



UNIVERSIDAD CARLOS III DE MADRID

Revisión entre pares como instrumento de aprendizaje. Una experiencia práctica



Raquel M. Crespo García
Julio Villena Román

working
papers



UNIVERSIDAD CARLOS III DE MADRID

<http://biblioteca.uc3m.es/WORKINGPAPERS>
Serie de Innovación Docente
nº 05-03-01
2005



Revisión entre pares como instrumento de aprendizaje Una experiencia práctica

Raquel M. Crespo García, Julio Villena Román
{rcrespo, jvillena}@it.uc3m.es
Departamento de Ingeniería Telemática
Universidad Carlos III de Madrid

Marzo 2005

Resumen

Este artículo describe la experiencia de innovación docente llevada a cabo este último curso basada en la aplicación de la metodología de revisión entre iguales como instrumento para el aprendizaje, desarrollada en las asignaturas Organización de Contenidos Audiovisuales (Ingeniería Técnica de Telecomunicación, especialidad Sonido e Imagen) durante el curso 2003-04 e Inteligencia en Redes de Comunicaciones (Ingeniería de Telecomunicación) durante el curso 2004-05. Los experimentos realizados muestran la utilidad de la metodología de revisión entre pares como instrumento para el aprendizaje, percibido de forma subjetiva por los alumnos participantes y demostrado empíricamente en un experimento de control mediante preguntas test, descrito en este documento.

Palabras clave: Co-evaluación formativa, evaluación entre iguales, evaluación entre pares

Keywords: Peer review, peer grading

1. Introducción

La revisión entre pares consiste en evaluar el trabajo de un compañero, con similar nivel de conocimientos y de experiencia, por lo que también se denomina revisión entre iguales. En el ámbito educativo, se define como “aquel proceso de evaluación donde los individuos juzgan la cantidad, nivel, valor, calidad o éxito de los productos o resultados del aprendizaje de compañeros de un nivel similar” [1].

Esta metodología estructura un ciclo de desarrollo revisión del trabajo en cuatro fases, tal como se ilustra en la Figura 1.

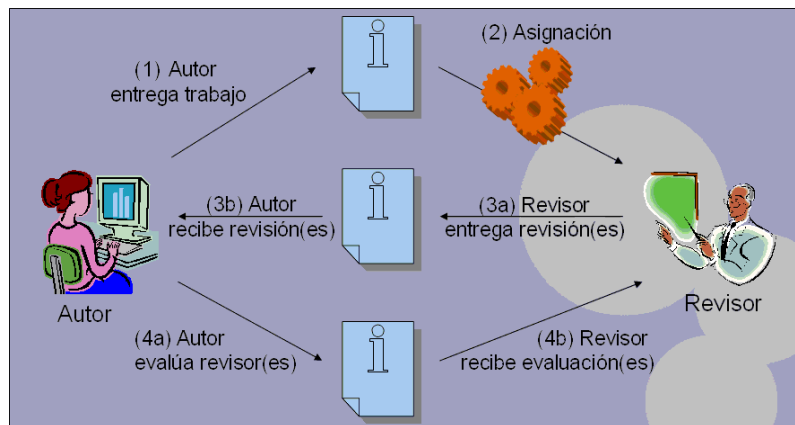


Figura 1: Proceso de revisión

1. Desarrollo del trabajo

Los alumnos realizan un determinado trabajo. Puede tratarse de un estudio teórico, una aplicación práctica, etc. Asimismo, puede realizarse individualmente o en grupos.

2. Asignación de revisores

Se asigna un conjunto de revisores a cada trabajo. El número de revisores, así como el criterio de asignación, son parámetros a decidir por el profesor, dependiendo del contexto. Se trata de una fase oculta, transparente desde el punto de vista de los alumnos.

3. Revisión de trabajos de otros compañeros

Cada revisor debe evaluar los trabajos que le hayan sido asignados. Para ello, es habitual que se le proporcione un formulario como guía, que deberá rellenar valorando y comentando distintos aspectos del trabajo. Dichos aspectos dependerán del tipo de trabajo propuesto (por ejemplo, variarán si se trata de un documento o una aplicación). Conviene que sean definidos a priori por el profesor, en función de los objetivos didácticos perseguidos por el trabajo planteado. La revisión puede realizarse de forma individual o discutirse en equipo, y suele ser anónima¹.

4. Valoración de las revisiones recibidas

Por último, puede existir, de forma opcional, una fase final en la que cada autor valora la realimentación recibida. Básicamente, su utilidad, su conformidad con los comentarios recibidos, si las explicaciones eran claras, si la revisión había sido exhaustiva... Estas valoraciones habitualmente se realizan de forma individual si incluyen parámetros subjetivos (como por ejemplo la utilidad), aunque depende del contexto concreto en que se aplique.

Este ciclo puede repetirse las veces que se considere oportuno a lo largo del curso. Desde el punto de vista de los alumnos, sólo se perciben tres fases en el proceso (entrega de

¹ En grupos reducidos, garantizar el anonimato es una cuestión compleja, puesto que a menudo se conocen los trabajos realizados por cada alumno.

trabajos, revisión y evaluación de revisiones), puesto que la etapa de asignación de revisores es oculta y transparente para ellos.

