

XV JENUI. Barcelona, 8-10 de julio de 2009

ISBN: 978-84-692-2758-9

<http://jenui2009.fib.upc.edu/>

Evaluación entre compañeros: estudio de su correlación con la evaluación del profesor

Javier Oliver, Verónica Canivell

Facultad de Ingeniería - ESIDE

Universidad de Deusto

Avda. de las Universidades 24, 48007 Bilbao

{oliver, vcanivel}@eside.deusto.es

Resumen

La evaluación entre compañeros es una técnica docente que permite desarrollar en los estudiantes capacidades cognitivas superiores como el pensamiento crítico y la capacidad de análisis. Es una buena herramienta para profundizar en el aprendizaje, ya que evaluar implica entender, ponderar, considerar alternativas, etc. Pero, ¿se pueden utilizar las calificaciones obtenidas de la evaluación entre compañeros para calcular la nota final de los alumnos? En este trabajo se analiza la correlación entre las calificaciones de los alumnos y del profesor en un intento de responder a la pregunta anterior.

1. Introducción

La evaluación entre compañeros consiste en que los estudiantes revisen los trabajos y resultados de aprendizaje de sus iguales y les proporcionen algún tipo de información acerca de dicha evaluación [11]. La evaluación puede ser de tipo formativo, cuyo objetivo fundamental es mejorar el aprendizaje de los alumnos a través de los comentarios de los compañeros, o incluir también un componente sumativo, es decir, ayudar a definir la calificación final de los alumnos. La evaluación entre compañeros puede facilitarse mediante la utilización de herramientas informáticas adecuadas [5, 12]

Uno de los aspectos más positivos de la evaluación entre compañeros es que fomenta el desarrollo de habilidades cognitivas superiores, como el pensamiento crítico, ya que los estudiantes deben utilizar su propio conocimiento para analizar y valorar el trabajo de los demás, proporcionándoles elementos de crítica constructiva para mejorar su trabajo [10]. Estas ventajas parecen obtenerse incluso si se hace una

evaluación rápida del trabajo de los compañeros durante unos pocos minutos en clase [3].

La búsqueda de la opinión de los compañeros para mejorar el trabajo propio es práctica habitual en el mundo académico y empresarial, por lo que la capacidad de criticar de manera constructiva el trabajo de los demás y de recibir críticas supone una competencia deseable en el mundo laboral [1].

La evaluación entre compañeros es una técnica docente que puede ser beneficiosa para todos los agentes implicados: el autor del trabajo, el revisor del trabajo y el propio profesor [2]. En clases con muchos estudiantes, los autores de un trabajo no pueden recibir mucha información de la corrección del profesor. La corrección de los compañeros supone una realimentación añadida, que aunque proviene de una fuente menos experta, puede proporcionar información muy útil, sobre todo si cada trabajo es evaluado por varios compañeros. La evaluación de los compañeros puede basarse en los mismos criterios de corrección que usará el profesor en su evaluación sumativa, con lo que los comentarios de los compañeros pueden ayudar a los autores a obtener mejores calificaciones.

El proceso de evaluación también ofrece ventajas para quien lo realiza, como la mejor comprensión de los criterios de evaluación después de haberlos usado uno mismo o la familiarización con la evaluación entre iguales, muy habitual en el mundo académico.

Si la evaluación entre compañeros se hace correctamente, el profesor podría incluso utilizarla como parte de su evaluación sumativa y ahorrar algo de tiempo en la corrección, aunque es cierto que la organización de la actividad, también lleva su tiempo, y que hay que comprobar la validez de las evaluaciones de los estudiantes.

Hay autores que recalcan la importancia de que los alumnos que realizan la evaluación

dispongan de unas criterios de corrección claros, concretos y explícitos proporcionados por el profesor [7].

En este trabajo se investiga el grado de correlación entre las calificaciones otorgadas por los alumnos y por el profesor para unos ejercicios de diseño de bases de datos, así como la opinión subjetiva que los alumnos tienen de la evaluación entre compañeros.

2. Metodología

El estudio se realizó en un grupo de alumnos de la asignatura “Sistemas de Gestión de Bases de Datos” de 3^{er} curso de la titulación de Ingeniería Técnica en Informática de Gestión.

