

Actas Simposio-Taller JENUI 2012, Ciudad Real, 10-13 de julio 2012
I.S.B.N. 10: 84-695-3941-8 | I.S.B.N. 13: 978-84-695-3941-5
Páginas 1-8

Aplicando Evaluación por Pares: Análisis y Comparativa de distintas Técnicas

Carlos Blanco y Pablo Sánchez
Dpto. de Matemáticas, Estadística y Computación
Universidad de Cantabria
Facultad de Ciencias, Avda. Los Castros S/N
39071 Santander
{carlos.blanco, p.sanchez}@unican.es

Resumen

Las actividades docentes han de diseñarse para favorecer que los alumnos adquieran las habilidades y competencias necesarias para el desarrollo futuro de su profesión. Una de estas habilidades es la capacidad de evaluar de forma crítica y objetiva la calidad de proyectos o documentos técnicos, y para adquirirla, los alumnos han de practicar durante la carrera. Para ello, sería interesante involucrar a los alumnos en el proceso de evaluación de los trabajos de una determinada asignatura, algo que tradicionalmente ha sido realizado unilateralmente por el docente. Sin embargo, hay que tener en cuenta que, por naturaleza, los alumnos tienden a calificarse de forma poco objetiva favoreciendo o perjudicando a ciertos compañeros de forma deliberada. Por lo tanto, en este artículo se proponen y analizan, mediante experimentos, variantes de aplicación de la técnica de corrección entre pares buscando aquella que maximice la objetividad de las calificaciones resultantes.

Summary

Teaching activities have to be designed in order to improve the acquisition of competences by students, that is, the skills that they will need in their professional career. One of them is the capability for assessing a piece of professional work objectively. To acquire this capability, students have to be trained for assessing works as part of their degrees. To achieve this goal, it would be helpful to involve students in the assessment process of certain subjects, a process that has been traditionally carried out by teachers exclusively. Nevertheless, students tend to assess their mates taking into subjective factors such as personal relationships. Moreover, they use to rate high marks, always above pass. This paper pro-

poses several assessment techniques and evaluates them empirically by means of experiments in order to analyse their effectiveness regarding objectivity.

Palabras clave

Evaluación por pares, Técnicas de evaluación, Análisis empírico, Investigación en educación.

1. Introducción

El marco Europeo de Educación Superior intenta favorecer que los alumnos desarrollen las habilidades o competencias necesarias para el desarrollo futuro de su profesión [2, 6, 5]. Una de estas habilidades es la capacidad de evaluar de forma crítica y objetiva la calidad de proyectos relacionados con su profesión, pero para adquirir esta competencia es necesario que el alumno durante su formación tome el papel de evaluador y practique evaluando varios proyectos [7]. Esta tarea puede realizarse de forma individual, en la que el alumno desarrollará su capacidad de crítica y objetividad, aunque también puede realizarse organizando a los alumnos en grupos de evaluación que hayan de tomar una decisión unificada, por lo que los alumnos habrán de debatir, expresar y defender sus opiniones y llegar a un consenso como grupo, trabajando así competencias adicionales.

Una buena forma para lograr que los estudiantes adquieran estas competencias es involucrándolos en el proceso de evaluación de una determinada asignatura, de forma que se propongan varios trabajos técnicos que los alumnos desarrollarán y posteriormente los propios alumnos se encargarán de evaluar [3]. Todos ellos tendrán los conocimientos técnicos básicos para desempeñar esta labor, ya que están siguiendo la asignatura y han realizado un trabajo de temática y dificultad similar.

Sin embargo, el proceso de evaluación ha sido una

labor tradicionalmente realizada de forma unilateral por el docente, y al involucrar a los alumnos en la evaluación hay que tener en cuenta que por naturaleza tienden a realizar estas calificaciones de forma poco objetiva, para favorecer o perjudicar a ciertos compañeros de forma deliberada [4, 8, 1].

Aun así, creemos que las ventajas obtenidas al incorporar a los alumnos en el proceso de evaluación compensan los inconvenientes, aunque habría que intentar minimizarlos lo máximo posible, intentando que los alumnos realmente otorguen calificaciones objetivas.

Para ello, en este artículo además de la clásica técnica de evaluación por pares en la que cada estudiante evalúa a los demás mediante una nota numérica de 0 a 10, hemos propuesto y analizado dos variantes adicionales que intentan maximizar la objetividad obtenida.

