



Análisis de tres técnicas diferentes para la evaluación individualizada de tareas realizadas por equipos de estudiantes*

Pablo Sánchez

Dpto. Ingeniería Informática y Electrónica
Universidad de Cantabria
Santander (Cantabria)
p.sanchez@unican.es

Resumen

Actualmente las empresas demandan profesionales con no sólo una serie de cualidades técnicas, sino también formados en una serie de habilidades y competencias, tales como habilidades para la gestión, que le permitan desarrollar con éxito su labor profesional. Entre estas competencias, muchas veces denominadas transversales, la capacidad de trabajo en equipo suele ser una de las más demandadas. Por tanto, con el objetivo de desarrollarla, diferentes metodologías de enseñanza-aprendizaje han ido incorporando actividades que requieran trabajar en equipo. Sin embargo, la incorporación de estas actividades añade una serie de dificultades a dichos procesos de aprendizaje. Entre ellas destacan la resolución de conflictos entre los miembros del equipo y la capacidad de calificar a los alumnos individualmente, de acuerdo a su grado de adquisición de competencias, a partir de elementos producidos colectivamente. Con el objetivo de minimizar esas dificultades, hemos ido refinando a lo largo de 3 años una técnica para calificar individualmente a los alumnos de una asignatura de Ingeniería del Software en la que trabajan en equipo. El presente artículo describe la evolución de dichas técnicas, analizando luces y sombras de la evolución.

Palabras clave: Trabajo en Grupo, Aprendizaje Basado en Proyecto, Calificación, Evaluación

1. Introducción

En un mercado laboral con un creciente número de licenciados y graduados, los empleadores demandan egresados universitarios que posean ya no sólo una serie de competencias técnicas, sino además una serie de competencias genéricas o transversales, tales como capacidad de trabajo en equipo [17] o capacidad de liderazgo [19], que les permitan desarrollar plenamente dichas competencias técnicas en entornos productivos reales.

De entre las competencias transversales, el trabajo en equipo suele encontrarse entre las competencias más demandadas [9]. Con el objetivo, entre otros, de desarrollar esta competencia, diversas metodologías docentes, tales como el *aprendizaje basado en proyectos* [8, 15], han ido incorporando dentro de sus procesos de enseñanza-aprendizaje actividades que deben ser desarrolladas en equipo.

No obstante, la incorporación de actividades en equipo plantea una serie de nuevas dificultades que deben ser solventadas para un adecuado aprendizaje. Por un lado, al trabajar en equipo pueden surgir conflictos por diversos motivos entre los miembros de dicho equipo. Dichos conflictos deberán

ser convenientemente resueltos tanto para asegurar el éxito del equipo como para preservar un clima favorable al aprendizaje. Por otro lado, aunque los alumnos trabajen en equipo, no todos los miembros de un equipo adquieren las competencias trabajadas a lo largo del proceso de aprendizaje con la misma intensidad. Por tanto, cada alumno debería ser calificado individualmente de acuerdo a su grado real de adquisición de competencias.

Con el objetivo de solventar estas dificultades, hemos ido refinando, a lo largo de tres cursos académicos, una técnica para la calificación individual de alumnos que han trabajado en equipo. Este artículo describe el proceso de evolución hasta el estado actual, en el cual se considera la técnica como estable y madura. Dicho proceso de evolución se ha llevado a cabo dentro de una asignatura de Ingeniería del Software I, la cual estaba organizada de acuerdo a una metodología de aprendizaje basado en proyectos [12].

La técnica actual, fruto de la mencionada evolución, posee las siguientes ventajas:

1. Permite calificar a los alumnos de forma individual;
2. Permite resolver de forma fácil ciertos tipos de conflic-

*Este artículo recibió uno de los dos premios a mejor ponencia de las Jenui 2014.

tos;

3. Equilibra la carga de trabajo de cada alumno del grupo;
4. Asegura que cada alumno trabaja las competencias por las cuales será evaluado al final de la asignatura;
5. Procura maximizar la productividad del grupo mediante una adecuada división del trabajo.

La efectividad de la técnica se ha evaluado mediante diversos indicadores objetivos y una encuesta de satisfacción del alumnado. No obstante, ciertos factores externos, como problemas con otras asignaturas, han hecho que consideremos los datos recabados como no conclusivos. Por tanto, consideramos necesaria una posterior investigación para poder determinar empíricamente el rendimiento de la técnica.

El siguiente apartado describe el contexto en el cual se enmarca el trabajo presentado en este artículo.

2. Contexto

Las técnicas de calificación que se describen en este artículo se han utilizado dentro de la asignatura de Ingeniería del Software I, del Grado en Ingeniería Informática de la Universidad de Cantabria. Dicha asignatura se organizaba de acuerdo a una *metodología de aprendizaje basada en proyectos* [12], ya que entendemos que la mejor manera de aprender a desarrollar un proyecto *software* es precisamente mediante el desarrollo de un proyecto *software*. Éste se desarrollaba en grupos de 4 ó 5 alumnos que eran establecidos por el equipo docente al principio del curso.

Dado que parece probado que los grupos heterogéneos son más productivos [4, 5], se procuraba maximizar la heterogeneidad de grupo. Para ello se determinaban diferentes aspectos de la personalidad de cada alumno mediante el *test de Lüscher* [7]. Dicho test permite identificar ciertos rasgos de la personalidad de un individuo mediante una simple selección de colores de entre un conjunto determinado. Una vez identificados los rasgos, procurábamos formar grupos que contuviesen personalidades complementarias. Los detalles acerca de cómo se realiza este proceso de formación de grupos se pueden encontrar en un trabajo anterior [12].

Para el desarrollo del proyecto los alumnos debían completar un portafolio electrónico [18] definido sobre la plataforma *Moodle*. Cada etapa del portafolio se correspondía con una de las primeras fases del ciclo de vida *software*.