Los beneficios que se pretenden obtener con esta metodología son múltiples y de diversa naturaleza. A continuación se citan los objetivos principales al abordar este proyecto:

a) Objetivos específicos (de la materia en estudio)

- Mayor motivación de los alumnos.
- Fomentar el análisis en profundidad de la materia tratada.
- Fomentar el trabajo continuo de los alumnos, como medio para conseguir un aprendizaje de calidad.
- Revisión de otras soluciones, ideas, planteamientos... alternativos.
- Proporcionar realimentación a los alumnos a lo largo del curso sobre el trabajo desarrollado, con el fin de detectar y resolver dudas inmediatamente.
- En definitiva, afianzar los conocimientos adquiridos por el alumno y profundizar en los conceptos de la asignatura mediante su aplicación práctica y análisis crítico.

b) Objetivos transversales

- Fomentar la asunción de un rol más activo, autónomo y responsable por parte del alumno en el proceso de aprendizaje y evaluación.
- Promover una formación integral del alumno, potenciando el desarrollo de habilidades no técnicas imprescindibles, como:
 - trabajo en equipo (si alguna de las fases se desarrolla en grupo),
 - capacidad de expresión
 - capacidad de síntesis,
 - extracción de las ideas fundamentales,
 - articulación de las propias ideas,
 - razonamiento y justificación de las mismas,
 - pensamiento crítico,
 - incluso simplemente su redacción y ortografía.

c) Objetivos docentes

Desde el punto de vista docente, podría parecer que esta metodología puede ayudar a reducir la carga docente asociada a la corrección de las prácticas. Sin embargo, como se verá más adelante, su uso como herramienta de evaluación debe abordarse con extraordinaria precaución.

Se trata de una metodología genérica, de modo que es aplicable en cualquier tipo de asignatura. La revisión entre pares [1] tiene una larga historia de aplicación en ámbitos de humanidades [2][3][4][5] (escritura, psicología...). Pero también se ha utilizado en entornos científicos [6] y tecnológicos [7][8][4][5] y, en el caso de programación [9][10][11][12], resulta particularmente interesante por las ventajas inherentes en este contexto.

En el presente artículo se describe una experiencia de aplicación de la metodología de revisión entre pares como instrumento de aprendizaje (también denominada co-evaluación formativa por algunos autores [13]), proyecto de innovación docente desarrollado durante el curso 2003-04 en la asignatura *Organización de Contenidos Audiovisuales* y durante el curso 2004-05 en la asignatura *Inteligencia en Redes de Comunicaciones*. En los siguientes apartados describiremos las experiencias desarrolladas en cada una de las asignaturas, y, por último, las conclusiones extraídas.

2. Caso de estudio I: OCA

La experiencia descrita en este apartado se ha desarrollado durante el segundo cuatrimestre de 2003-04, en la asignatura *Organización de Contenidos Audiovisuales*. No se dispone de antecedentes ni información previa de la asignatura puesto que es la primera vez que se imparte, ya que es cuando comienza la titulación de *Ingeniería Técnica de Telecomunicación, especialidad Sonido e Imagen*, a la que pertenece. Se trata de una asignatura obligatoria de primer curso. Es de 6 créditos (3 de teoría y 3 de prácticas) y se imparte en dos sesiones semanales de 100 minutos de duración cada una. Cada semana se dedica una sesión a los aspectos teóricos de la asignatura y en la otra se abordan aspectos prácticos.

Cuadro 1: Organización de Contenidos Audiovisuales

Asignatura	Organización de Contenidos Audiovisuales (OCA)
Curso	2003-04
Titulación	Ingeniería Técnica de Telecomunicación (esp. Sonido e Imagen)
Tipo	Obligatoria
Créditos	6 (3 teoría + 3 prácticos)
Cuatrimestre	2º

Los objetivos principales de la asignatura son:

- Fundamentos teóricos de multimedia, diseño y organización de contenidos.
- Introducción al desarrollo de sistemas multimedia con Java:
 - Afianzar conocimientos sobre el lenguaje Java y la programación orientada a objetos adquiridos en el primer cuatrimestre.
 - Familiarizar al alumno con técnicas de diseño y pruebas de programa.
 - Interfaces gráficas con Java.
 - Todo ello orientado al diseño y desarrollo de interfaces gráficas y aplicaciones multimedia.

La parte práctica de la asignatura se desarrolla en el laboratorio, una sesión cada semana a lo largo del cuatrimestre. Se organiza en tres grandes bloques de contenidos:

1. Interfaces gráficas de usuario con Java: repaso de conceptos básicos de programación y Java e introducción al desarrollo de aplicaciones gráficas en este lenguaje.
2. Aplicaciones multimedia con Java: incorporación de elementos multimedia en las aplicaciones, presentación de audio y vídeo con Java.
3. Formatos de representación: introducción a XML.

2.1. Motivación

La experiencia desarrollada se apoya en dos líneas fundamentales de actuación: realización de trabajos en grupo, concretamente desarrollo de aplicaciones y applets Java, complementados con el análisis crítico de trabajos desarrollados por otros compañeros.

Las prácticas tradicionales, que se corresponderían con la primera línea de actuación, cubren la mayoría de los objetivos planteados, pero no todos. La metodología propuesta incorpora mecanismos de revisión de código de los propios compañeros, que tienen un nivel similar, lo cual redundará en una mayor motivación, fomenta el trabajo reflexivo frente al riesgo de repetición mecánica de procedimientos y potencia el desarrollo de habilidades no técnicas fundamentales para el futuro profesional del alumno.

Pese a ser este curso el primero en impartirse la asignatura, esta iniciativa surge motivada por la experiencia de los profesores en asignaturas similares. Los conocimientos procedimentales son objetivo prioritario en asignaturas de programación. Por ello, suelen articularse en torno a la realización en el laboratorio de prácticas semanales, para facilitar la asimilación de los conceptos teóricos, y la entrega de una práctica global evaluable con un peso importante en la calificación.