Este artículo está organizado de forma que en primer lugar se describen las diversas variantes que hemos considerado para implicar a los alumnos en el proceso de evaluación (Sección 2). Posteriormente, se presenta el diseño de varios experimentos que tratan de evaluar cada una de las técnicas propuestas (Sección 3). Las secciones siguientes (Sección 4, 5 y 6) muestran los resultados obtenidos al aplicar cada experimento y razonan sobre el grado de objetividad alcanzado por cada técnica analizada. Finalmente, a modo de conclusión, se realiza un análisis final de las técnicas analizadas y los resultados obtenidos (Sección 7).

2. Técnicas de Evaluación

Como se ha comentado anteriormente, la participación de los alumnos en el proceso de evaluación de una determinada asignatura asumiendo el papel de evaluadores, es un buen método que permite que mediante la práctica adquieran las competencias necesarias para poder evaluar de forma objetiva y crítica documentos técnicos relacionados con su profesión.

Sin embargo, por experiencias previas hemos comprobado como la participación de los alumnos en el proceso de evaluación, tiene asociados varios problemas relacionados con la objetividad de las calificaciones asignadas por ellos. Principalmente, evalúan guiados no sólo por motivos de la calidad del trabajo analizado, sino por otros factores como la relación personal con sus compañeros, para favore-

cer o perjudicar a ciertas personas deliberadamente. Además, puntúan con notas altas y son reticentes a asignar una calificación por debajo del 5. Por ejemplo, aunque exista un conjunto de trabajos a evaluar en el que claramente se diferencien uno o varios muy buenos y trabajos muy mejorables, las puntuaciones asignadas por los estudiantes suelen estar todas alrededor del notable y no necesariamente los trabajos mejores consiguen la mejor puntuación.

Dado que consideramos que otorgar a los alumnos el papel de evaluadores proporciona la práctica necesaria para que adquieran las competencias que deseamos, hemos analizado distintas variantes de aplicación con objeto de buscar aquella que mitigue en mayor medida los problemas asociados. Describimos dichas técnicas a continuación.

2.1. Evaluación Numérica Individual

Es la forma clásica de involucrar los alumnos en el proceso de evaluación. Los alumnos, de forma individual, califican los trabajos de cada uno de sus compañeros mediante una nota numérica de 0 a 10. A priori, esta técnica tendrá asociados los problemas comentados ya que al actuar cada alumno de forma individual, tenderá a favorecer o perjudicar a quien él decida.

2.2. Evaluación Numérica por Grupos

Al igual que la técnica anterior, esta técnica califica los trabajos numéricamente de 0 a 10 pero organiza a los alumnos en grupos, tanto para realizar los trabajos como para evaluar los del resto de compañeros. Al incluir el concepto de grupo evaluador, la evaluación ya no es una tarea individual, sino que cada grupo habrá de debatir y llegar a un consenso sobre la calificación a asignar. Con esta medida se pretende mejorar la objetividad de las calificaciones.

2.3. Evaluación por Ordenación por Grupos

Esta técnica continúa con la organización de los alumnos en grupos de trabajo y de evaluación, pero sin embargo, modifica la forma de calificar. En este caso, en vez de asignar calificaciones numéricas de 0 a 10 puntos, cada grupo ha de ordenar los trabajos del resto de grupos de mejor a peor. Con este cambio se intenta mitigar el problema de que los es-

tudiantes siempre asignan calificaciones numéricas altas y abriendo el abanico de calificaciones.

3. Evaluación Empírica

Considerábamos importante analizar cada una de las técnicas anteriores de forma empírica, con el objetivo de buscar aquella que mitigue en mayor medida los problemas comentados y otorgue unas calificaciones más objetivas y cercanas a las que asignaría el docente. Por tanto, realizamos una serie de experimentos para analizar la objetividad de cada una de estas técnicas. En esta sección describimos como se realizaron dichos experimentos.

Los experimentos se han realizado en la asignatura de Ingeniería del Software I de cuarto curso de Ingeniería Informática, de la Universidad de Cantabria. Han participado 21 alumnos del curso académico 2010-2011, en el que se utilizó la técnica de evaluación numérica individual, 28 alumnos del curso 2010-2011 que aplicaron la técnica de evaluación numérica por grupos y 28 alumnos del curso académico 2011-2012, donde se analizó la técnica de evaluación por ordenación por grupos.