En primer lugar, debían elaborar, sobre una idea genérica común a todos los grupos, su propia idea de negocio. Durante el curso 2011–2012 los alumnos desarrollaron variaciones de un juego en línea para la gestión de clubes de fútbol denominado *Comunio* (<http://www.comunio.es/>). Durante el curso 2012–2013, desarrollaron plataformas en línea para la venta de bienes y servicios con grandes descuentos, al estilo de las plataformas comerciales *Groupon* o *Planeo*. En el curso 2013–2014, desarrollaron plataformas web para la realización de trueques de objetos o servicios.

A continuación, en una segunda fase, cada equipo debía completar las fases de *ingeniería de requisitos* [10], *diseño arquitectónico* [14] y *diseño microarquitectónico* o *detallado* de un proyecto de desarrollo *software*. Para cada etapa, los alumnos debían entregar un informe de resultados de acuerdo a una plantilla predefinida.

Las fases de *construcción*, *pruebas* y *despliegue* se completaban en la asignatura Ingeniería del Software II, la cual complementaba a la primera. No obstante, dicha asignatura se impartía por un equipo docente distinto siguiendo un enfoque metodológico distinto, no basado en proyectos. Por esta razón, en este artículo nos centraremos en la asignatura de Ingeniería del Software I.

Dentro de cada etapa del portafolio, o del proceso de desarrollo *software*, cada grupo debía completar una serie de sub-tareas. Por ejemplo, dentro de la fase de *ingeniería de requisitos*, cada grupo debía ejecutar las siguientes acciones:

1. Elaborar un modelo de casos de uso de alto nivel;
2. Por cada caso de uso de alto nivel identificado, refinar dicho caso de uso como un conjunto de casos de uso de bajo nivel que representasen interacciones concretas de los actores con el sistema;
3. A continuación, cada caso de uso de bajo nivel debía especificarse de forma completa mediante plantillas [3];
4. Paralelamente a la descripción funcional del sistema, debía crearse un diagrama de clases que modelase el dominio del sistema;
5. Finalmente, había que realizar un análisis de la influencia de ciertos requisitos no funcionales sobre el sistema especificado.

La calificación final de la asignatura se determinaba por la media ponderada entre la calificación obtenida en el proyecto y la prueba final escrita de la asignatura. Los pesos de cada elemento oscilaban entre el 40 % y el 60 % en función del curso académico.

En la prueba final escrita el alumno debía realizar exactamente las mismas actividades que había realizado durante el desarrollo del proyecto. Además, el alumno debía responder una serie de cuestiones cortas destinadas a verificar que había asimilado ciertos principios teóricos. Las cuestiones cortas tenían un peso aproximado de 2.5 puntos sobre 10, mientras que la repetición de las actividades prácticas tenía un peso aproximado de 7.5 puntos.

En cualquier caso, en la prueba final escrita nunca preguntábamos nada que el alumno tuviese que memorizar. Además, permitíamos hacer uso de todo el material escrito que los alumnos considerasen necesario, incluidos los diferentes elementos del portafolio con las correcciones del profesor. El objetivo era que el alumno que había adquirido las competencias deseadas pudiese superar la prueba con facilidad, incluso sin dedicar tiempo alguno a la preparación de dicha prueba.

Esta idea está basada en el *principio de alineamiento constructivo de Biggs* [1], wue establece que lo que el alumno tiene que hacer para aprobar y aprender debe ser exactamente lo mismo. Por lo tanto, si diseñáramos la prueba final de manera correcta, el alumno que hubiese aprendido antes de la misma podía considerarse, sin miedo alguno de equivocarse, como aprobado.

Se puede argumentar, no sin razón, que aquellos alumnos que hubieran desarrollado con éxito el proyecto, y por tanto se les consideraba ya como aprobados, no había necesidad de realizarles la prueba final escrita. No obstante, mantuvimos la prueba final por dos razones:

1. Para verificar que el alumno no sólo había adquirido ciertas habilidades y destrezas, sino que además era capaz de ejercerlas dentro de un periodo de tiempo acotado. Es decir, se trataba de comprobar, por ejemplo, que el alumno no sólo era capaz de crear modelos del dominio de un sistema, sino que además era capaz de producirlos con una cierta agilidad.
2. Para verificar que el alumno era realmente el autor del trabajo entregado como resultado del proyecto. Por desgracia, no son residuales los casos en los cuales los alumnos acuden a academias para que les desarrollen el proyecto.

De forma complementaria al desarrollo del proyecto, otro de nuestros objetivos era enseñar a nuestros alumnos que el trabajo en equipo no implica necesariamente que todo el equipo tenga que reunirse físicamente en un mismo lugar y a una misma hora para poder trabajar juntos. Esta forma de trabajar en equipo es la que, por desgracia, traían asimilada la mayoría de ellos del instituto o de cursos inferiores. Sin embargo, esta no es la forma en la que se trabaja en equipo en las empresas de desarrollo *software*.

En dichas empresas se planifica el trabajo de forma que cada trabajador realice una serie de tareas de manera individual. El equipo sólo se reúne cuando es necesario coordinar actividades, alinear los objetivos del equipo o cuando la propia naturaleza de la actividad a realizar demanda de la presencia de varios miembros del equipo, tal como podría ser el caso de una *tormenta de ideas*.

Para que los alumnos aprendiesen a trabajar en equipo de esta nueva forma, al principio del curso impartíamos un pequeño seminario en el que explicábamos técnicas de planificación de forma que cada miembro de un equipo pudiese trabajar de forma individual y aislada, pero coordinada con sus compañeros, tal como se trabaja en una empresa real.

En las siguientes secciones describimos la evolución de la estrategia aplicada para calificar los informes de resultados entregados por los equipos, destacando las ventajas e inconvenientes de cada variación de la técnica.

3. Técnica 1: Café con leche para todos

La primera vez que aplicamos la metodología de aprendizaje basado en proyectos a la asignatura de Ingeniería del Software I fue durante el curso 2011–2012 [12]. Este primer año, debido fundamentalmente a nuestra inexperiencia, adoptamos la estrategia de calificación más simple: cada informe de resultados se evaluaba como un todo y se asignaba la misma calificación a todos los miembros del grupo. Además, cada grupo era responsable de organizarse internamente, dividiendo el trabajo entre sus miembros de la forma que considerase más conveniente.