En dichos entornos se han detectado una serie de características y efectos que desvirtúan la eficacia del proceso de aprendizaje, principalmente:

- Desmotivación de los alumnos, principalmente respecto a las prácticas semanales. Al no ser evaluables, los alumnos no dedican el esfuerzo necesario, sino que centran su atención en la práctica global, perdiéndose en buena medida su eficacia docente.
- Trabajo colaborativo mal entendido. Los alumnos, especialmente en los primeros cursos, no perciben claramente el objetivo didáctico de la metodología propuesta ni, por tanto, su importancia. Concretamente, la frontera entre cooperación y ayuda entre compañeros y realizar la práctica en grupo les resulta muy difusa. Prácticas dimensionadas para realizarse en parejas a menudo se realizan en grupos de 2 ó más parejas. Viceversa, es frecuente que los alumnos de un equipo se repartan el trabajo

encargándose uno de una asignatura y otro de otra, con lo cual en realidad no existe colaboración, aparte de que algunos no realicen el trabajo.

- Los alumnos limitan a menudo su percepción de las prácticas evaluables a su dimensión evaluativa, obviando su dimensión formativa. En consecuencia, su objetivo es alcanzar el resultado final y obtener la nota en vez del propio proceso de desarrollo. Lo cual lleva a problemas de plagio e incluso prácticas realizadas por otros (profesores particulares, academias, familiares, amigos, etc.). En consecuencia, se pierde la componente didáctica que constituye la razón de ser de dichas prácticas evaluables, generándose además conflictos en el proceso de evaluación.

Esta experiencia pretende transformar dichas características, originalmente nocivas, en ventajas y apoyos que potencian al aprendizaje. Así hay una parte de trabajo en equipo y una parte individual (que obviamente resultará más sencilla si se ha realizado correctamente la primera). Igualmente, ya que los alumnos tienden a observar el trabajo de sus compañeros, se fomenta que lo hagan, pidiéndoles que lo revisen seriamente, en vez de intentar evitarlo.

Ayuda al alumno en la organización y planificación del estudio, fomentando el trabajo continuo y el razonamiento sobre los conceptos tratados en clase para una mejor comprensión de los mismos. Se basa asimismo en la experiencia de que a programar se aprende programando, pero también leyendo y analizando código ya desarrollado[14]. Por otra parte, recibir realimentación sobre el trabajo realizado durante el curso permite detectar errores y subsanarlos.

Un fallo que se ha detectado con frecuencia en los alumnos es que una vez consiguen resolver un problema no contemplan posibles alternativas. Analizar soluciones ajenas probablemente les enfrente con esas posibles alternativas que ellos no se plantean. Asimismo, otro fallo muy común es que no saben enfrentarse a sus errores. Cuando se produce un fallo rara vez intentan resolverlo de forma sistemática, intentando aislar las posibles causas para facilitar su corrección. En su lugar, demasiado a menudo intentan resolverlo mediante la estrategia de prueba y error, llegando incluso a intentar variaciones aleatorias (que lógicamente suelen perjudicar más que ayudar). La consecuencia es que si en algún momento un programa de cierta complejidad funciona, a menudo no tienen claro el proceso ni entienden exactamente el porqué. Analizar código desarrollado por otras personas, con el consiguiente esfuerzo de comprensión, ayuda a afianzar los conocimientos procedimentales más allá de la pura repetición mecánica.

La metodología adoptada estimula, además, la curiosidad del alumno e incentiva una sana competitividad, enfrentando al alumno a retos asequibles, puesto que han sido superados por compañeros de su nivel.

Por último, permite facilitar a los alumnos la oportunidad de desarrollar habilidades transversales, no técnicas, habitualmente descuidadas en este tipo de estudios, y hacer hincapié en su importancia, ya que son fundamentales para un ingeniero profesional. La capacidad de expresión escrita es básica en prácticamente cualquier profesión. Así como la capacidad de análisis crítico del trabajo desarrollado por otros colegas; en programación, concretamente, la prueba de programas es una actividad fundamental a la que los alumnos no suelen conceder la importancia debida.

2.2. Desarrollo de la experiencia

Las sesiones del laboratorio desarrollan prácticas guiadas para familiarizar a los alumnos con los temas fundamentales tratados. Sin embargo, dada la naturaleza de sus contenidos, la asignatura es idónea para desarrollar la iniciativa y creatividad de los alumnos, permitiéndoles idear sus propios proyectos, lo cual redundará en una mayor motivación e interés.

Partiendo de esta idea, se propuso a los alumnos la realización de tres prácticas de carácter voluntario, para afianzar las principales líneas temáticas de la asignatura:

1. Aplicación gráfica: juego [15].
2. Applet interactivo: mapa [16].
3. Applet multimedia [17].

Los enunciados se plantearon abiertos, para respetar la iniciativa y creatividad de los alumnos. Se incluyeron una serie de requisitos mínimos, básicamente, los elementos vistos en

clase que debían incluirse en el trabajo, para asegurar su valor didáctico. Y se sugería una posible aplicación, como guía. El alumno podía desarrollar esta idea, o bien, otra temática que se ajustase mejor a sus intereses.

Por ejemplo, en el caso de la aplicación gráfica se pidió la realización de un juego, sugiriéndose concretamente el de las tres en raya por su simplicidad. La mayoría de los alumnos optó por seguir esta sugerencia, probablemente por las pautas que se facilitaban en el enunciado sobre cómo enfocarlo. Sin embargo, una parte significativa desarrolló además una aplicación alternativa: puzzle, adivinación de números...