En cada experimento disponemos de un conjunto de trabajos realizados por los alumnos. Por cada experimento, los alumnos califican el conjunto de trabajos (a excepción del suyo propio). Para ello se les proporciona una plantilla con los aspectos de los trabajos que deben evaluar. La evaluación se realiza de forma no anónima en todos los experimentos, intentando así que las calificaciones sean más objetivas. Por otro lado, el docente también evalúa los trabajos.

Una vez finalizado cada experimento procedimos a analizar los resultados obtenidos. Por cada trabajo evaluado disponíamos del conjunto de calificaciones realizadas por el resto de estudiantes. A estas calificaciones le calculamos la media y desviación, y asignamos la nota media como calificación del trabajo. Era interesante comprobar que el valor de la desviación de estas medias no fuese demasiado elevado, ya que una alta desviación sería señal de subjetividad.

Por otro lado, disponíamos de las calificaciones que el docente había asignado a cada trabajo. A continuación, calculamos la media y desviación de las calificaciones finales del total de los trabajos, tanto de las calificaciones realizadas por el profesor, como

de las realizadas por los estudiantes. Si los valores para las calificaciones de la media y desviación para las calificaciones del profesor difiriese mucho de las de las calificaciones otorgadas por los alumnos, se entendería que pertenecen a poblaciones de datos distintas, lo cual sería síntoma de cierta subjetividad (asumiendo que el docente es objetivo).

Finalmente, calculamos la diferencia media (y desviación) existente entre las calificaciones del profesor y de los estudiantes por cada trabajo. Si la media de estas diferencias (o la desviación) resultase muy alta, sería también síntoma de subjetividad.

A continuación se detalla cada experimento realizado, mostrando y analizando los resultados obtenidos.

4. Técnica de Evaluación Numérica Individual

Este experimento analiza la técnica de evaluación numérica individual, en la que cada alumno puntúa de forma individual el trabajo del resto de alumnos, utilizando una nota numérica de 0 a 10 puntos. Los trabajos a evaluar fueron realizados individualmente y consistían en una memoria y presentación ante la clase de un patrón de diseño. Este experimento se realizó en el curso académico 2010-2011, contando con un total de 21 trabajos.

La Tabla 1 muestra las calificaciones obtenidas para cada trabajo, tanto por parte de los alumnos como del profesor.

Como puede observarse, las notas que asignan los alumnos son bastante altas y todos los trabajos quedan puntuados en torno al 7 (media 7.36 y desviación 0.43). Si se analizan individualmente las puntuaciones de cada alumno, vemos como son reticentes a asignar una calificación menor a unos 6 puntos, ya que lógicamente asocian las notas inferiores a 5 a suspender el trabajo y no quieren perjudicar al resto de compañeros. Además, no distinguen claramente los peores trabajos de los mejores de forma objetiva, aunque sí que puntúan mejor unos trabajos que otros. Esto puede estar motivado posiblemente por factores subjetivos como el afecto a la persona que presenta o lo simpatía al tema elegido. Estos factores hacen que al final todas las notas converjan hacia el notable, con muy poca variación (media 7.36 y desviación 0.43). Por otro lado, la evaluación reali-

Técnica de Evaluación Numérica Individual

Trabajos	Profesor	Alumnos		Diferencias Prof-Alum
		Media	Desviación	
T1	9	7,97	0,92	1,05
T2	6,5	7,75	1,08	1,56
T3	3	6,60	0,82	12,96
T4	7	7,48	1,21	0,23
T5	5,5	7,45	0,87	3,80
T6	6	7,70	1,23	2,89
T7	7	7,19	0,91	0,04
T8	4	6,55	1,36	6,50
T9	6	7,43	0,83	2,03
T10	5	7,05	1,33	4,20
T11	10	7,90	0,85	4,41
T12	6	7,95	1,05	3,80
T13	7	7,65	0,95	0,42
T14	6,5	7,68	0,91	1,38
T15	7	7,65	0,80	0,42
T16	4	6,99	0,88	8,94
T17	6	7,30	1,09	1,69
T18	4	6,63	1,09	6,89
T19	6,5	7,33	1,38	0,68
T20	5	7,20	1,34	4,84
T21	4	7,08	1,87	9,46
Media	5,95	7,36	1,08	8,8
Desviación	1,69	0,43		

Cuadro 1: Resultados de la Técnica de Evaluación Numérica Individual

zada por el profesor, más objetiva, da como resultado una nota media más baja (5.95) pero sobre todo, un rango mayor de notas que oscila desde el 3 al 10 (desviación 1.69) en el que claramente se diferencian trabajos malos de buenos.

