencias y resultados de la asignatura sobre la que se realiza la experiencia, antes de pasar a describir la metodología docente y el método de evaluación utilizado. A continuación se comparan las dos alternativas de evaluación aplicadas, mostrando los resultados obtenidos y un detallado análisis comparativo. Seguidamente se presenta la herramienta que ha facilitado la implementación de uno de estos métodos de evaluación, incidiendo en sus características más interesantes y comentando líneas de mejora. Finalmente se detallan las conclusiones del trabajo.

2. Contexto

Los autores de este trabajo son profesores de la asignatura *Programación Concurrente y Paralela (PCP)*. El principal objetivo de esta asignatura es que los alumnos adquieran las competencias que les permitan desarrollar aplicaciones que aprovechen al máximo las características de las nuevas arquitecturas. Esta asignatura es una materia Obligatoria en los Grados de Ingeniería en Informática (II) y de Matemática Computacional (MC), titulaciones ambas pertenecientes a la Escuela Superior de Tecnología y Ciencias Experimentales (ESTCE) de la Universitat Jaume I (UJI).

Su docencia se realiza durante el primer semestre pero el curso en el que se enmarca difiere en cada una de las titulaciones, ya que aparece en el tercer curso de II y en el cuarto curso de MC. La razón que justifica esta diferencia es que las competencias previas de la asignatura PCP han sido planificadas en cursos distintos de cada titulación. Dado que la docencia en los grados se inició en el curso académico 2010/11, el presente curso (2013/14) es el segundo año en el que PCP se imparte en II y el primer año que se imparte en MC.

La asignatura consta de 6 créditos con 60 horas presenciales y 90 horas no presenciales. Del total de horas presenciales, 22 horas se dedican a clases teóricas, 14 horas a clases de problemas, 20 horas a prácticas de laboratorio y 4 horas a la elaboración de un examen. Por su parte, las horas no presenciales se dividen en 75 horas de trabajo personal y 15 horas de trabajo de preparación de los exámenes.

2.1. Competencias

En los planes de estudio de estos grados, II y MC, la asignatura PCP sólo tiene asignada una competencia:

- Conocimiento y aplicación de los principios fundamentales y técnicas básicas de la programación paralela, concurrente, distribuida y de tiempo real.

Pero tiene asociada varios resultados de aprendizaje:

- Describir los aspectos fundamentales de la programación de sistemas en tiempo real.
- Diseñar e implementar algoritmos paralelos eficientes sobre memoria compartida.
- Diseñar e implementar algoritmos paralelos eficientes sobre memoria distribuida.
- Diseñar e implementar algoritmos que funcionen de forma concurrente.
- Explicar los principios y técnicas básicas de la programación concurrente.
- Resolver los problemas fundamentales de la programación concurrente y paralela.

Durante el diseño de los planes de estudios de ambas titulaciones, se acordó que Java fuese el lenguaje

de programación vehicular en ambas titulaciones, razón por la que ha sido el lenguaje empleado en PCP. También se introduce OpenMP para la programación de procesadores multinúcleos y multiprocesadores, así como la librería de paso de mensajes MPI para la programación de multicomputadores.

2.2. Metodología

La metodología docente varía según cada uno de los tipos de clases: teoría, problemas y laboratorio.

Los *contenidos teóricos* de la asignatura se han dividido en 11 sesiones de dos horas en las que el profesor combina la docencia expositiva e interactiva [1]. En primer lugar, el profesor presenta los conceptos básicos que son necesarios para alcanzar los resultados de aprendizaje asociados a cada tema. Durante esta exposición y tras cada concepto básico, se fomenta de modo expreso la participación activa del alumno mediante la resolución de tareas y ejercicios relacionados. Estos ejercicios son resueltos por los alumnos con la supervisión del docente, lo cual permite una inmediata retroalimentación [5]. El principal material que se utiliza en estas clases son las transparencias y los entregables de aula. Los primeros incorporan de un modo estructurado la información requerida por el alumno, mientras que los segundos contienen los ejercicios que se realizarán en el aula.

Cada sesión de teoría tiene asociada una *sesión de problemas* de una duración de una hora y cuarto, que se realiza una semana después, en la que se evalúa la labor realizada por los estudiantes fuera del aula. El trabajo requerido aparece en el entregable de casa que es asignado a los alumnos al finalizar cada sesión de teoría. Los entregables incluyen un conjunto de tareas y/o ejercicios cuya duración estimada varía entre las tres y cuatro horas, y que los estudiantes tienen una semana para completar.