Siguiendo el proceso de revisión descrito anteriormente (Figura 1), cada práctica se dividía en tres fases (desde el punto de vista de los alumnos):

1. Desarrollo de la aplicación.

En parejas (algunos alumnos trabajaron de forma individual y excepcionalmente un grupo constaba de tres personas), los alumnos desarrollaban una aplicación conforme a los requisitos pedidos.

2. Revisión de aplicaciones de otros compañeros.

A cada práctica se asignaban tres revisores, cada uno de los cuales debía rellenar un formulario valorando distintos aspectos de la aplicación [18]. La revisión se hacía de forma individual y, por supuesto, anónima².

3. Valoración de las revisiones recibidas.

Por último, cada autor podía rellenar otro formulario valorando las evaluaciones recibidas [19]. Básicamente, la utilidad de las mismas, su grado de conformidad y una valoración de la revisión en sí: si las explicaciones eran claras, si la revisión había sido exhaustiva... Estas valoraciones se rellenaban de forma individual, al igual que las revisiones, independientemente de si el trabajo se había hecho en grupo o no.

Cada formulario incluía además un apartado en el que el alumno debía indicar en qué grado le había resultado útil en su aprendizaje, acompañado de un campo de texto en el que podía describir los puntos más destacados. Asimismo, se incluía también un apartado abierto mediante el cual los alumnos podían hacer llegar comentarios y sugerencias libremente a los profesores.

2.3. Resultados

Dada la carga implicada en el desarrollo de las prácticas, era razonable reflejar el trabajo realizado en la nota final de la asignatura. Cada práctica podía sumar hasta 0.5 puntos a la nota del examen práctico, sumando un máximo de 1.5 en total. La puntuación obtenida en cada práctica se calculaba como la media obtenida por el alumno como autor y como revisor. Pese al escaso peso asignado, acorde con la prioridad concedida al valor didáctico del proceso frente a su dimensión evaluativa, el grado de participación superó nuestras expectativas. La Figura 2 resume la participación de los alumnos en cada fase de los tres ciclos propuestos.

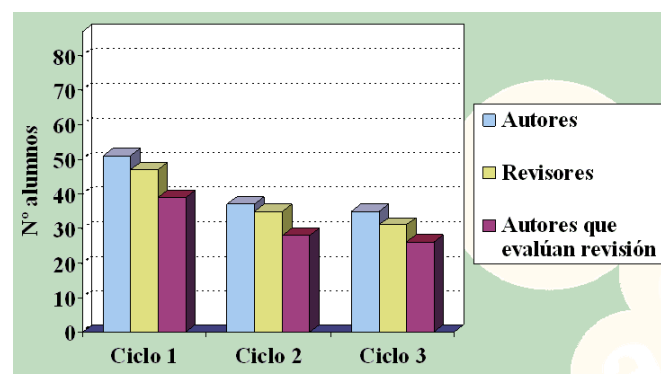


Figura 2: Participación de los alumnos

² En grupos reducidos como éste, garantizar el anonimato es una cuestión compleja, puesto que las aplicaciones más llamativas a menudo podían reconocerse.

En una primera impresión puede parecer que la participación ha sido escasa: menos de la mitad de los alumnos matriculados. Sin embargo, analizando el contexto en que se desarrolla la experiencia, los datos adquieren un significado distinto. Esta asignatura se apoya en gran medida en los conocimientos de programación adquiridos en el primer cuatrimestre. Se trata de una materia en la que habitualmente los alumnos encuentran serias dificultades y a menudo abandonan. Una vez abandonada en el primer cuatrimestre, o bien, si los conocimientos no son suficientemente sólidos, son raros los alumnos que realmente continúan con la materia en el segundo cuatrimestre. En consecuencia, un alto porcentaje abandona la asignatura sin ni siquiera intentarla. Medida respecto al grupo de alumnos que realmente abordan su estudio (mediante indicadores como la asistencia a clase o el número de alumnos presentados) la participación resulta muy significativa.

La escasa familiaridad de los alumnos con este tipo de metodologías se refleja en la diferente atención que prestaron a cada fase del proceso. En general, dedicaron un esfuerzo considerable a la realización de los trabajos propuestos. Sin embargo, las revisiones no siempre fueron igual de exhaustivas. Además, hay un porcentaje, afortunadamente pequeño, de alumnos que realizaron el trabajo pero no entregaron la revisión asignada, pese a requerir un esfuerzo mucho menor que el trabajo previo. Hay otro porcentaje significativo que simplemente olvidó enviarlas a tiempo, pero las entregó inmediatamente cuando se les recordó.

Es interesante notar también la evolución a lo largo del curso de la calidad del trabajo realizado por los alumnos. En general, el nivel de los trabajos prácticos se mantuvo más o menos constante. Mientras que la calidad de las revisiones disminuyó en la última fase. Así, las revisiones de la primera entrega fueron (siempre en general) muy detalladas; mientras que en la tercera, los comentarios fueron mucho más breves o incluso inexistentes. La carga de trabajo de los alumnos, muy próximos ya a las fechas de los exámenes; cansancio y, por tanto, menor entusiasmo; así como un menor dominio de la materia final, más compleja, son las hipótesis que barajamos como posibles causas de esta degradación.

Un resultado particularmente interesante es la relación que aparece entre el nivel de la práctica revisada y lo que el alumno aprende de ella (según encuesta subjetiva), reflejada en la Figura 3. Aunque es necesario realizar más estudios para confirmar esta hipótesis, parece razonable que los alumnos aprendan más de los mejores trabajos.