La Figura 1 muestra las calificaciones del profesor y los alumnos para cada trabajo, y permite observar gráficamente las diferencias entre ambas.

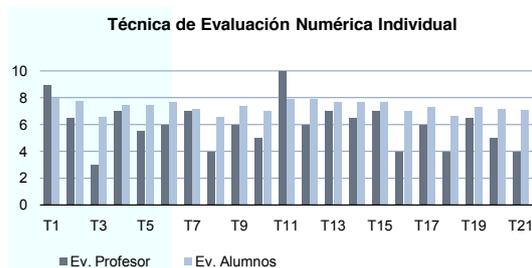


Figura 1: Comparativa de Calificaciones de la Técnica de Evaluación Numérica Individual

Algunos ejemplos concretos son el trabajo T3, que según el profesor fue el peor de todos y obtuvo una calificación de 3, pero que sin embargo los alumnos lo puntuaron con 6.60. Y en el otro extre-

mo, el trabajo T11, que el profesor puntuó con la nota más alta, un 10, y los alumnos calificaron con un 7.90. Con estos ejemplos podríamos pensar que los alumnos están diferenciando trabajos malos de buenos, pero que otorgando notas siempre entre un 6.50 y un 8. Por otro lado, encontramos casos como el trabajo T12, al que los alumnos identifican como muy bueno y le asignan un 7.95, pero el docente le asigna un 4.

Como conclusión de esta técnica de evaluación, podríamos decir que las calificaciones de los alumnos son poco objetivas, ya que aun existiendo trabajos buenos y malos, todos los trabajos obtienen una puntuación más o menos homogénea. Y si comparamos las calificaciones de los alumnos con las del profesor, observamos que hay una gran diferencia (dispersión 8.84), por lo que podemos concluir que este método no es adecuado.

5. Técnica de Evaluación Numérica por Grupos

En este experimento se analizó la técnica de evaluación numérica por grupos, en la que la calificación del resto de trabajos no era individual como en el caso anterior, sino que los alumnos se organizaron en grupos y cada uno puntuaba como grupo, debiendo así debatir la nota que quería asignar al resto de grupos. Esta nota era una nota numérica entre 0 a 10. En principio, con esta técnica deberían mitigarse algunos de los problemas identificados en la técnica anterior y obtenerse calificaciones más objetivas.

Este experimento se realizó en el curso académico 2010-2011 y se contó con 7 trabajos, realizados cada uno por un grupo de 4 alumnos. El trabajo consistía en describir una metodología de desarrollo software realizando una memoria y una presentación ante la clase.

La Tabla 2 muestra las calificaciones asignadas a cada trabajo, siguiendo la misma estructura que la Tabla 1, pero esta vez, al contar con un menor número de trabajos a evaluar se detallan las puntuaciones asignadas por cada grupo de estudiantes.

Al analizar las calificaciones de los estudiantes, puede observarse como la puntuación media de los trabajos sigue siendo más alta (6.98 de media) que la proporcionada por el docente (5.96 de media). La desviación aunque es mayor que la obtenida en

Técnica de Evaluación Numérica por Grupos

Trabajos	Profesor	Alumnos										Diferencias Prof-Alum
		Media	Desviación	R1	R2	R3	R4	R5	R6	R7		
T1	2,5	5,67	0,41		6	5	6	5,5	6	5,5		10,03
T2	9	7,42	1,23	9		7,75	7,75	7,5	8	5		2,51
T3	7	7,21	0,46	7	7		7,75	7,5	7,5	6,5		0,04
T4	4,2	6,58	1,02	7	7	6		8	5	6,5		5,68
T5	8,5	7,54	0,81	7	8	7	8,25		8,5	6,5		0,92
T6	4	7,00	1,14	7	9	7	6,5	7		5,5		9,00
T7	6,5	7,46	1,10	7	6,5	8	8,25	9	6,5			0,92
Media	5,96	6,98	0,88									5,39 Dispersión
Desviación	2,45	0,67										