A diferencia de las clases teóricas, en las que el profesor imparte docencia a todos los estudiantes, las clases de problemas se imparten a grupos más pequeños, reduciendo así la distancia entre el profesor y los estudiantes. Este hecho permite al profesor comentar las diferentes soluciones que los alumnos han realizado. De esta manera, los estudiantes pueden conocer tanto sus propios errores como los errores cometidos por sus compañeros, a la vez que examinan las diferentes formas de resolver un determinado problema. Todo ello redundará en un aumento de su nivel de aprendizaje [7].

Las diez *sesiones de prácticas* de dos horas en las que se descomponen los correspondientes créditos de la asignatura PCP, se realizan en aulas informáticas en las que todos los estudiantes tienen acceso a un ordenador. Obligatoria los alumnos deben realizar un trabajo previo en casa antes de realizar la práctica en el laboratorio. Al principio de la clase, el pro-

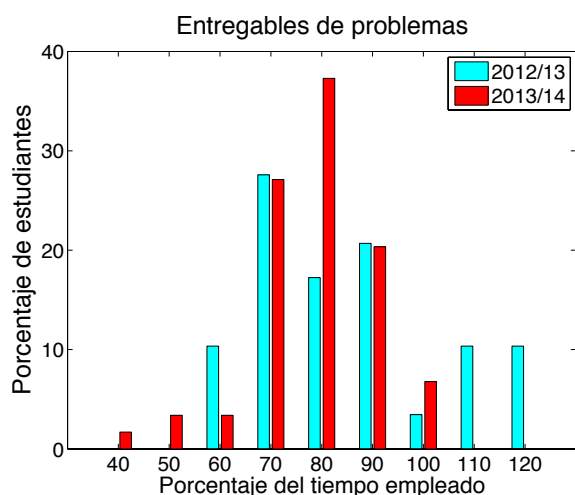


Figura 1: Histograma del porcentaje del tiempo utilizado por los alumnos para completar los entregables de problemas respecto del tiempo inicialmente estimado.

feesor comprueba la realización del trabajo previo, a la vez que explica y resuelve los posibles problemas que se hayan podido encontrar. Este trabajo autónomo previo fomenta la responsabilidad y el autoaprendizaje del alumno. Aunque existe cierta relación en el contenido de las prácticas, cada una de las sesiones es autocontenida de modo que al final de cada una los alumnos deben presentar sus resultados para ser evaluados.

Además, se proponen diferentes *trabajos adicionales* (como la lectura de libros en inglés) que complementan los contenidos de la asignatura. Estas tareas son voluntarias y sólo tienen impacto en la nota final si se ha aprobado la asignatura.

La metodología docente de los cursos 2012/13 y 2013/14 ha sido básicamente la misma, considerando los tres tipos de docencia y la dinámica de las clases. Tras impartir la asignatura por primera vez, se modificó el contenido del material de algunas sesiones de la asignatura para asegurar que la duración real de los entregables para casa se ajustase mejor a la duración planificada. Estas modificaciones fueron muy pocas en número y sólo afectaron a algunos temas puntuales.

El cálculo del tiempo asociado a los distintos entregables de problemas se calcula a partir de la información que los propios estudiantes deben anotar obligatoriamente en el propio entregable. Si se acumula toda esta información y se compara respecto del tiempo asignado para estas tareas en la planificación inicial, que es de 42 horas, se puede construir un gráfico como el que aparece en la Figura 1. En él se observa una reducción del porcentaje de tiempo medio empleado sobre el total planificado, que ha pasado del 85 % en el curso 2012/13 al 78 % del curso 2013/14, así como un descenso muy importante de los valores máximos, que han pasado del 118 % al 101 %. De este modo se verifica que el ajuste realizado ha sido adecuado.

2.3. Evaluación

En la evaluación de la asignatura se tienen en cuenta todos las tareas realizadas por los estudiantes a lo largo del semestre, con el fin de obtener una valoración más acorde al trabajo realizado y al nivel de aprendizaje adquirido. Seguidamente se comenta el peso que cada parte tiene así como sus requisitos:

- **Problemas (4 puntos):** Cada sesión de problemas se evalúa de modo independiente, y la media de las calificaciones obtenidas conforma el 75 % del peso de esta parte. Los entregables no realizados se evalúan a cero.

El restante 25 % se obtiene a partir de la media de dos pruebas escritas, una a mediados y otra al final de semestre, que permite a los estudiantes identificar el nivel de aprendizaje obtenido en los diferentes contenidos.