Las desventajas de este método eran obvias. En primer lugar, la calificación de cada alumno no reflejaba su grado real de adquisición de competencias, sino el grado del grupo como un ente global e indivisible. Esto, como era de esperar, generaba quejas y distorsiones.

Por un lado, había alumnos que veían incrementada su calificación sin merecerlo, lo que se conoce como *calificación parásita* [2]. Por otro lado, aquellos alumnos muy capacitados o comprometidos con la asignatura se enfrentaban a la siguiente disyuntiva: realizar un sobreesfuerzo y asumir la parte del trabajo que sus compañeros no querían o no sabían realizar o asimilar que su calificación se vería reducida por el menor desempeño de sus compañeros.

Por ejemplo, en un caso extremo, un alumno asumió él solo la realización de todo el proyecto. No confiaba en las habilidades de sus compañeros, por lo que llegó incluso a negarles la participación en el proyecto con tal de no ver disminuida su calificación. Aún así, todos los alumnos recibieron la misma calificación. A pesar de no haber realizado apenas trabajo, algún alumno que se benefició del ingente trabajo de su compañero llegó incluso a considerar que la calificación recibida era fruto de su buen hacer. Esto es claramente contraproducente, pues genera en el alumno una visión distorsionada de su situación real.

Estas distorsiones se podían observar claramente en la diferencia entre las calificaciones del proyecto y de la prueba final escrita. Estas calificaciones, en el caso ideal, debían ser lo más parecidas posible, si aspirábamos a cumplir el principio de alineamiento constructivo de Biggs. Sin embargo, detectamos ciertos casos extremos. Por ejemplo, en un caso la calificación de un grupo fue de 7.69, mientras que uno de sus integrantes obtuvo un 2.70 en la prueba final escrita y otro ni siquiera llegó a presentarse a dicha prueba.

Existía una variación de esta técnica, aplicada por algunos compañeros de departamento, cuyo objetivo era reducir estas distorsiones. Dicha variación consistía en hacer que el equipo se pusiese de acuerdo para establecer qué porcentaje del mérito de un trabajo correspondía a cada miembro del equipo. Por ejemplo, en un equipo de cinco miembros, si un miembro había aportado más que el resto, podría atribuirse un 25 % del esfuerzo del trabajo, mientras que el resto se asignaría equitativamente un 18.75 %. Si todos habían colaborado equitativamente, se asignarían un 20 % cada uno.

A la hora de asignar las calificaciones, se multiplicaba la

calificación global del informe por el número de participantes del grupo. Luego se asignaba a cada alumno el porcentaje que le correspondiese del resultado de esa multiplicación. Por ejemplo, en el caso anterior, si el informe del grupo obtenía una calificación de 8, se multiplicaba por 5, que era el número de integrantes del grupo, obteniendo un global de 40 puntos a repartir. A continuación, se asignaba a cada alumno como calificación el porcentaje de esos 40 puntos que le correspondiese. Por ejemplo, al miembro que se le había asignado un porcentaje de esfuerzo del 25 % se le otorgaba una calificación de 10, mientras que para el resto de los miembros, con un porcentaje de 18.75 %, la calificación sería de 7.5.

Experiencias previas con esta técnica en otras asignaturas nos hicieron descartarla de inicio. Los problemas detectados eran básicamente los tres siguientes. Por una parte, los grupos en los que el porcentaje de esfuerzo no era el mismo para todos sus miembros eran poco frecuentes. Los equipos tendían a asignarse automáticamente porcentajes iguales sin detenerse a reflexionar sobre si tal asignación era o no justa. En segundo lugar, aquellos grupos en los cuales un miembro abandonaba se veían claramente favorecidos. Si en un grupo de 5 miembros, un alumno abandonaba la asignatura al poco de comenzar el cuatrimestre, el resto de sus miembros asumían su carga de trabajo, la cual no era excesiva, y también su porcentaje en la calificación final. Dicho porcentaje extra les servía para rebajar las exigencias para superar la asignatura.

Por ejemplo, en el caso de un grupo de 4 miembros en el cual abandonase uno, era suficiente, para obtener una calificación final de 5, con que los 3 miembros que permanecían en el grupo obtuviesen una calificación de 3.75 en el informe, asumiendo que se habían asignado un porcentaje de esfuerzo del 33.3 % cada uno. Por tanto, durante la formación del grupo, la presencia de un alumno con una alta probabilidad de abandono de la asignatura, como ciertos tipos de estudiantes Erasmus o alumnos con muchas asignaturas pendientes de otros cursos, era bastante bien recibida por el resto de compañeros.

Por último, el tercer problema asociado con esta variación de la estrategia era que los alumnos tendían a confundir *productividad* con *esfuerzo*, un mal endémico de la economía española que considerábamos necesario combatir. Los alumnos tendían a asignar los porcentajes de mérito en función de la horas trabajadas, no de los resultados obtenidos. Sin embargo, más horas trabajadas no tienen por qué implicar un mejor resultado o un trabajo de mayor calidad. Se trataba, y se trata, de obtener los mejores resultados con el menor esfuerzo posible.

Por tanto, dado que éramos concededores de estos problemas, decidimos no aplicar esta variación de la técnicas de *café con leche para todos* durante este primer año de aplicación de la metodología basada en proyectos.

Aparte del problema de las calificaciones distorsionadas, también pudimos observar que los equipos en su mayoría hacían caso omiso del seminario sobre trabajo en equipo impartido al principio del cuatrimestre. La mayoría de los grupos seguían reuniéndose, a una hora determinada y en un lugar

concreto, para poder trabajar de forma conjunta. Cuando faltaba alguno de los compañeros, el grupo simplemente se disolvía y no trabajaba.

Consecuentemente, la productividad de estos grupos era bastante baja. Muchos grupos seguían esgrimiendo la conocida excusa de «es que no podemos reunirnos para trabajar en el proyecto a casi ninguna hora porque se nos solapan los horarios». Por tanto, la formación en la competencia *trabajo eficaz en equipo* era claramente mejorable.