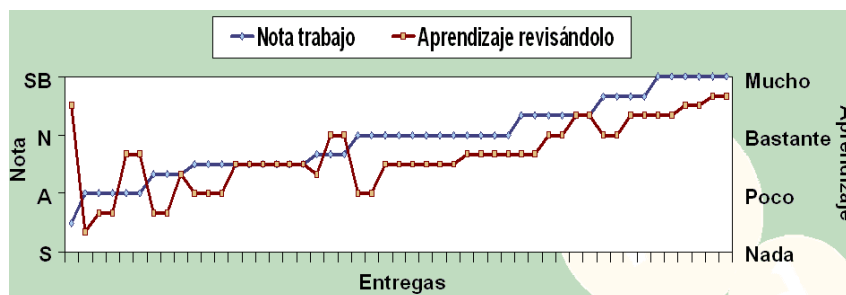


Figura 3: Nota del trabajo vs. Aprendizaje

Respecto a la metodología en sí, hemos comprobado que las revisiones suponen una carga razonable, especialmente después de haber realizado un trabajo similar. El tiempo requerido, según estimaciones de los alumnos, variaba desde los 20 minutos a la hora. Aunque el tiempo dedicado claramente se reflejaba en la calidad de la revisión.

3. Caso de estudio II: IRC

La asignatura de *Inteligencia en Redes de Comunicaciones* se imparte en 5º curso de *Ingeniería de Telecomunicación*, optativa de la especialidad Sistemas y Redes de Telecomunicaciones. Es una asignatura de 4,5 créditos (3 teoría y 1,5 prácticas) que se imparte en dos sesiones semanales, de 50 y 100 minutos de duración respectivamente, en horario de tarde.

Cuadro 2: Inteligencia en Redes de Comunicaciones

Asignatura	Inteligencia en Redes de Comunicaciones (IRC)
Curso	2004-05
Titulación	Ingeniería de Telecomunicación
Tipo	Obligatoria
Créditos	4,5 (3 teoría + 1,5 prácticos)
Cuatrimestre	1º

En los últimos años, el número de alumnos ha ido incrementándose desde los 11 alumnos del curso 2002-03, pasando por 46 alumnos el curso 2003-04, hasta los 51 alumnos del curso 2004-05, que la sitúan en la titulación, según el número de alumnos, como la 1ª asignatura de especialidad de 5º del bloque A (de un total de 5 asignaturas) y la 3ª entre las asignaturas de especialidad de primer cuatrimestre de 5º (de 10 asignaturas).

Al preguntar a los alumnos por los motivos de su elección de esta asignatura, la respuesta más frecuente es que lo hacen fundamentalmente atraídos por el temario, que resulta sustancialmente diferente al resto de las asignaturas de la carrera. Otro motivo es por recomendación de alumnos de años anteriores. Aunque la alta tasa de aprobados puede ser también un factor de elección de la asignatura, no es el factor determinante puesto que esta tasa es similar a la de algunas otras asignaturas optativas de especialidad.

El objetivo de la asignatura es analizar el concepto de inteligencia en los sistemas informáticos y de comunicaciones y estudiar las técnicas para incorporar estos comportamientos en redes y servicios. Es decir, proporcionar al alumno los fundamentos básicos de la Inteligencia Artificial y las áreas donde estas tecnologías pueden proporcionar avances significativos.

La asignatura está planteada de modo que durante las clases se exponen a los alumnos los diferentes puntos del temario, que está dividido fundamentalmente en dos grandes áreas: elementos de la inteligencia artificial (esto es, definiciones, conceptos, principios básicos) y campos de aplicación (a diferentes campos de la ingeniería) (ver Cuadro 1). Intercaladas con las clases teóricas, se realizan prácticas en el laboratorio sobre puntos concretos del temario, que resultan de especial interés por su valor didáctico o su aplicación en ingeniería.

El final del curso se dedica al desarrollo de una práctica global en la que los alumnos, trabajando en equipo, deben enfrentarse a la aplicación práctica de los conceptos aprendidos, diseñando e implementando un sistema informático dotado de inteligencia que resuelva aspectos concretos de ingeniería.

Cuadro 3: Programa de la asignatura

- Elementos de la inteligencia artificial
 - Representación del conocimiento
 - Fundamentos de la lógica formal y sistemas inferenciales
 - Tratamiento de la incertidumbre
 - Resolución de problemas y estrategias de búsqueda
 - Aprendizaje y adquisición del conocimiento
- Aplicaciones de la inteligencia artificial
 - Sistemas basados en conocimiento
 - Agentes
 - Ingeniería lingüística y recuperación de información
 - Minería de datos
 - Sistemas inteligentes y servicios de comunicaciones

3.1. Motivación

Con este planteamiento de la asignatura, se puede considerar que los resultados son altamente satisfactorios tanto para los alumnos como para los docentes, constatado por los

resultados positivos de las encuestas de valoración de la docencia [20] de los últimos años y la evolución creciente del número de alumnos.

Pese a todo, de forma recurrente todos los años, a principio de cada curso, se plantean ciertas inquietudes sobre la asignatura. Sin ser problemas a resolver en sí mismos, son aspectos de mejora donde podría incidirse con vistas al objetivo ambicioso de la educación integral y de calidad, centrada en el alumno.