- **Laboratorio (3 puntos):** El trabajo realizado en cada una de las sesiones de laboratorio también se evalúa por separado y la nota asociada a esta parte se obtiene a partir de la media de las calificaciones obtenidas en cada sesión. Como en los problemas, se evalúan a cero los entregables no presentados.
- **Examen (3 puntos):** El examen escrito incluye tanto cuestiones cortas, que permiten evaluar el nivel de aprendizaje de los fundamentos, como ejercicios prácticos más extensos, que permiten evaluar la capacidad de combinar diversos conceptos sobre problemas más prácticos.
- **Trabajos Adicionales (puntos extra):** Los alumnos deben demostrar que la tarea se ha realizado de modo satisfactorio, para lo cual realizan una entrevista con un profesor en el que resumen su contenido y dan su opinión sobre el tema.

Para superar la asignatura es necesario que un estudiante obtenga al menos el 50 % de la calificación en las dos primeras partes y el 40 % de la calificación en la tercera parte, debiéndose cumplir además que la suma de todas las calificaciones sea mayor o igual a 5. Sólo en dicho caso se consideran y añaden a la calificación final los puntos obtenidos en el cuarto apartado.

3. Evaluación de los entregables de problemas

En el inicio del curso 2012/13 se decidió que el objetivo principal de las clases de problemas fuese la corrección y evaluación del trabajo realizado en casa por los alumnos, para que así dispusieran de una realimentación adecuada, temprana y frecuente sobre su trabajo. La siguiente cuestión fue decidir el método de evaluación a utilizar, optando por la *evaluación entre compañeros*, uno de los métodos presentados en el SPIE.

El análisis de la docencia en ese curso permitió detectar varios problemas, entre los que destacan la elevada carga de los profesores y la homogeneización sobre valores altos de las notas de problemas. El origen del primero fue la revisión semanal de los ejercicios y anotación de las notas, que en algunos casos llegó a ser superar los mil ejercicios por profesor. Por su parte, la homogeneización en valores tan altos de las notas hacían que el método no sirviera a los alumnos como predictor de sus resultados en el examen final.

Este último dato fue el que motivó el cambio del método de evaluación para el curso 2013/14, eligiendo una combinación de la *autoevaluación* y la *realización de pruebas objetivas*. El impacto de este cambio sobre las calificaciones de los alumnos ha sido muy interesante, poniendo de manifiesto alguno de los problemas que presenta el uso de estas técnicas.

3.1. Evaluación entre compañeros

Esta técnica permite cubrir, de un modo sencillo, los aspectos que se consideran más importantes:

- **Retroalimentación rápida:** Ésta es una de las herramientas fundamentales para aumentar el nivel de aprendizaje de los estudiantes, ya que les permite identificar sus errores y conocer la solución de un modo inmediato.
- **Estudio de varias soluciones:** La puesta en común de las respuestas de los alumnos permite conocer nuevas alternativas para resolver los ejercicios. Tanto si la respuesta del alumno es correcta como si no, el esfuerzo de analizar y comprender la posible validez de otras alternativas abre nuevas líneas de razonamiento muy beneficiosas para el alumno [2].

Siendo ciertas todas estas afirmaciones, la implementación fue un tanto más compleja.

La dinámica en el aula era muy sencilla. Al inicio de la clase el profesor recogía los entregables realizados por los estudiantes y los reasignaba de modo aleatorio. Tras presentar y comentar las posibles soluciones de cada uno de los apartados de los ejercicios, éstos se evaluaban utilizando valores enteros entre 0 y 10, lo que les permitía interiorizar los estándares de calidad [5]. Seguidamente el profesor recogía y supervisaba todas las calificaciones, para después calcular su valor medio e incorporarlo como nota final del entregable. Esta metodología dio lugar a diferentes problemas:

- **Trabajo del docente:** La supervisión e introducción de las calificaciones fue una tarea bastante ardua, ya que el número medio de notas por alumno y semana fue de 23, aunque hubo entregables con hasta 50 notas. Por ello, en algunas semanas se llegó a superar los mil ejercicios por

profesor. Además, algunos alumnos no calificaban algunas respuestas de sus compañeros, lo que alargó aún más el proceso.

- **Información compartida:** Aunque suele resultar muy positiva la colaboración entre alumnos en la realización del trabajo en casa, su uso indiscriminado puede tener un impacto muy negativo sobre los niveles de aprendizaje. La proliferación actual de herramientas electrónicas permite a los alumnos compartir rápidamente las soluciones sin un análisis previo ni puesta en común. La presencia de algunos errores triviales muy repetidos inducían a pensar que algunos alumnos tomaban directamente las soluciones de otros compañeros sin examinarlas críticamente, lo que anulaba todo su proceso de aprendizaje.
- **Método de evaluación:** La simplificación del método de calificación ayudó a integrar a los estudiantes en el proceso, pero ayudó a elevar la nota media de problemas, homogeneizándola sobre valores bastante altos.