Detectadas estas deficiencias, decidimos que había que adoptar dos medidas: (1) calificar a los alumnos de manera individualizada, para evitar distorsiones y conflictos entre los miembros de un equipo; y (2) calificar de alguna forma la competencia trabajo eficaz en equipo, de forma que los alumnos no pudiesen obviarla. El siguiente apartado describe cómo adoptamos dichas mejoras para el siguiente curso académico.

4. Técnica 2: Grupos autoorganizados con tareas y calificaciones individuales

Para el curso 2012–2013 decidimos, en primer lugar, que si queríamos que los alumnos se tomasen en serio la competencia de *trabajo eficaz en equipo* necesitábamos calificar dicha competencia de alguna manera. Experiencias previas nos habían demostrado que, desafortunadamente, los alumnos tienden a ignorar aquellas actividades que no son calificables, aún a pesar de ser interesantes y relevantes para su aprendizaje [13].

Por tanto, decidimos que necesitábamos saber cómo se estaba organizando cada grupo para trabajar en equipo. Para ello hicimos que cada equipo en cada etapa del portafolio, además del informe de resultados, entregase un pequeño informe que indicase cómo se había organizado. Se valoraba que dicha organización contuviese una adecuada división de tareas, que la carga de trabajo fuese equitativa y que la planificación permitiese maximizar la productividad del grupo. Por ejemplo, se buscaba que no hubiese una excesiva dependencia entre tareas, de forma que no hubiese alumnos ociosos a la espera de que otros completasen sus tareas.

Además, para hacer que la competencia de trabajo eficaz en equipo fuese calificable decidimos que el informe que describía el plan de trabajo de cada etapa también fuese calificable. De esta forma, la calificación para cada etapa del portafolio venía determinada por la media ponderada de las calificaciones del informe de resultados y del plan de trabajo entregado, con pesos del 90 % y el 10 % respectivamente. Con esto se pretendía solucionar el problema de la formación en la competencia trabajo eficaz en equipo. Restaba solventar el problema de cómo calificar individualmente a los alumnos.

La solución a este segundo problema era sencilla. Al disponer del documento de la planificación, podíamos saber qué partes del informe de resultados había elaborado cada alumno. Por tanto, podíamos evaluar a cada alumno de forma indivi-

dualizada en función del trabajo personal que cada uno de ellos había realizado.

No obstante, no se trataba de que los alumnos trabajasen de forma completamente autónoma. Trabajar en equipo también implica coordinarse y complementarse, de forma que el todo adquiriera más valor que la simple suma de sus partes. Por ejemplo, cada equipo debía implementar técnicas de control de la calidad en la cual el trabajo individual de un alumno fuese supervisado por al menos uno de sus compañeros, con el objeto de detectar y eliminar posibles errores.

Por ello, para evitar que un alumno pudiese trabajar de forma completamente autónoma, desentendiéndose del resto de los compañeros, decidimos que la calificación del informe fuese de nuevo una media ponderada entre la calificación individual de cada alumno y la del informe como un todo, con pesos del 80 % y del 20 %, respectivamente.

Esta técnica basada en la división del trabajo y su posterior calificación individual permitía además resolver ciertos conflictos. Por ejemplo, si un alumno ofrecía un rendimiento extraordinariamente bajo, se indicaba al correspondiente equipo que no modificasen nada de lo que entregaba dicho compañero ni se preocupasen por hacerlo trabajar.

A la hora de calificar los informes, si se detectaba un alumno con un rendimiento notoriamente más bajo que el del grupo, se le calificaba de manera completamente individual, impidiéndole beneficiarse de la calificación del grupo. Con esto evitábamos las *calificaciones parásitas*. Además, el trabajo realizado por dicho alumno con rendimiento extraordinariamente bajo tampoco computaba para la calificación del informe como un todo. Por tanto, alumnos con un buen rendimiento no veían disminuida su calificación por el bajo desempeño o implicación de otros miembros del equipo.

Nuestra idea era que no podíamos exigirles a los alumnos aquello que nosotros mismos no éramos capaces de realizar. Así que dado que nosotros mismos nos confesábamos incapaces de hacer trabajar a ciertos tipos particulares de alumnos, entendíamos que no podíamos exigirles dicha tarea a sus compañeros. Por tanto, la solución pasaba por aceptar que dichos alumnos no iban a trabajar e intentar evitar que ello perjudicase a sus compañeros.

Aunque esta técnica de calificación mejoraba notoriamente la empleada en el curso anterior, aún adolecía de algunos problemas. En primer lugar, las organizaciones de trabajo que realizaban los alumnos solían ser bastante ineficientes. Dicha organización solía estar basada en las tareas a completar dentro de cada etapa, o en las secciones del informe de resultados a entregar.

Por ejemplo, si para la fase de ingeniería de requisitos, como ya hemos comentado (cf. Sección 2), había que realizar cinco tareas diferentes, cada uno de estos apartados se asignaba a un único alumno. De esta forma, un alumno se encargaba en exclusiva, por ejemplo, del análisis de los requisitos no funcionales, pero no especificaba ningún caso de uso mediante plantillas, ni participaba en la elaboración del modelo de dominio.

Esta división del trabajo, aparte de ser ineficiente, no fo-

mentaba el aprendizaje, pues un alumno sólo aprendía a realizar algunas de las tareas que teóricamente debía realizar. De hecho, algún alumno manifestó al término de la asignatura que la única plantilla de especificación de casos de uso que había realizado era la de la prueba final escrita, ya que durante el desarrollo del proyecto se había encargado de otras tareas.

Aunque en este fenómeno una gran parte de responsabilidad la tiene el propio alumno, pues de sobra sabe cuáles son los objetivos de aprendizaje de la asignatura y por tanto lo que se le va a exigir en la prueba final escrita, no es menos cierto que, al ser la primera vez que los alumnos realizan este tipo de tareas, les es difícil saber como dividir las y organizarlas de forma efectiva.

El siguiente apartado describe cómo resolvimos estos inconvenientes para el curso 2013–2014.