Se pidió a los alumnos que, de forma voluntaria, realizaran en casa un ejercicio de diseño de una base de datos relacional en tercera forma normal. El ejercicio consistía en un enunciado en el que se recogía la realidad de una empresa y la información que habría que incluir en una base de datos relacional. El alumno debía analizar dicha realidad, añadir algunos datos que considerara necesarios, además de los descritos, y aplicar al conjunto de datos de partida el proceso de normalización para obtener un diseño de base de datos relacional en tercera forma normal. Para asegurar el anonimato, los ejercicios solo se identificaron con el DNI. Una vez entregados y fotocopiados por cuadruplicado, se explicó en clase la resolución del ejercicio y se informó a los alumnos de cómo debían acometer el proceso de corrección. Estas informaciones incluyeron:

- Con el fin de aliviar la sensación de responsabilidad que algunos manifestaban, se les dijo que las notas que pusieran a sus compañeros no iban a servir para calcular la calificación final de los mismos.
- Los ejercicios debían calificarse con una nota numérica sobre 10.
- Ante las preguntas de para qué les servía a ellos la realización de esa tarea, se les explicó que la corrección entre compañeros era una técnica de aprendizaje en sí misma, y que además de desarrollar su capacidad de análisis y espíritu crítico, podrían aprender cómo habían resuelto sus compañeros el mismo problema y apreciar la dificultad y tiempo que supone corregir un trabajo.

- Al calificar los ejercicios, había que tener en cuenta tanto el procedimiento de normalización seguido, como el resultado final obtenido, valorando con cinco puntos cada uno de ellos.

A continuación se repartieron los ejercicios, de modo que cada uno fue corregido por cuatro alumnos, y posteriormente el profesor hizo también su corrección con lo que se, obtuvieron cinco calificaciones por ejercicio.

La semana siguiente se volvió a solicitar a los alumnos que, voluntariamente, realizaran un nuevo ejercicio de diseño de una base de datos relacional similar al primero. Todo el proceso fue análogo al del primer ejercicio.

Finalmente se pidió a los alumnos que dieran su opinión en formato libre acerca del proceso de corrección entre compañeros, resaltando los aspectos positivos y negativos que encontraran en el mismo.

3. Resultados

De un total de 51 alumnos matriculados en la asignatura de “Sistemas de Gestión de Bases de Datos”, en el grupo en el que se realizó el estudio, y con una asistencia media a clase de en torno a 40 alumnos, 19 alumnos entregaron el primer ejercicio y 27 entregaron el segundo. En el segundo ejercicio hubo dos correcciones no válidas, por lo que solo pudieron usarse 25 evaluaciones. En las figuras 1 y 2 se muestran las correcciones del profesor y la media de las correcciones de 4 alumnos para los Ejercicios 1 y 2 respectivamente.

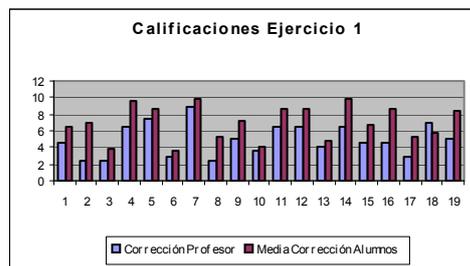


Figura 1. Calificaciones sobre 10 del Ejercicio 1. En color claro, las calificaciones del profesor, y en color oscuro la media de las calificaciones de 4 alumnos.

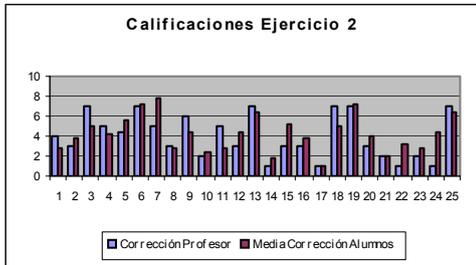


Figura 2. Calificaciones sobre 10 del Ejercicio 2. En color claro, las calificaciones del profesor, y en color oscuro la media de las calificaciones de 4 alumnos.

Al comparar las correcciones del profesor y los alumnos, se aprecia que éstos tienden a otorgar calificaciones más altas. Este efecto es más marcado en el Ejercicio 1 que en el 2.

Además de la comparación de las correcciones del profesor y los alumnos mediante un diagrama de barras, se pueden representar los datos en un diagrama de dispersión o nube de puntos. Las figuras 3 y 4 muestran los diagramas de dispersión de los Ejercicios 1 y 2.

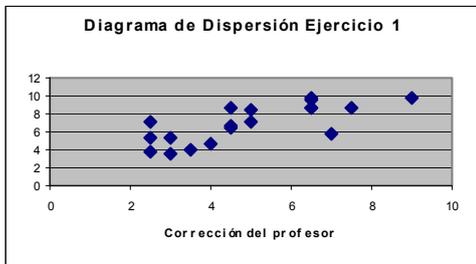


Figura 3. Diagrama de dispersión con la corrección del profesor en el eje de las X, y la de los alumnos en el eje de las Y.