Cuadro 2: Resultados de la Técnica de Evaluación Numérica por Grupos

el experimento anterior, sigue siendo baja (0.67) y bastante menor que la del docente (2.45). Podemos comprobar cómo el rango de notas asignadas por los estudiantes varía algo más que en el experimento anterior, de 5.67 puntos para el peor trabajo a 7.54 puntos para el mejor, aunque sigue siendo un rango mucho más limitado que el usado por el profesor, que asigna al peor trabajo 2.5 puntos y al mejor 9 puntos.

Observando las puntuaciones realizadas por cada grupo (columnas R1 a R7), podemos ver cómo hay grupos menos objetivos, como el R7, que han puntuado todos los trabajos con notas entre 5 a 6.5 puntos, o como el grupo R1 que puntúan casi todos los trabajos con 7 puntos.

Para analizar la objetividad de los alumnos se han comparado sus calificaciones con las del profesor. La Figura 2 muestra gráficamente estas diferencias y podemos comprobar como existen varios casos en los que la diferencia de calificaciones es más que evidente. Puede observarse cómo los alumnos son reticentes a suspender a sus propios compañeros y aunque identifican trabajos malos, como el trabajo T1, al que le asignan una calificación de 5.67 puntos mientras que el docente le asigna 2.5 puntos. Podríamos considerar que están detectando bien trabajos buenos y malos, y por lo tanto, el que asignen notas altas podría ponderarse, pero sin embargo encontramos casos en los que aparecen puntuaciones altas no objetivas, como dos 9 con los que puntúan a los trabajos T6 y T7, a los que el docente ha calificado con un 4 y un 6.5 respectivamente.

Tras analizar los resultados, podemos concluir que este método ha mejorado el caso anterior expandiendo un poco más las calificaciones de los alum-

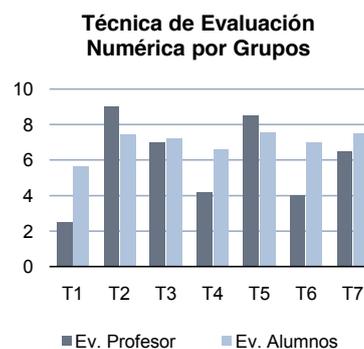


Figura 2: Comparativa de Calificaciones de la Técnica de Evaluación Numérica por Grupos

nos, que ya no varían en torno al 7 sino que nos encontramos notas desde el 5.5 al 7.5 aproximadamente. Por otro lado, con este método ha mejorado la objetividad de los alumnos, ya que la diferencia de sus calificaciones con respecto a las del profesor es inferior (ha pasado de una desviación de 8.84 a 5.39). Esto es debido a que con este método, los alumnos no actúan individualmente sino que han de debatir y generar una única calificación consensuada por todo el grupo.

6. Técnica de Evaluación por Ordenación por Grupos

El tercer experimento se centra en la técnica de evaluación por ordenación en grupos. En este caso, las calificaciones en vez de realizarse de forma numérica de 0 a 10 puntos, se realizan ordenando los

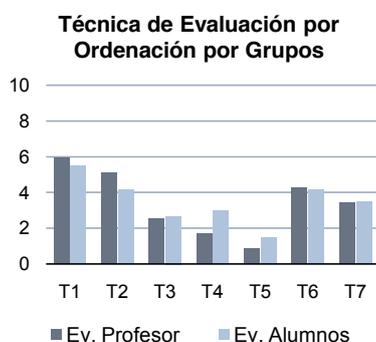


Figura 3: Comparativa de Calificaciones de la Técnica de Evaluación por Ordenación por Grupos

trabajos de mejor a peor. Con este cambio se pretende conseguir un mejor equilibrio de las calificaciones, para evitar que estén centradas en el notable, como en las dos técnicas de evaluación analizadas anteriormente.

Este experimento puede ser considerado como una réplica del experimento anterior realizada en el siguiente curso académico, 2011-2012. En este curso, volvió a realizarse el trabajo de metodologías de desarrollo software en grupos, que debían generar una memoria y presentación ante el resto de la clase. Y además, se analizaron el mismo número de trabajos que en el experimento anterior, 7 trabajos realizados en grupos de 4 alumnos.