5. Técnica 3: Grupos preplanificados con tareas y calificaciones individuales

Las deficiencias detectadas en el curso 2012–2013 tenían una solución bastante trivial. La base del problema era que los equipos producían planificaciones de trabajo ineficientes las cuales, además, no favorecían el aprendizaje. Una fácil solución era que dichas planificaciones las proporcionase el propio equipo docente. Por tanto, para cada etapa del proyecto, entregábamos a los diferentes equipos un plan de trabajo predefinido, al cual debían adherirse.

Como ejemplo, mostramos el plan que se le entregaba para la fase de ingeniería de requisitos, compuesto por 5 tareas.

T1: Identificación de casos de uso de alto nivel. Entre todos los miembros del grupo se debían identificar los casos de uso de alto nivel del sistema mediante alguna dinámica de grupo, tipo tormenta de ideas.

T2: Identificación de los casos de uso de bajo nivel. A continuación se debían repartir de manera equitativa los casos de uso de alto nivel previamente identificados entre los diferentes miembros del grupo. Para cada caso de uso de alto nivel asignado cada miembro del equipo debía identificar los casos de uso de bajo nivel asociados.

T3: Especificación detallada de los casos de uso. Para cada caso de uso de bajo nivel identificado cada alumno debía especificar de forma detallada dicho caso de uso mediante una plantilla. Si algún alumno tuviese asociado un número de casos de uso de bajo nivel muy superior al de sus compañeros, se debía redistribuir la carga de trabajo para hacerla equitativa.

T4: Elaboración del modelo de dominio. Para cada caso de uso de alto nivel el alumno responsable debía elaborar un modelo de su dominio asociado. A continuación, entre todos los miembros del grupo, se debía componer

un modelo de dominio común a partir de los modelos de dominio parciales creados por cada alumno.

T5: Análisis de requisitos no funcionales. Finalmente, se repartían los requisitos no funcionales vistos en clase entre los miembros del grupo. Cada alumno se debía encargar de analizar la influencia sobre el sistema a desarrollar de los requisitos no funcionales que tuviese asignados.

De esta forma conseguíamos organizar el trabajo de cada equipo de manera que cada miembro pudiese trabajar de la forma más autónoma posible, con momentos puntuales en los que debía coordinarse con sus compañeros. Además, por el diseño de la planificación de trabajo, nos asegurábamos que cada alumno trabajase todas las habilidades que estaban incluidas en los objetivos de aprendizaje.

Durante el curso 2013–2014 también mantuvimos la calificación del informe como la media ponderada de la calificación del trabajo individual de cada alumno y la calificación de cada informe como un todo, con los mismos pesos del curso anterior. De esta forma, al igual que en el curso anterior, un alumno no podía desentenderse completamente del trabajo de sus compañeros, necesitando cooperar con ellos.

De manera similar, a aquellos alumnos que no realizaban su parte del trabajo con un mínimo de calidad se les aislaba de la calificación de sus compañeros, de forma que los alumnos no se viesen afectados por la presencia de un alumno con un rendimiento extraordinariamente bajo en su equipo. Esto, como ya comentamos anteriormente, ayudaba a solventar los conflictos asociados a los alumnos con muy bajo rendimiento.

Esta estrategia nos funcionó bastante bien durante el curso 2013–2014, por lo que, salvo pequeñas y sutiles mejoras, la consideramos como estable y la seguiremos aplicando en los próximos cursos. Como punto negativo comentar que esta estrategia disminuye el nivel que alcanzan los alumnos con respecto a la competencia de trabajo eficaz en equipo. Los alumnos ya no elaboran planificaciones para trabajo, sino que se limitan a ejecutar planificaciones prefijadas.

En cualquier caso, tal como hemos comentado en el apartado anterior, quizás el nivel exigido de inicio a los alumnos para esta competencia fuese excesivo. Dado que se trata de una asignatura de tercer curso, entendemos que no es malo trabajar esta competencia a un primer nivel básico, en el cual los alumnos se limitan a familiarizarse con el concepto de planificación de trabajo. Una vez adquirido este nivel básico, posteriormente, en alguna asignatura posterior, tal como Diseño Software o Procesos de Ingeniería Software, se podría continuar trabajando esta competencia, exigiendo ya que los alumnos elaboren sus propias planificaciones. Por tanto, se trataría de diseñar un mapa para el desarrollo de esta competencia a través de varias asignaturas, siguiendo un modelo similar al propuesto por Sánchez *et al.* [11].

El siguiente apartado describe diversos indicadores utilizados para medir el rendimiento de esta técnica, analizando si nos ha ayudado a conseguir los objetivos que perseguíamos.

6. Análisis de los resultados obtenidos

Para medir el nivel de eficacia de nuestra técnica utilizamos dos indicadores diferentes: (1) los resultados de una encuesta de satisfacción realizada a los estudiantes durante el curso 2013–2014; y (2) la comparación entre las calificaciones obtenidas en el proyecto y en la prueba final. Desafortunadamente, debido a la influencia de factores externos, consideramos los datos que aquí presentamos como no conclusivos.

6.1. Encuesta de satisfacción

Al término del curso 2013–2014 realizamos una encuesta de satisfacción general entre el alumnado. Dicha encuesta contenía diversas preguntas sobre el proceso de enseñanza-aprendizaje utilizado para la asignatura de Ingeniería del Software I. En dicha encuesta general incluimos durante el curso 2013–2014 una serie de preguntas específicamente destinadas a analizar la satisfacción de los estudiantes con técnica de calificación individual con grupos preplanificados. La encuesta se envió a los 45 alumnos matriculados, de los cuales 33 la completaron.

Las preguntas formuladas se muestran a la izquierda del Cuadro 1. Las preguntas se respondían con valores numéricos para una escala de *Likert* con valores entre 0 y 4, donde el 0 significaba *Nada*, el 1, *Poco*, el 2, *Normal*, el 3, *Bastante* y el 4, *Mucho*. La Figura 1 muestra los resultados obtenidos de forma gráfica, mientras que el Cuadro 1 muestra los resultados numéricos. La parte central del cuadro muestra el número de respuestas recibidas por cada valor numérico, mientras que en la derecha se muestran la media y la dispersión para cada pregunta.