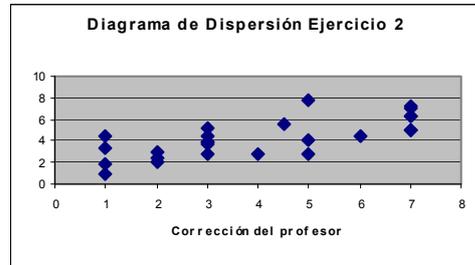


Figura 4. Diagrama de dispersión con la corrección del profesor en el eje de las X, y la de los alumnos en el eje de las Y.

Parece haber una cierta relación entre la corrección del profesor y la corrección de los alumnos: a mayor nota del profesor, mayor nota tienden a poner los alumnos (o viceversa), pero la nube de puntos no está excesivamente concentrada.

Para cuantificar la relación entre los datos, calculamos el *coeficiente de correlación lineal* entre las correcciones de profesor y alumnos, que es de 0,76 para el Ejercicio 1 y de 0,73 para el Ejercicio 2. Estos valores son similares a los obtenidos por otros autores [7], pero solo miden hasta qué punto la relación entre los datos es lineal y positiva, y no miden otra cosa: dos variables pueden tener un coeficiente de correlación lineal igual a cero, y tener una clara relación, pero de tipo cuadrático [8].

También nos interesa saber el grado de dispersión de las puntuaciones otorgadas por los alumnos, por lo que calculamos su desviación típica. En las figuras 5 y 6 se muestran las desviaciones típicas de las calificaciones de los alumnos de los Ejercicios 1 y 2. Se observa una dispersión importante, con un 63% de los valores en el Ejercicio 1, y un 72% de los valores en el Ejercicio 2 con una desviación típica >1.

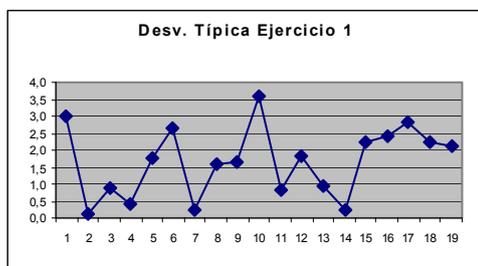


Figura 5. Desviación típica de las calificaciones de los alumnos al Ejercicio 1.

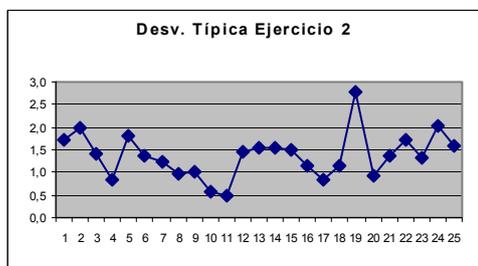


Figura 6. Desviación típica de las calificaciones de los alumnos al Ejercicio 2.

Además del análisis cuantitativo de los datos, también se ha hecho una encuesta para averiguar la opinión que los alumnos tienen de la evaluación entre compañeros como metodología de aprendizaje. Un total de 16 alumnos dieron su opinión por escrito acerca de los aspectos positivos y negativos que encontraban en la técnica de evaluación entre compañeros. A continuación se describen las opiniones más frecuentemente expresadas.

En seis casos, los alumnos opinan que esta técnica puede ayudar en el aprendizaje, como refleja este ejemplo:

“El ejercicio a corregir se valora y se discute de un modo más intenso. Los errores encontrados en el ejercicio se estudian con más profundidad y se aprende de una forma más efectiva.”

En cinco ocasiones piensan que sirve para valorar la dificultad del trabajo del profesor:

“Las ventajas de este ejercicio son que nos hemos dado cuenta de lo difícil que es valorar un ejercicio de diseño.”

También en cinco ocasiones opinan los alumnos que ayuda a comprender la necesidad de realizar los trabajos y exámenes de manera clara y legible:

“Creo que ha servido para hacer un examen lo más claro y explicado posible puesto que es difícil valorar algo que está mal detallado o poco organizado.”

Las valoraciones críticas incluyen en cinco ocasiones que los profesores ponen peor nota que los alumnos:

“Quizá el profesor tiende a ser más riguroso y estricto, ya que sus conocimientos y experiencia son mayores, y espera equivocadamente lo mismo de sus alumnos.”

También se opina en cinco ocasiones que las notas que ponen los alumnos son muy dispares:

“Hay gente que infla las notas, y otros que ponen notas más bajas de las debidas.”

En conjunto podemos decir que los alumnos valoran positivamente la evaluación entre compañeros como método de aprendizaje, pero reconocen discrepancias entre las notas de los alumnos y las notas del profesor.