Los principales aspectos de mejora se pueden resumir en:

- Intentar evitar una visión sesgada de la asignatura. El temario resulta muy interesante, según la opinión de los propios alumnos [20], pero está muy compartimentado. Como puede verse en el programa resumido en el Cuadro 3, bajo el paraguas global de la inteligencia artificial se engloban técnicas y aplicaciones muy diversas, que la asignatura trata de cubrir (en amplitud más que en profundidad), no dando tiempo, en muchos casos, a la adquisición o consolidación de aspectos clave. Todos los años se intenta redefinir y aligerar el temario eliminando aquellos puntos que no aporten fundamentos básicos o que resulten de menor interés para los alumnos, pero queda muy poco margen de maniobra. Además, la práctica final de la asignatura profundiza en un único aspecto de la asignatura, el tema de la propia práctica, incidiendo en este problema.
- Aumentar la motivación de los alumnos. A pesar del elevado interés declarado por los alumnos, se trata de una asignatura optativa, que compite con otras asignaturas (obligatorias, de más créditos, pendientes de cursos anteriores, etc.) por captar la atención del alumno y asegurar el trabajo y esfuerzo necesario para su aprovechamiento.
- Potenciar el desarrollo de habilidades transversales. Pese a la edad de los alumnos, la proximidad de la lectura del Proyecto Fin de Carrera y, por tanto, de su salida al mercado laboral, se detectan en muchos casos carencias y lagunas en ciertas habilidades como la comunicación, capacidad de expresión oral y escrita, análisis crítico y autocrítica, confianza en uno mismo... En definitiva, habilidades transversales a los conocimientos técnicos estudiados en la carrera pero que son igual de importantes (o incluso más) en la formación de los alumnos y futuros profesionales.

En este contexto surge la idea de aplicar la metodología de revisión entre iguales. La idea fundamental en la que se basa es utilizar la participación activa del alumno como factor impulsor del aprendizaje, tanto de los aspectos técnicos de la asignatura como en las mencionadas habilidades transversales. A su vez, la participación activa se fundamenta en la motivación del alumno, puesto que sin motivación no hay participación.

De manera formal, los objetivos específicos que se plantean con el proyecto se resumen a continuación.

- Aprendizaje de calidad: afianzar los conocimientos adquiridos por el alumno y profundizar en los conceptos de la asignatura, fomentando el trabajo continuo, mediante su aplicación práctica y análisis crítico.
- Aprendizaje completo: ampliar la visión del estudiante (centrada generalmente en el tema elegido), complementando su aprendizaje y permitiéndole profundizar en otros aspectos del temario.
- Formación integral: promover una formación integral del alumno, potenciando el desarrollo de habilidades no técnicas imprescindibles, como el trabajo en equipo, la capacidad de expresión y la capacidad de crítica.
- El alumno como centro del proceso educativo: fomentar la asunción de un rol más activo, autónomo y responsable por parte del alumno en el proceso de aprendizaje y evaluación.
- Mejorar el proceso de evaluación: ayudar a los alumnos cuyo rendimiento en los exámenes no es acorde a sus posibilidades por factores ajenos a su preparación y conocimientos (presión, nervios, miedo escénico), afianzando sus conocimientos y permitiéndoles demostrarlos en entornos de menor presión, y aumentando, por tanto, su confianza de cara a la prueba definitiva.

3.2. Desarrollo

El planteamiento del curso 2004-05 incluyó desde el primer momento la participación de los alumnos en el desarrollo de las clases teóricas, con presentaciones orales en el aula por parte de los alumnos voluntarios sobre aspectos complementarios del tema principal del día.

Con estas presentaciones se lograban varios objetivos: ampliar de manera fácil los contenidos de la asignatura a aspectos complementarios de especial interés, desarrollar habilidades comunicativas en los alumnos (desarrollo de los materiales y presentación en público), desarrollar un cierto espíritu crítico por observación de aciertos o fallos de los compañeros, y hacer las clases más amenas implicando a los alumnos.

La idea inicial era realizar una presentación por cada clase teórica, pero la alta participación (fomentada por una mínima "recompensa" en la nota) provocó que hubiera que cambiar la planificación incluyendo dos días de mesas redondas sobre un tema, para dar posibilidades de participar a más alumnos. Todos los alumnos valoraron muy positivamente las presentaciones en clase.

Con estos antecedentes se llega a la práctica final de la asignatura. Como ya se explicó anteriormente, los alumnos tienen que desarrollar en grupos, como parte de la evaluación, una práctica consistente en el desarrollo de un sistema inteligente, aplicando los conceptos presentados a lo largo del curso para solucionar un determinado problema. La lista de prácticas propuestas, así como las normas de desarrollo, se especifican en la página web de la asignatura [21]. La Cuadro 4 resume las prácticas realizadas por los alumnos en este curso.

Cuadro 4: Prácticas desarrolladas

Área	Práctica	Nº alumnos
Simulación	Simulador: Robocode	10
	Simulador: CodeRuler	5
	Simulador: RARS	6
Minería de Datos	MD: Predicción meteorológica	9
	MD: Diagnóstico en cardiología	8
	Reconocimiento de caras con RRNN	2
Sistemas Expertos	SE en turismo con Prodigy	2
	SE "Concesionario de coches"	2
Varios	IA y proyectos de desarrollo sostenible	2

Cada grupo tenía que entregar por escrito una memoria descriptiva del trabajo realizado (incluyendo la descripción del problema, la solución aportada y las técnicas empleadas, y, en su caso, el código desarrollado) y realizar una breve presentación en clase. Además, debían realizar una evaluación de las presentaciones y memorias realizadas por sus compañeros.