En general, la técnica parece tener una aceptación positiva entre el alumnado. No obstante, es una aceptación positiva pero no entusiasta. Ello puede deberse en parte a la presencia de ciertos factores externos que pensamos que pueden haber distorsionado los resultados de esta encuesta. O podría deberse simplemente a que las mejoras introducidas no han tenido el impacto esperado, a pesar del esfuerzo realizado por el equipo docente.

En cualquier caso, merece la pena comentar que el curso, el 2013–2014 estuvo claramente afectado por la presencia en el mismo cuatrimestre y curso de lo que se denomina como una *asignatura tiburón* [6]. Una asignatura tiburón es aquella que plantea una carga tal de actividades de evaluación continua que hace que los alumnos abandonan o descuiden el trabajo en el resto de asignatura, centrándose sólo en ella. En nuestro caso, de 45 alumnos matriculados en la asignatura, sólo una media de 5 asistían a clase en presencia de una prueba parcial o entrega importante de prácticas de la asignatura tiburón.

Ello provocó que los resultados obtenidos por los alumnos en fases intermedias del proyectos fuesen bastante pobres en la mayoría de los casos. Dichos resultados generaron un descontento y malestar generalizados del alumnado con el profesorado de la asignatura. Pensamos, tras interpretar los resulta-

Pregunta	0	1	2	3	4	Promedio	Desviación
[P1] Prefiere el sistema de evaluación individualizada al grupal	1	7	10	7	8	2,42	1,16
[P2] Prefiere que se le proporcione una planificación predefinida	3	6	7	12	5	2,30	1,19
[P3] El plan le permite trabajar todas las habilidades requeridas	2	4	10	15	2	2,33	0,97
[P4] La carga de trabajo del plan es equilibrada	4	6	9	11	3	2,09	1,16
[P5] El plan le permite trabajar de forma coordinada pero equitativa	2	4	13	8	6	2,36	1,10

Escala Respuestas: (0) Nada (1) Poco (2) Normal (3) Bastante (4) Mucho
 Respuestas Recibidas: 33

Cuadro 1: Satisfacción del alumnado con la estrategia aplicada

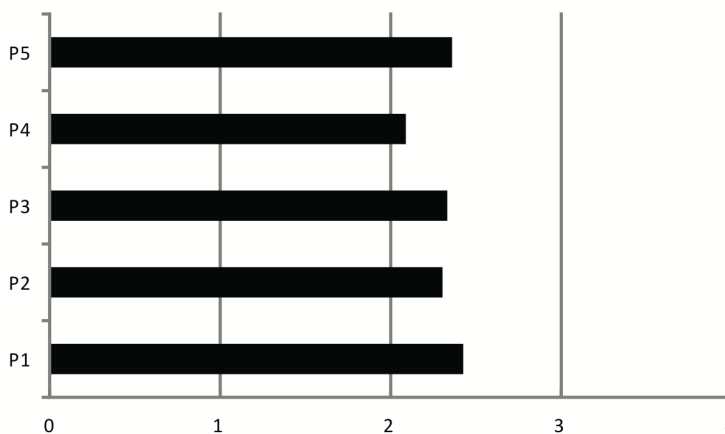


Figura 1: Resultados en escala *Likert* para la encuesta de satisfacción realizada

dos de la encuesta, que este clima pudo condicionar de algún modo las respuestas de los alumnos.

Por ejemplo, la pregunta *P3* podemos considerarla como una pregunta de control. Todos los planes de trabajo habían sido diseñados de forma que cada alumno cubriese todas y cada una las tareas que se le requerirían más tarde en la prueba final. Por tanto, ante la pregunta de si el plan de trabajo le permite ejercitar todas las habilidades que se le requerirán en la prueba final, las respuestas deberían ser todas *Bastante* (valor 3) o *Mucho* (valor 4), por la propia construcción de las planificaciones. Sin embargo, hay 2 respuestas con valor *Nada*, y el segundo valor más predominante es el *Normal*. Por tanto, entendemos que muchas respuestas han sido realizadas con cierto desdén o incluso rechazo a la asignatura y/o al equipo docente.

6.2. Similitud de las calificaciones del proyecto y la prueba final

El segundo elemento objetivo que utilizamos para medir la eficacia de nuestra técnica era la diferencia entre las calificaciones del proyecto y de la prueba final escrita. Entendíamos que si lo que se le exigía a los alumnos en ambas actividades era prácticamente idéntico, las calificaciones obtenidas en cada actividad debían de ser similares. Asumiendo que la calificación de la prueba final era objetiva, en el caso ideal, la

calificación del proyecto y de la prueba final debían ser prácticamente idénticas.

No obstante, hay argumentos razonables en contra de esta afirmación. Por un lado, es de esperar que los alumnos aprendan a partir de la retroalimentación recibida durante el desarrollo del proyecto. Por tanto, los alumnos aprenderían a subsanar ciertos errores cometidos durante el desarrollo del proyecto y nos los cometerían en la prueba final. Por tanto, la calificación de la prueba final debía ser superior a la del proyecto, si se ha producido cierto aprendizaje durante el desarrollo del proyecto.

Por otro lado, en la prueba final existen una serie de limitaciones de tiempo que eran inexistentes o despreciables durante el desarrollo del proyecto. Además, la realización de pruebas escritas provoca en ciertos alumnos situaciones de estrés y ansiedad que suelen repercutir negativamente en su calificación. Atendiendo a este razonamiento, la calificación de la prueba final debería ser menor que la del proyecto. Por tanto, entendemos que ambos fenómenos tienden compensarse dentro un grupo de alumnos heterogéneo. Consecuentemente, aunque haya diferencias a nivel individual, dichas diferencias deberían tener media cero.