Las calificaciones asignadas a cada trabajo por los estudiantes y el docente, se presenta en la Tabla 3. Al analizar los órdenes establecidos por los estudiantes, podemos observar cómo ningún trabajo está en los extremos, es decir, ni está identificado claramente como el mejor (posición 1) ni como el peor (posición 6). Para que esto pasara debería haber unanimidad, lo cual es muy difícil que ocurra. Quizás el caso más claro es el trabajo T1, al que tres grupos identifican como el peor (posición 6), dos grupos como el segundo peor (5), y un grupo, extrañamente, lo califica como de los mejores (2). En este experimento, las medidas de media y varianza no son determinantes, ya que al forzar el tener que ordenar todos los trabajos, resultarán medidas parecidas.

Los órdenes asignados por el docente nos sirven para calcular la objetividad de los alumnos y comprobar si el orden establecido por ellos es muy dispar al establecido por él. Para ello se calcula por cada

trabajo la diferencia entre el orden asignado por el docente y el orden asignado por los estudiantes. Estas diferencias han sido representadas gráficamente en la Figura 3.

Puede observarse cómo las diferencias obtenidas, por lo general, no son tan marcadas como en los experimentos anteriores. La diferencia es máxima para algún trabajo, como el T4, que el docente ha considerado como de los mejores, pero los estudiantes no lo han considerado por unanimidad. Si observamos las calificaciones realizadas por cada grupo (columnas R1 a R7) podemos comprobar como sí que existe cierta unanimidad en la mayoría de los casos, y aunque cada grupo haya introducido algunas calificaciones poco objetivas, en general se han identificado los trabajos buenos y malos.

Al calcular la desviación total de los órdenes establecidos por el docente y por los alumnos, obtenemos una puntuación bastante inferior que en el experimento anterior (pasando de 5.39 a 1.79), lo que indica que este método es el que nos ha proporcionado calificaciones más objetivas de los tres.

7. Conclusiones

La participación de los alumnos en el proceso de evaluación les otorga la práctica necesaria para en un futuro poder evaluar de forma crítica y objetiva trabajos técnicos relacionados con su profesión. Además, ayuda a la adquisición de otras competencias, tales como la capacidad de crítica, la autocrítica, el debate, la defensa de opiniones propias, el asumir consensos para las calificaciones, etc. Sin embargo, el involucrar a los alumnos en las evaluaciones tiene problemas asociados principalmente a la falta de objetividad de las calificaciones que efectúan.

En este artículo, se han presentado varias técnicas para incorporar a los estudiantes en el proceso de evaluación: (i) cada alumno califica de forma individual al resto, con notas numéricas de 0 a 10; (ii) cada grupo de alumnos califica al resto de grupos con notas numéricas de 0 a 10; y (iii) cada grupo califica ordenando el resto de grupos de mejor a peor. A continuación el artículo presenta los resultados de evaluar cada una de estas técnicas mediante experimentos realizados con alumnos de Ingeniería Informática durante dos cursos académicos.

Los resultados obtenidos indican que la técnica de evaluación numérica individual, no es adecuada, ya

que las puntuaciones resultantes son muy bondadosas y sitúan todos los trabajos evaluados en el notable, y además, difieren mucho de las asignadas por el docente. La técnica de evaluación numérica por grupos, al puntuar en grupo debatiendo las notas, mejora la anterior en dos aspectos, pero levemente: (i) mejora que la calificación de los trabajos se hace con un mayor rango de notas, pero aún centrado en el notable, y (ii) que es un poco más cercana a la evaluación del docente.

En estos experimentos hemos observado como los alumnos, puntúan los trabajos en muchos casos guiados más por relaciones personales que por razones objetivas, lo que genera que trabajos buenos sean mal puntuados y viceversa. Además, son reticentes a suspender a los compañeros, así que utilizan un rango corto de puntuaciones que suelen ser de 6 a 9 puntos, y que al final hace que las notas medias converjan al notable. De este modo, trabajos buenos y malos no quedan claramente diferenciados, ya que en algunos casos su puntuación difiere pocas décimas.