4. Discusión

La evaluación entre compañeros está recibiendo un interés creciente en la formación universitaria, y hay muchos estudios que analizan la validez de las calificaciones comparadas con las de los profesores [4, 9, 7 y 11]. Aunque en muchos de estos estudios se obtienen altos grados de correlación entre las calificaciones de profesores y alumnos, la comparación entre distintos estudios es compleja, ya que hay muchos factores que pueden afectar las calificaciones de los estudiantes:

- Grado de anonimato con el que se desarrolla el proceso para los autores de los trabajos y para los evaluadores.
- Puesta a disposición de los estudiantes de unos criterios de corrección claros y explícitos.
- Naturaleza de los trabajos a corregir. Parece claro que aquellos ejercicios en los que hay muchas posibles respuestas válidas, tendrán una corrección más variable que aquellos ejercicios con respuestas más cerradas y concretas.
- Notificación a los estudiantes del uso que se va a hacer de sus correcciones: puede ocurrir que si los estudiantes saben que sus calificaciones se van a usar para el cálculo final de las notas, se vean influenciados para corregir al alza (o a la baja).

Nosotros hemos trabajado con dos grupos de 19 y 25 ejercicios y hemos obtenido un índice de correlación lineal de 0,76 y 0,73 respectivamente. En [7] se hace un estudio sobre 13 ejercicios y se halla un índice de correlación de 0,758 (aunque usa la correlación intra-clase en lugar del índice de Pearson de nuestro estudio). Por su parte, en [9] se trabaja con un grupo de 48 estudiantes y se analizan hasta 30 parámetros de la corrección de unas páginas web desarrolladas por los alumnos. El índice de correlación de Pearson alcanza en este estudio 0,991.

Aunque los índices de correlación que hemos hallado se aproximan bastante a 0,8, que es un valor suficiente para muchos autores, si se calificara a los alumnos con la nota de sus compañeros, en el caso del Ejercicio 1, nos encontraríamos con que, de 19 correcciones:

- En 12 casos las notas tendrían una diferencia de dos o más puntos con la del profesor.
- En seis casos la nota pasaría de un aprobado a un suspenso (o viceversa).

Para el Ejercicio 2, de 25 correcciones:

- En tres casos las notas tendrían una diferencia de dos o más puntos con la del profesor.
- En cinco casos la nota pasaría de un aprobado a un suspenso (o viceversa).

Por todo ello, no parece que la proximidad entre las notas de los alumnos y del profesor sea suficiente para poder utilizar aquéllas en la calificación de los trabajos, al menos en el contexto de ejercicios de respuesta abierta, como los diseños de bases de datos.

5. Conclusiones

La evaluación entre compañeros es una buena herramienta de aprendizaje activo, y está bien valorada por los alumnos en ese papel, pero las calificaciones de los alumnos se alejan demasiado de las del profesor como para utilizarlas de manera rutinaria en la calificación de trabajos de respuesta abierta.

Referencias

- [1] Anewalt K. *Using Peer Review as a vehicle for communication skill development and active learning*. Consortium for Computing Sciences in Colleges, 2005.
- [2] Bostock S. *Web Support for Student Peer Review*. Keele University Innovation Project Report, 2002.
- [3] Denning T., Kelly M., Lindquist D., Malani R., Griswold W. G. and Simon B. *Lightweight Preliminary Peer Review: Does In-Class Peer Review Make Sense?* SIGCSE'07, Covington, Kentucky, March 7-10, 2007.
- [4] Falchikov N. y Goldfinch J. *Student peer assessment in higher education: a meta analysis comparing peer and teacher marks*. Review of Educational Research, 70:3, 2000.
- [5] Gehringer E. F. *Electronic Peer Review and Peer Grading in Computer science Courses*. SIGCSE'01, Charlotte, NC, 2001.
- [6] Gehringer E. F., Chinn D. D., Ardis M. A., Pérez-Quiñones, M. A. *Using Peer Review in teaching Computing*. SIGCSE'05, St. Louis, February 23-27, 2005.
- [7] Mirielli E. J. *Using Peer Evaluation in a Website Design Course*. Consortium for Computer Sciences in Colleges, 2007.
- [8] Narvaiza J. L., Laka J. P., Madariaga J. A. y Ugarte J. V. *Estadística descriptiva y probabilidad*. Biblioteca de Gestión Desclee de Brouwer, 1998.

- [9] Sahin S. *An application of peer assessment in higher education*. The Turkish Online Journal of Education Technology – TOJET. April 2008 ISSN: 1303-6521 volume 7, Issue 2, Article 1.
- [10] Sitthiworachart J. and Joy M. *EffectivePeer Assessment for Learning Computer Programming*. ITiCSE'04, Leeds UK, June 28-30, 2004
- [11] Topping K. *Peer assessment between students in colleges and universities*. Review of Educational Research, 68, 249-276, 1998
- [12] Wolfe W. J. *Online Student Peer Reviews*. SIGITE'04, Salt Lake City, October 28-30, 2004.