La evaluación de las presentaciones se realizó por grupos, durante las propias presentaciones, utilizando un cuestionario-guía que valoraba tanto aspectos técnicos como comunicativos. Cada grupo debía evaluar a los demás, a todos, de modo que debía asistir a las clases de presentación de todas las prácticas. Para obligar a tomar decisiones se pedía además un ranking, y para asegurar un análisis en profundidad se pedían tres aspectos a destacar (positivos) y otros tres a mejorar.

En cambio, la evaluación de las memorias se realizó de forma individual. A cada alumno se le asignaron cuatro prácticas para revisar: una práctica como la suya, una práctica del mismo dominio que la suya, una práctica de otro dominio, y su propia práctica (autoevaluación). El cuestionario de revisión se muestra en la Cuadro 5. La asignación de las prácticas que debía revisar cada alumno se hizo empleando un esquema pseudo-aleatorio: una asignación aleatoria (al azar) entre diferentes grupos de prácticas, combinada con una asignación guiada entre los componentes del mismo grupo de prácticas, para garantizar que cada alumno revisara prácticas distintas que su compañero y eliminar así el posible sesgo del aprendizaje en grupo.

Cuadro 5: Cuestionario de revisión

Puntúa de 0 (peor) a 5 (mejor):

1. El trabajo está tratado en general con la profundidad adecuada. Aspectos en que se hace demasiado hincapié. Aspectos que necesitarían más esfuerzo
2. Valoración global del trabajo
3. Faltas de ortografía (de 5=ninguna a 1=más de 10).
4. Claridad: la redacción es correcta (clara, comprensible...)
5. Estilo: el lenguaje utilizado es el adecuado (lenguaje científico, formal, evita expresiones coloquiales...)
6. Semántica: el trabajo realizado está correctamente descrito
7. Valoración global de la redacción (de 5=excelente a 1=mal)
8. Las hipótesis, proceso y resultados descritos son correctos
9. Una vez visto en detalle el trabajo, valora la exposición oral
10. Valoración global de la memoria
11. ¿Has aprendido algo revisando este trabajo? (de 1=nada a 5=mucho)
12. Mi nivel de conocimientos del tema es (de 1=nulo a 5=muy alto)
13. Mi interés en el tema es (de 1=nulo a 5=muy alto)

3.3. Resultados

La Figura 4 muestra el tiempo de revisión estimado por los propios alumnos de cada una de las prácticas. Aunque la gráfica muestra sólo valores medios, la desviación típica es bastante baja y similar en todos los casos, con lo que se puede asumir que las respuestas son bastante acertadas o que todos los alumnos se han puesto de acuerdo en sus respuestas, lo que resulta poco probable. Así, se puede observar que el proceso de revisión requiere de un cierto tiempo no despreciable (considerando que cada alumno revisó 3 prácticas más la suya propia), que hay que valorar en este tipo de iniciativas por el grado de exigencia y sobrecarga a los alumnos, considerando además la cercanía con el final del cuatrimestre y el periodo de exámenes.

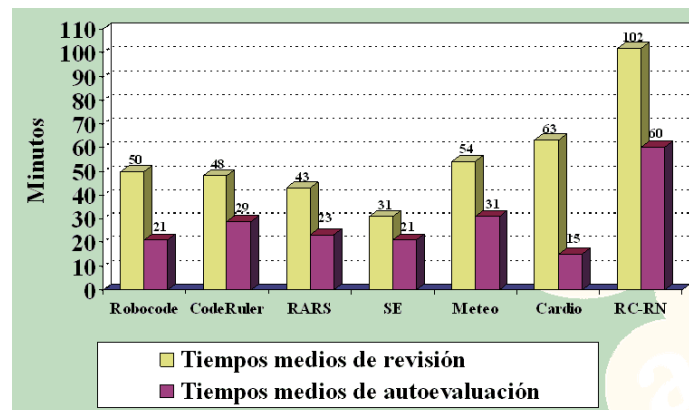


Figura 4: Tiempos de revisión

La figura muestra además que las prácticas consideradas por los alumnos como más difíciles (según su valoración en el cuestionario de evaluación) requieren más tiempo de revisión, lo que resulta lógico teniendo en cuenta además la mayor extensión de las memorias de resultados. El caso de la práctica sobre "reconocimiento de caras empleando redes neuronales" es significativo porque los revisores dedicaron casi el doble de tiempo que a las demás, que se explica por el gran interés que suscitó la memoria (puesto de manifiesto por los comentarios de los revisores en el informe de evaluación).

Además de un cuestionario subjetivo incluido en las propias revisiones para valorar la utilidad y eficacia del proceso percibidas por los alumnos, se realizó también una medida objetiva del aprendizaje de los alumnos. Para ello, se intentaron cuantificar dos efectos: en

primer lugar, el aprendizaje de los alumnos mediante la realización de la práctica (algo comúnmente asumido), En segundo, el aprendizaje de los alumnos en el transcurso de la revisión (quizás en menor medida que el autor, pero en cualquier caso, más que alguien que no haya realizado la revisión), con el objetivo de confirmar la hipótesis que justifica el uso de esta metodología.

Como instrumento de medida, se incluyeron en el examen teórico 10 preguntas relacionadas con las prácticas, dos preguntas de cada una de las prácticas más realizadas por los alumnos: Robocode, CodeRuler y RARS –del área de simulación–, y predicción en meteorología y diagnóstico en cardiología –del área de minería de datos–. Para aumentar la fiabilidad de los resultados, estas cuestiones (de tipo test: verdadero/falso) se intercalaron camufladas entre el resto de preguntas del examen, de modo que los alumnos pensaran que las respuestas puntuaban como en el resto del examen, aunque en realidad no se utilizaron para la calificación final.