Todas estas razones aconsejaron la revisión del método de evaluación de los entregables de problemas para el curso 2013/14.

3.2. Autoevaluación y pruebas objetivas

La nueva metodología debía mantener los aspectos positivos que se alcanzaban con la evaluación entre compañeros, pero reduciendo el impacto de los inconvenientes de su implementación. La solución por la que se optó fue separar el proceso en dos etapas: la *evaluación formativa* y la *evaluación calificativa* [7].

En la primera de las etapas, implementada mediante una *autoevaluación*, se mantenía el nivel de implicación de los estudiantes, que seguían comentando en el aula las diferentes soluciones, lo que permitió aumentar su nivel de aprendizaje. Pero la diferencia fundamental es que cada estudiante comprobaba sólo la validez de su propio trabajo pero no lo calificaba. No obstante, en la puesta en común los alumnos continuaban examinando las respuestas de sus compañeros.

Por su parte, la segunda de las etapas se realizó mediante una *prueba objetiva* relacionada con el contenido del entregable de casa. Esta prueba incidía en sus aspectos más representativos, dándoles el peso que se consideraba más adecuado en el conjunto de la evaluación. Además, la realización de esta prueba dio a los alumnos mayor responsabilidad sobre su nivel de aprendizaje, ya que eran conscientes de su valor de modo casi inmediato.

El procesamiento de los resultados de estas pruebas fue muy sencillo, reduciendo de forma muy significativa la carga de trabajo del profesorado. En ello influyó, en gran medida, el desarrollo de una herramienta de co-

Curso	Num	Problemas				Examen				Problemas-Examen			
		Max	Min	Med	DSt	Max	Min	Med	DSt	Max	Min	Med	DSt
2012/3	28	9,3	6,0	8,1	0,8	8,0	1,8	5,1	1,5	5,3	1,1	3,0	1,1
2013/4	51	9,0	5,1	6,8	2,6	8,7	2,3	5,6	2,6	4,5	-1,6	1,2	1,2

Cuadro 1: Comparación de las calificaciones de los entregables de problemas y del examen durante los cursos 2012/13 y 2013/14.

rrección y validación automática de pruebas objetivas, que se describe en detalle en la sección 4.

3.3. Comparación de las alternativas

El cambio en el método de evaluación de los entregables de problemas permitió resolver los problemas detectados: la carga de los profesores, la compartición sin aprendizaje y una evaluación poco ajustada al nivel de aprendizaje. Pero además reveló un impacto sobre la calificación de los estudiantes mayor del esperado.

Existen muchas alternativas para analizar e interpretar un conjunto de valores. El Cuadro 1 muestra los valores estadísticos más significativos (máximos, mínimos, medias y desviaciones estándares) de las calificaciones de los entregables de problemas, del examen, y la diferencia entre ambos valores, para los estudiantes que se presentaron a primera convocatoria de los cursos 2012/13 y 2013/14. Para asegurar un correcto análisis, cabe indicar que el nivel de dificultad del examen en ambos cursos fue similar. En el posterior análisis se ha excluido a los alumnos de MC debido a su pequeño número (5) y al reducido impacto que tiene sobre los valores contenidos en el Cuadro 1.

Si se comparan los resultados de las calificaciones de los *entregables de problemas* se observa un descenso acusado de la nota media, que pasa de 8,0 a 6,8. La interpretación que parece más plausible es que el nuevo método de evaluación permite individualizar el nivel de aprendizaje de los alumnos, ya que elimina tanto el impacto de compartir la solución de los entregables sin un aprendizaje real, como la problemática asociada a la implementación de la evaluación entre compañeros. Esta interpretación se podrá confirmar cuando se analicen los resultados en nuevos cursos académicos. En ningún caso se considera que los cambios en la asignatura haya tenido ningún impacto sobre esta diferencia, debido a su poco número y calado.