Es la última técnica, la evaluación por ordenación por grupos, la que ofrece mejores resultados, ya que al forzar a los alumnos a ordenar los trabajos de mejor a peor obtenemos un orden más cercano al establecido por el docente. Es decir, podríamos concluir con que los alumnos sí identifican correctamente trabajos buenos de malos, pero al tener la presión de asignar una nota numérica y no querer perjudicar a los compañeros, no asignan calificaciones objetivas. En este caso, al no utilizar una nota numérica y verse forzados a asignar un orden, sí que lo hacen de forma más objetiva.

Como conclusión final tras los experimentos realizados, creemos que la participación de los alumnos en el proceso de evaluación ha de realizarse siempre de forma supervisada, ya que en muchos casos hay una falta de objetividad más que evidente. En caso de utilizar una de las técnicas presentadas en este artículo, recomendamos utilizar la calificación mediante ordenación de trabajos de mejor a peor, y realizada en grupos para que haya debate y consenso. Esta técnica es la que ha proporcionado unas calificaciones más cercanas a las del docente. Sin embargo, posteriormente el docente debería asignar notas numéricas a cada uno. En caso contrario, si se realizara de forma automática (por ejemplo al mejor un 10, al segundo un 8, etc. hasta asignarle un 0 al peor)

habría casos en los que las puntuaciones no serían realistas. Por ejemplo, si ningún trabajo mereciera suspender, si existieran varios trabajos buenos, etc. Es por esto que el docente debería asignar las notas numéricas finales, aunque teniendo en cuenta el orden de trabajos obtenidos y la opinión del alumnado.

Agradecimientos

Este trabajo se encuadra dentro del proyecto de innovación docente “Aprendizaje basado en proyectos para Ingeniería del Software”, financiado por la Universidad de Cantabria.

Referencias

- [1] A. Abelló and X. Burgués. Puntuación entre iguales para la evaluación del trabajo en equipo. In *Jornadas de Enseñanza Universitaria de la Informática (JENUI)*, pages 73–80, Sevilla, 5-8 de Julio de 2011 2011.
- [2] A. Irons Alexander and S. *Effective learning & teaching in computing*. Effective Learning and Teaching in Higher Education Series. RoutledgeFalmer. 2004.
- [3] J. Biggs. *Aligning teaching and assessing to course objectives*, 2003.
- [4] G. Brown. *Assessment: A guide for lecturers*. Itsn generic centre, 2001.
- [5] S. Ketteridge H. Fry and S. Marshall. *A Handbook for Teaching and Learning in Higher Education: Enhancing Academic Practice*. 2009.
- [6] T. Ahoniemi D. Cukierman I. Hernán-Losada J. Jackova E. Lahtinen T. L. Lewis D. M. Thompson C. Riedesel U. Fuller, C. G. Johnson and E. Thompson. Developing a computer science-specific learning taxonomy. *ACM SIGCSE Bulletin*, 39 (4):152–170, 2007.
- [7] J. van der Pol, B. A. M. van den Berg, W. F. Admiraal, and P. R. J. Simons. The nature, reception, and use of online peer feedback in higher education. *Computers and Education*, 51 (4):1804–1817, 2008.
- [8] M. van Zundert, D. Sluijsmans, and J. van Merriboer. Effective peer assessment processes: Research findings and future directions. *Learning and Instruction*, 20 (4):270–279, 2010.

Técnica de Evaluación por Ordenación por Grupos

Trabajos	Profesor	Alumnos									Diferencias Prof-Alum
		Media	Desviación	R1	R2	R3	R4	R5	R6	R7	
T1	5,95	5,50	0,55		5	6	6	6	5	5	0,2025
T2	5,1	4,17	1,47	4		5	5	3	6	2	0,87
T3	2,55	2,67	1,21	2	2		3	1	4	4	0,01
T4	1,7	3,00	1,10	1	4	3		4	3	3	1,69
T5	0,85	1,50	0,84	3	1	2	1		1	1	0,42
T6	4,25	4,17	1,47	5	3	4	2	5		6	0,01
T7	3,4	3,50	2,17	6	6	1	4	2	2		0,01
Media	3,40	3,50	1,26								1,79 Dispersión
Desviación	1,84	1,28									

Cuadro 3: Resultados de la Técnica de Evaluación por Ordenación por Grupos