Para el cálculo de la similitud entre las calificaciones del proyecto y la prueba final, eliminamos de cada actividad aquellas partes que no tenían un análogo en la otra actividad. De la calificación de la prueba escrita final sustraíamos la parte co-

respondiente a las preguntas cortas sobre cuestiones teóricas, normalizando el resultado para que el valor resultante quedase dentro del rango de 0 a 10. En la calificación del proyecto repetimos la misma operación con la idea de negocio, la cual no tenía un elemento análogo en la prueba escrita final.

Tras procesar de esta forma las calificaciones del proyecto y la prueba final, calculamos la diferencia entre ambas calificaciones. A continuación, para cada técnica aplicada calculamos la media de dichas diferencias, la diferencia máxima y el número de casos especiales. Por casos especiales entendemos aquellos en los cuales la calificación del proyecto es igual o superior a 7 y la calificación de la prueba final es inferior a 5. Es decir, alumnos con una calificación de notable en el proyecto que no superan la prueba escrita final.

La Figura 2 muestra la evolución de estos indicadores en función de la técnica aplicada. La media de las diferencias no parece disminuir ni incrementarse significativamente para ninguna de las tres técnicas. Por contra, la diferencia máxima, en contra de lo que cabría esperar, sube ligeramente para la técnica de los grupos autoorganizados. Finalmente, sí parece haber una mejora continuada en el número de casos especiales, los cuales se reducen a uno para la técnicas de los grupos autoorganizados.

La técnica de *café con leche para todos* se aplicó durante el curso 2011–2012 que corresponde al último curso de la titulación de Ingeniería Informática, antes de la implantación de los planes de grado. La técnica de calificación individual en grupos autoorganizados se utilizó durante el curso 2012–2013, correspondiente al primer año de la implantación del nuevo plan de Grado en Ingeniería Informática. Por tanto, se trata de un plan de estudios distinto al de la técnica anterior. En este curso se aplicó la técnica de autoorganización del trabajo y calificación individualizada. Este curso fue un curso especial, con un grupo de alumnos con un rendimiento extraordinariamente bajo en casi todas las asignaturas de la titulación, por lo que sus datos hay que interpretarlos con cautela, ya que suelen estar distorsionados. Por último, el curso 2013–2014 se corresponde al segundo año de la implantación del grado, y como ya se ha comentado, estuvo afectado por la presencia de una asignatura tiburón.

Por tanto, consideramos que hay demasiada influencia de factores externos para poder extraer conclusiones sólidas y definitivas de la comparación de estos tres cursos. De cualquier forma, nuestra percepción es que, aún mejorando la técnica de calificación, es complejo reducir la diferencia entre las calificaciones del proyecto y la prueba final, debido a dos fenómenos que hemos identificado informalmente, aunque no podemos cuantificar su influencia, y que describimos a continuación:

- **Efecto academia** Como ya hemos comentado, ciertos alumnos acuden a academias donde les realizan el proyecto. Por tanto, el alumno obtiene unos resultados que no se corresponden con sus habilidades reales. Por tanto, a la hora de realizar la prueba final las carencias en dichas habilidades queda patente, produciendo una di-

ferencia excesiva entre la calificación del proyecto y de la prueba final.

- **Efecto supervisión excesiva** Sin necesidad de acudir a la academia, el equipo docente, en ciertas ocasiones, proporciona una asistencia y retroalimentación excesivas al alumno. En estos casos, los comentarios, sugerencias e indicaciones del profesor prácticamente contienen la solución al problema que debe resolver el alumno. En estos casos, el alumno mejora los resultados, pero en muchas ocasiones no adquiere las correspondientes competencias. Por tanto, se produce la misma situación que en el fenómeno anterior.

El segundo fenómeno es evitable y estamos trabajando en fórmulas y soluciones para proporcionar retroalimentación al alumno de forma que éste no pueda llegar a una solución correcta sin adquirir las competencias necesarias para ello. El primer fenómeno nos parece más difícil de evitar, siendo la principal causa por la que somos muy reticentes a eliminar la prueba final de la asignatura.

7. Sumario y trabajos futuros

Este artículo ha presentado la evolución a lo largo de tres cursos académicos de la técnica de calificación de una serie de informes producidos por equipos de alumnos, dentro de una metodología de aprendizaje basada en proyectos. La evolución de dicha técnica se ha centrado en encontrar una fórmula que permita que: (1) los alumnos trabajen en equipo pero puedan ser calificados de forma individual, conforme a las competencias adquiridas por cada uno de ellos; (2) permita formar a los alumnos en técnicas de trabajo eficaz en equipo, mostrándoles, entre otras cosas, que trabajar en equipo no implica necesariamente trabajar todos juntos, a una misma hora y en un mismo espacio físico.

El primer año se aplicó la técnica más simple: se otorgaba una calificación al grupo, y dicha calificación era la misma para todos sus miembros. En este caso, la calificación muchas veces no reflejaba las habilidades reales del alumno. Además, la competencia de trabajo eficaz en equipo no era ni ejercitada ni evaluada.

El segundo año cada grupo debería elaborar una planificación que dividiese las tareas a realizar. Siguiendo esta división, se evaluaba a cada alumno de forma individual por las tareas que había realizado. Además, se calificaba el plan de trabajo. En este caso, observamos que los planes de trabajo realizados eran ineficientes y además no aseguraban que un alumno trabajase todas las competencias que se le iban a demandar en la prueba final.

Por tanto, el tercer año se optó por proporcionar un plan de trabajo predefinido, evitando que tuviesen que realizarlo ellos.