Por lo que respecta a la *calificación de los exámenes*, el Cuadro 1 muestra el efecto beneficioso de la nueva evaluación de los problemas sobre la nota del examen, que ha permitido aumentar en un 10% la calificación media de los alumnos, que pasa de 5,1 a 5,6. Este incremento se puede justificar por la mayor responsabilidad e involucración de los estudiantes en su aprendizaje, propiciando un mayor esfuerzo por su parte, y por tanto un aumento de su nota del examen. El incremento de

0,5 puntos se mantiene en el análisis de los valores mínimos, que pasa de 1,8 a 2,3, mientras que es de 0,7 en los valores máximos, que pasan del 8,0 a 8,7. Por tanto se puede concluir que lo que realmente se ha producido es un desplazamiento de la nota del examen.

El tercer grupo de columnas del Cuadro 1 muestra la *diferencia de las calificaciones* que aparecen en los dos bloques anteriores, definiendo la distancia entre la calificación de la evaluación continua y del examen escrito. Siendo lógico que haya una diferencia entre ambas calificaciones, los valores para el curso 2012/13 pueden entenderse como poco habituales, ya que la diferencia media es de 3 puntos a favor de la nota de problemas. Además el análisis del valor mínimo muestra que ningún alumno mejoró la calificación de problemas en su examen escrito. Sin embargo en el curso 2013/14 aparece una diferencia media de 1,2 puntos, un valor que parece más equilibrado, mientras que el valor mínimo pasa a ser de -1,6 puntos. Este último dato indica que hubo alumnos que obtuvieron una mejor calificación en su examen final que en los problemas, hecho que no tendría por qué ser excepcional, dado que algunos alumnos prefieren esforzarse para el examen final antes que involucrarse en la evaluación continua.

Las gráficas incluidas en la Figura 2 muestran, respectivamente, la distribución asociada a las diferencias de calificaciones de los cursos académicos estudiados, junto con un histograma que agrupa las calificaciones en grupos de 0,5 puntos. Se observa claramente cómo en 2013/14 las notas de problemas se acercan más a las de los exámenes que en 2012/2013

Por su parte, la Figura 3 muestra en un plano bidimensional la relación existente entre ambas calificaciones para todos los alumnos en los dos cursos estudiados. La relación ideal sería que la calificación del examen fuera muy parecida a la de problemas. Por ello, los puntos del gráfico deberían situarse cerca de la diagonal (línea negra). Puede verse fácilmente cómo los alumnos del curso 2013/14 se acercan más a la situación ideal que los del curso anterior. Para ambos casos se muestran las correspondientes rectas de regresión.

En resumen, se puede concluir que el cambio en la evaluación de los entregables ha permitido lo siguiente:

- Ajustar la nota de la evaluación continua al nivel de aprendizaje real del estudiante.
- Mejorar la nota media del examen final debido a una mayor involucración de los estudiantes.

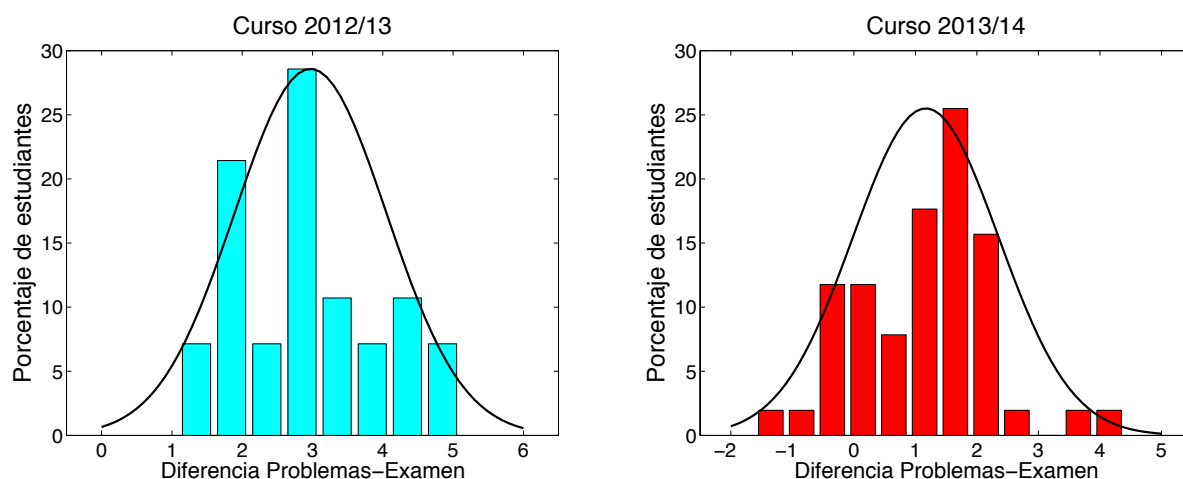


Figura 2: Histograma y campana de Gauss asociada a la diferencia entre las calificaciones del examen y de los entregables de problemas de los cursos 2012/13 y 2013/14.

4. Corrección y validación automática de pruebas objetivas

Otro cambio muy significativo en la docencia de la asignatura PCP entre los cursos 2012/13 y 2013/14 fue la herramienta utilizada para gestionar las calificaciones de los alumnos. Durante el primer curso se utilizó un conjunto de hojas de cálculo, agrupadas en un libro, en las que los profesores iban introduciendo las diferentes calificaciones obtenidas por los estudiantes. Posteriormente estos valores se iban propagando a través de las hojas para obtener la calificación de cada una de las partes de la asignatura. Como el procesamiento de los datos introducidos resultaba bastante complejo, al comienzo del curso 2013/14 se optó por utilizar un método alternativo: una base de datos. La introducción de los datos es muy similar en ambos casos, pero el análisis de los datos se puede realizar de un modo más completo con esta segunda alternativa, sin más que desarrollar un conjunto de consultas SQL que permitan obtener los datos deseados.

En esta sección se describe el módulo principal de la aplicación, el cual permite evaluar las respuestas de los alumnos a la vez que se analiza la validez de las preguntas que contiene la prueba.

4.1. Descripción básica del módulo

Las pruebas objetivas son una herramienta muy útil que pueden ayudar a reducir la carga del docente, pero a la vez puede ser muy injusta para los estudiantes en función del grado de dificultad de las pruebas y la ambigüedad de las opciones. El objetivo del módulo desarrollado no es el almacenamiento y generación de pruebas objetivas a partir de un banco de preguntas, sino la evaluación y análisis de pruebas objetivas. Este

módulo ha sido diseñado para ser utilizado en la calificación de pruebas de un número cualquiera de preguntas de respuesta única sobre 3 opciones, aunque se puede generalizar fácilmente para otros casos.

La primera tarea del docente es el desarrollo de la prueba y la introducción de la cadena que define la solución. Esta cadena está compuesta por un número de letras, una por pregunta (habitualmente unas 20). Cada carácter de la cadena define la respuesta correcta a una pregunta, y que para el caso de 3 opciones, se encuentra en el conjunto {'a', 'b', 'c'}. Para reducir el número de errores, se obliga a que el docente introduzca dos veces la cadena de soluciones para verificar que la introducción ha sido correcta, mostrándose una alerta si ambas cadenas no coinciden, si el tamaño de alguna de las cadenas no es igual al número de preguntas de la prueba, o si alguno de los caracteres no es una letra válida. También se impide el uso de la acción *Pegar* sobre estos campos, para forzar al usuario a que realice la introducción de la cadena dos veces.

Tras definir la cadena con las soluciones asociadas a una prueba, la siguiente tarea es la introducción de las respuestas de los alumnos. Como en el caso anterior, la cadena con las respuestas de un estudiante debe ser introducida dos veces, generándose una alerta si el tamaño de alguna de ellas no es correcto o si las dos cadenas no son iguales. De igual modo, también se ha inhabilitado la acción *Pegar* sobre estos campos. Además de las tres posibles respuestas {'a', 'b', 'c'}, existe otro valor válido, el carácter '.', que indica que el estudiante no ha respondido la pregunta correspondiente, ver Cuadro 2. Tras introducir todas las respuestas, el módulo puede generar un listado en el que se muestra para cada alumno el número de respuestas correctas, el número de errores y el número de respuestas en blanco, así como la calificación obtenida, ver Cuadro 3. Para el

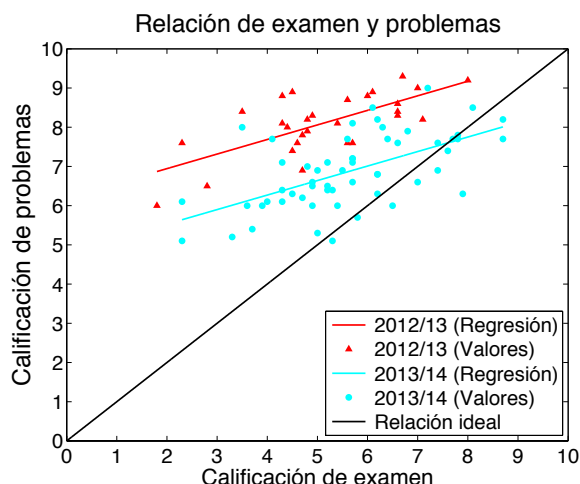


Figura 3: Relación entre la calificación del examen y la calificación de los entregables de problemas para los cursos 2012/13 y 2013/14.