Como trabajo futuro esperamos obtener mejores datos empíricos que nos permitan estudiar mejor el rendimiento de la técnica empleada. Además pretendemos implantar algún mecanismo de análisis interno del rendimiento del grupo que nos

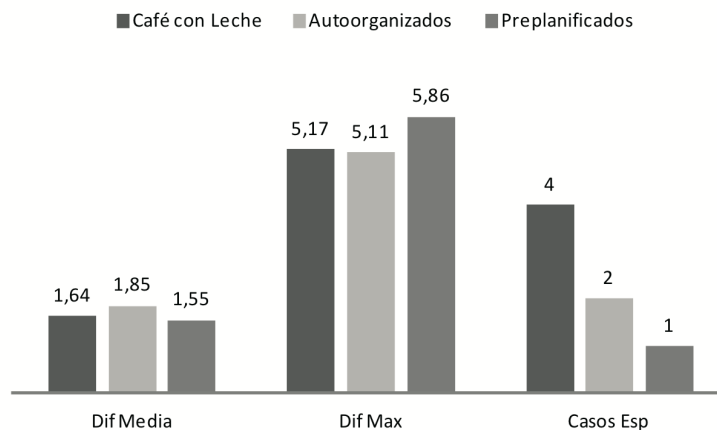


Figura 2: Indicadores extraídos de las diferencias entre calificaciones del proyecto y la prueba escrita final

permita evaluar el grado de compromiso de cada alumno con respecto a sus compañeros, al estilo de [6, 16]

Referencias

- [1] John Biggs y Catherine Tang: *Teaching for quality learning at university*. McGraw Hill, 4ª edición, 2011.
- [2] Sally Brown y Peter Knight: *Assessing learners in higher education*. Routledge. 1998.
- [3] Alistair Cockburn: *Writing effective use cases*. Addison-Wesley. Octubre de 2000.
- [4] José Antonio Cruz-Lemus, Marcela Genero, Marta N. Gómez y Silvia T. Acuña. *Formación de equipos de trabajo basada en factores de la personalidad de los integrantes: Un estudio empírico*. En Actas de las XVIII Jornadas sobre la Enseñanza Universitaria de la Informática, Jenui 2012, pp. 97–104. Ciudad Real, España. Julio de 2012.
- [5] José A. Cruz-Lemus, Marcela Genero, y Marta N. Gómez y Silvia T. Acuña: *Formación de equipos en función de la extroversión de sus integrantes: Un experimento replicado*. En Actas de las XIX Jornadas sobre la Enseñanza Universitaria de la Informática, Jenui 2013, pp. 135–142. Castellón de la Plana, España. Julio de 2013.
- [6] Manuel Enciso, Carlos Rossi y Eduardo Guzmán. *Evaluación del trabajo en grupo: ¿café para todos?* En Actas del Simposio-Taller de las XVII Jornadas de Enseñanza Universitaria de la Informática, pp. 91–98, Sevilla, España. Julio de 2011.
- [7] Max Lüscher: *The Lüscher Color Test*. Pocket. Septiembre de 1990.
- [8] Thom Markham, John Larmer y Jason Ravitz: *Project Based Learning handbook: A guide to standards-focused Project Based Learning for middle and high school teachers*. Buck Institute for Education, 2ª edición. Mayo de 2003.
- [9] Mercedes Marzo Navarro, Marta Pedraja Iglesias, y Pilar Rivera Torres: *Las competencias profesionales demandadas por las empresas: el caso de los ingenieros*. Revista de Educación, núm. 341, pp. 643–661. Septiembre-diciembre de 2006.
- [10] Klaus Pohl: *Requirements Engineering: Fundamentals, Principles, and Technique*. Springer. Julio de 2010.
- [11] Fermín Sánchez, Alicia Ageno, Lluís Belanche, Jose Cabré, Erik Cobo, Rafel Farré, Jordi García, David López, Pere Marés, Carme Martín, y Antonia Soler: *Desarrollo integral de las competencias genéricas mediante mapas competenciales*. En Actas de XVIII Jornadas de Enseñanza Universitaria de la Informática, Jenui 2012, pp. 185–192. Ciudad Real, España. Julio de 2012.
- [12] Pablo Sánchez y Carlos Blanco: *Experiencias y Lecciones Aprendidas de la Implantación de una Metodología de Aprendizaje basada en Proyectos para una Asignatura de Ingeniería del Software I*. En Actas de XVIII Jornadas de Enseñanza Universitaria de la Informática, Jenui 2012, pp. 41–48. Ciudad Real, España. Julio de 2012.
- [13] Pablo Sánchez y Carlos Blanco. *Una Metodología para Fomentar el Aprendizaje mediante Sistemas de Evaluación entre Pares*. En Actas de XVIII Jornadas de Enseñanza Universitaria de la Informática, Jenui 2012, pp. 37–44. Ciudad Real, España. Julio de 2012.
- [14] Richard N. Taylor, Nenad Medvidovic y Eric M. DASHOFY: *Software architecture: Foundations, theory, and practice*. Wiley. Enero de 2009.

- [15] Miguel Valero-García y Javier García Zubia: *Cómo empezar fácil con PBL*. En Actas de XVII Jornadas de Enseñanza Universitaria de la Informática, Jenui 2011, pp. 109–116. Sevilla, España. Julio de 2011.
- [16] Luis Velasco y Carlos Villavieja: *Cómo evaluar continua e individualmente en asignaturas basadas en proyectos*. En Actas de XV Jornadas de Enseñanza Universitaria de la Informática, Jenui 2009, pp. 347–354. Barcelona, España. Julio de 2009.
- [17] Michael A. West: *Effective teamwork: Practical lessons from organizational research*. Wiley-Blackwell. Diciembre de 2003.
- [18] Ann Bullock y Parmalee P. Hawk: *Developing a teaching portfolio: A guide to preservice and practicing teachers*. Prentice Hall. Julio de 2004.
- [19] Gary A. Yukl: *Leadership in organizations*. Prentice Hall. Enero de 2012.



Pablo Sánchez es Profesor Contratado Doctor del Departamento de Ingeniería informática y Electrónica de la Universidad de Cantabria. Dentro del campo de la investigación, sus trabajos se centran en el área de las líneas de productos software y el desarrollo software dirigido por modelos, con trabajos que intentan combinar ambos paradigmas.

En el ámbito educativo su interés se centra en la aplicación de la metodología de aprendizaje basado en proyectos a materias de Ingeniería del Software. Los resultados de este trabajo pueden encontrarse en las actas de diversas ediciones de JENUI entre 2011 y 2014. Par más información, consultar <http://personales.unican.es/sanchezbp/>.



2014 P. Sánchez. Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial-SinObraDerivada 4.0 Internacional 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 y no se haga un uso comercial.