DNI	horas	respuesta	respuesta2
XXXX	3:30	cabc.abb.c.cbbabaccc	cabc.abb.c.cbbabaccc
XXXX	3:15	cabbaabaacaabcabaccc	cabbaabaacaabcabaccc
XXXX	3:00	cabbacbbbcabbcbacccc	cabbacbbbcabbcbacccc
XXXX	3:00	cabba.bbbcaeb.ab.ccc	cabba.bbbcaeb.ab.ccc
XXXX	2:30	cabababbacacbbabaccc	cabababbacacbbabaccc
XXXX	2:00	cab...bb.ca.cbabc.cc	cab...bb.ca.cbabc.cc
XXXX	2:30	caba..bb.ca...ab...c	caba..bb.ca...ab...c
XXXX	2:00	cabcbabbccac.bab...cc	cabcbabbccac.bab...cc
XXXX	1:45	cabcaebbbcbcbcbacccc	cabcaebbbcbcbcbacccc
XXXX	2:00	cabbacba.cacbbabccccc	cabbacba.cacbbabccccc
XXXX	3:00	cabba.baaccb.bac...ca	cabba.baaccb.bac...ca

Cuadro 2: Ejemplo de introducción del tiempo requerido para completar un entregable y de las respuestas de una prueba objetiva.

caso de 20 preguntas y 3 opciones, la ecuación que calcula la calificación de un estudiante es la siguiente:

$$nota = [(aciertos - errores/2)/20] * 10.$$

4.2. Uso avanzado del módulo

Antes de generar un listado de notas definitivas, resulta conveniente analizar el número de aciertos y fallos de cada pregunta de la prueba. De este modo, se pueden detectar posibles redacciones incorrectas o ambiguas en una o más preguntas.

Este análisis se realiza a partir de un listado que incluye las estadísticas de las preguntas de una prueba, y que se muestra ordenado por el número de errores y/o respuesta en blanco que haya tenido cada pregunta. En el Cuadro 4 aparece el ejemplo de una prueba con unos resultados que requieren una revisión. Concretamente,

DNI	horas	nAcier	nErr	nBlan	Nota
XXXX	3:15	12	6	0	5,0
XXXX	2:00	14	2	2	7,2
XXXX	2:00	14	2	2	7,2
XXXX	2:00	14	3	1	6,9
XXXX	2:45	12	6	0	5,0
XXXX	3:00	10	5	3	4,2
XXXX	3:15	16	2	0	8,3
XXXX	2:30	14	3	1	6,9
XXXX	1:00	13	4	1	6,1
XXXX	1:30	17	1	0	9,2

Cuadro 3: Ejemplo del listado final de un entregable de problemas, incluyendo el tiempo de realización, los resultados de la prueba objetiva y la calificación.

las preguntas 3 y 13 tienen un número de aciertos muchísimo menor que el resto de preguntas, lo que invita a estudiar su causa.

Existen diferentes razones por las que las respuestas a una pregunta puede contener muchos más errores y blancos que el resto:

- Contenido mal asimilado.
- Pregunta mal formulada.
- Existencia de una o más opciones ambiguas.

La solución a un problema de este tipo en una pregunta puede conllevar una o varias de las siguientes acciones:

- A largo plazo (para el curso siguiente), modificar los contenidos del material de la asignatura, tanto del tema como de los entregables, para asegurar un nivel de aprendizaje adecuado.
- A corto plazo, anular la pregunta analizada, de modo que las respuestas no contabilicen en el cálculo de la calificación de los estudiantes.
- A corto plazo, permitir más de una respuesta válida a la pregunta analizada.

La aplicación ha sido diseñada específicamente para permitir realizar alguna de estas dos últimas opciones de forma rápida y automatizada. Para ello, tan sólo hay que modificar la cadena con la solución como sigue:

- El carácter '.' anula la correspondiente pregunta.
- La letra 'x' equivale a validar todas las opciones salvo 'a', es decir, 'b' y 'c' se consideran válidas.
- La letra 'y' equivale a validar todas las opciones salvo 'b', es decir, 'a' y 'c' se consideran válidas.
- La letra 'z' equivale a validar todas las opciones salvo 'c', es decir, 'a' y 'b' se consideran válidas.

La utilización de esta herramienta ha permitido detectar 8 cuestiones mal diseñadas de un total de 236, es decir, el 3,4 % del total. De estas cuestiones problemáticas, 4 se han resuelto invalidando la pregunta mientras que en las otras 4 se ha permitido varias opciones

nPreg	nResp	nAcier	nErr	nBlan
3	57	9	40	8
13	57	17	22	18
17	57	31	25	1
20	57	35	6	16
15	57	37	16	4
5	57	38	12	7
12	57	39	13	5
4	57	41	7	9
9	57	43	14	0
18	57	44	8	5
19	57	45	11	1
2	57	46	9	2
7	57	48	5	4
11	57	48	3	6
8	57	50	5	2
1	57	50	4	3
16	57	54	1	2
14	57	55	2	0
6	57	57	0	0
10	57	57	0	0

Cuadro 4: Ejemplo de listado de procesamiento de las respuestas de una prueba objetiva.

como válidas.

El uso de las letras { 'x', 'y', 'z' } no es generalizable para el caso de respuesta única sobre más de 3 opciones o de respuesta múltiple. En ambos casos se aconseja la utilización de un mapa de bits para identificar qué respuestas son válidas asignando un 1 en la posición correspondiente. De este modo, una pregunta se podría considerar anulada si el mapa de bits asociado fuese igual a 0.

5. Conclusiones

En este trabajo se ha analizado el impacto de un cambio en la evaluación de la labor realizada por los estudiantes fuera del aula sobre la calificación de este trabajo y sobre los resultados del examen final. Para ello, en primer lugar, se ha descrito la implementación de dos metodologías utilizadas al impartir una misma asignatura en dos cursos académicos distintos, indicando sus características más interesantes y presentando los resultados que han obtenido los alumnos en cada curso. Del análisis de todos estos datos se puede concluir que las técnicas de evaluación entre compañeros resultan de gran utilidad para mejorar la retroalimentación, pero pueden dar calificaciones que no se ajusten al nivel de aprendizaje de los alumnos, lo que puede llevar a un posterior fracaso en el examen final. La separación de la evaluación formativa y de la evaluación calificativa, mediante la inclusión de pruebas objetivas, da al alumno una idea más correcta de su ni-

vel de aprendizaje y una mayor responsabilidad sobre su aprendizaje continuo, lo que le permite corregir sus posibles carencias con la suficiente antelación.

También se ha presentado una aplicación en la que se evalúan las respuestas de los estudiantes a la vez que se validan los enunciados de las preguntas, permitiendo resolver de un modo rápido y sencillo sus posibles deficiencias. De este modo se agiliza todo el proceso de corrección, análisis y revisión de las preguntas y de las respuestas de los alumnos. Aún cuando la aplicación presentada está limitada a preguntas de respuesta única sobre 3 opciones, su generalización resulta muy sencilla, por lo que su planteamiento puede ser válido en muchos otros ámbitos.

Referencias

- [1] Maria Aymerich Andreu y Maria Eugènia Gras Pérez. Las metodologías docentes y su valoración por parte de los estudiantes universitarios. *Actas del UNIVEST'09*, Girona, 2009.
- [2] María Cavas, Francisco Chicano, Francisco Luna y Luis Molina. Autoevaluación y Evaluación entre Iguales en una Asignatura de Redes de Ordenadores. *Actas del Actas de JENUI'11*, ISBN: 978-84-615-8119-1, 321 - 328, Sevilla, 2011.
- [3] Mario De Miguel. Metodologías para optimizar el aprendizaje. Segundo objetivo del Espacio europeo de educación superior. *Revista Interuniversitaria de Formación del Profesorado*, 20(3), 71-91, 2006.
- [4] Grahah Gibbs. Uso estratégico de la evaluación en el aprendizaje Grahah Gibbs, , en *Evaluar en la universidad: problemas y nuevos enfoques (2ª edición)*, 61 - 74, Narcea Ediciones, ISBN: 978-84-277-1422-9, 2007.
- [5] Graham Gibbs y Claire Simpson. Condiciones para una evaluación continuada favorecedora del aprendizaje. *Cuadernos de Docencia Universitaria*, num. 13, ICE y Ediciones OCTAEDRO, S.L, ISBN: 978-84-8063-934-7, 2009.
- [6] Mercedes Marqués Andrés, José M. Badía Contelles y Ester Martínez-Martín. Una experiencia de autoevaluación y evaluación por compañeros. *Actas del JENUI'13*, ISBN: 978-84-695-8051-6, 93 - 100, Castellón, Julio 2013. DOI: 10.6035/e-TiIT.2013.13.
- [7] Miguel Valero-García y Luis M. Díaz de Cerio. Autoevaluación y co-evaluación: estrategias para facilitar la evaluación continuada. *Actas del SINDI'05 (AENUI)*, ISBN: 84-9732-443-9, 25-32, 2005.