Es importante puntualizar que somos conscientes del relativo rigor empírico de este tipo de experimentos sociales, con poblaciones tan reducidas, sesgadas en comportamientos o actuaciones, y con tan pocas preguntas de control, por lo que las conclusiones que se puedan obtener son de tipo cualitativo más que cuantitativo, y hay que considerarlas en su aspecto de aplicación más general.

La Figura 5 presenta el número medio de respuestas correctas de los alumnos en cada uno de los bloques de preguntas. Como se puede observar, en todos los casos, el número medio de respuestas correctas de los autores de las prácticas (esto es, respuestas a preguntas relacionadas con la propia práctica) es superior al número medio de respuestas correctas de toda la clase en general, y, a su vez, esta media de la clase es superior a la media de los alumnos que no habían realizado esta práctica. Estos resultados muestran una medida objetiva y cuantitativa del aprendizaje al realizar una práctica.

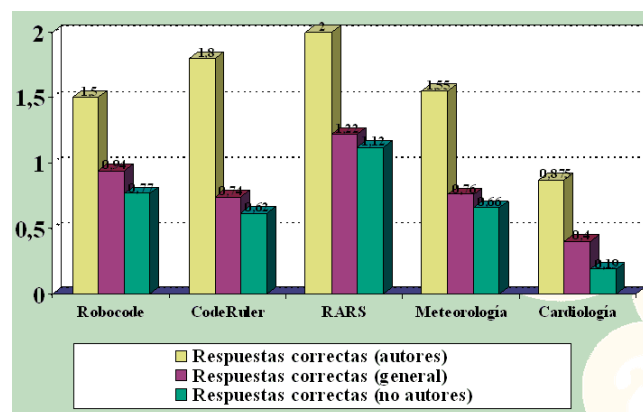


Figura 5: Número medio de respuestas correctas

El valor tan bajo del número medio de respuestas correctas en las preguntas sobre diagnóstico en cardiología, incluido el caso de las respuestas de los propios autores, se explica porque una de las preguntas estaba redactada de forma ambigua y resultaba confusa para los alumnos, lo que se puso de manifiesto tras la evaluación del experimento.

Frente a la figura anterior, la Figura 6 muestra la comparativa del número medio de respuestas correctas de los autores de las prácticas frente a las respuestas de los revisores y del resto de alumnos (que no son ni autores ni revisores). Según se observa en todos los casos, el número medio de respuestas correctas de los revisores es superior al número medio de respuestas correctas de alumnos que ni han hecho la práctica ni tampoco la han revisado, esto es, una evaluación objetiva del aprendizaje. Estos resultados permiten inferir (considerando el rigor relativo intrínseco a este tipo de experimentos con poblaciones reducidas) que la metodología de revisión entre pares es válida como instrumento de aprendizaje.

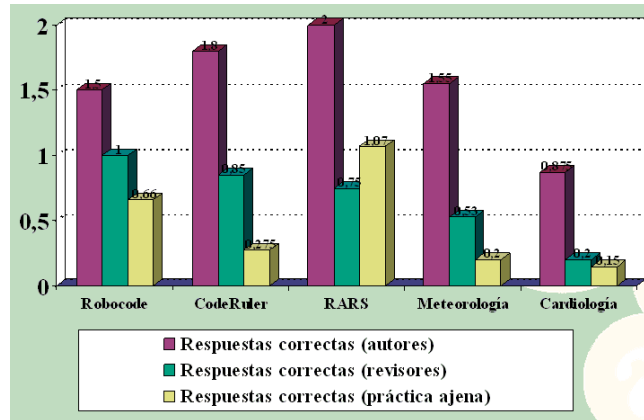


Figura 6: Comparativa autores/revisores

El hecho de que los resultados de los revisores sean peores que los de los propios autores resulta lógico y es lo que cabría esperar (simplemente teniendo en cuenta el tiempo dedicado a la revisión frente a la realización de la práctica), por lo que sirve de control de la calidad de los resultados de este experimento.

La Figura 7 presenta un resultado interesante respecto al grado de confianza que tienen los alumnos en sus respuestas. La primera figura muestra el número de respuestas correctas/incorrectas/en blanco de los alumnos sobre su propia práctica (esto es, autores), y la segunda figura muestra lo mismo pero para revisores. Dado que los alumnos estaban convencidos de que las respuestas incorrectas penalizaban en la nota del examen, se puede suponer que una pregunta que no estuviera en blanco indica que el alumno estaba convencido de conocer la respuesta.

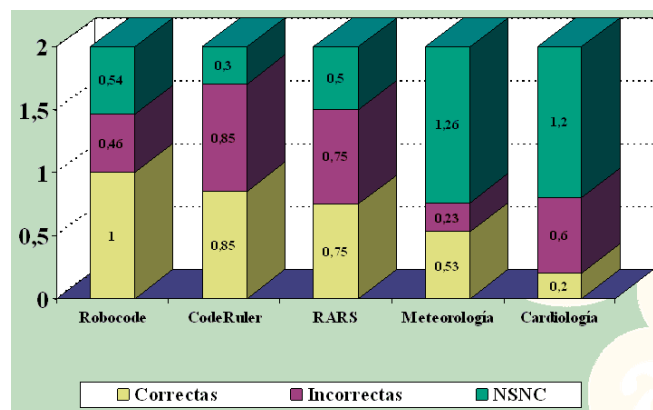
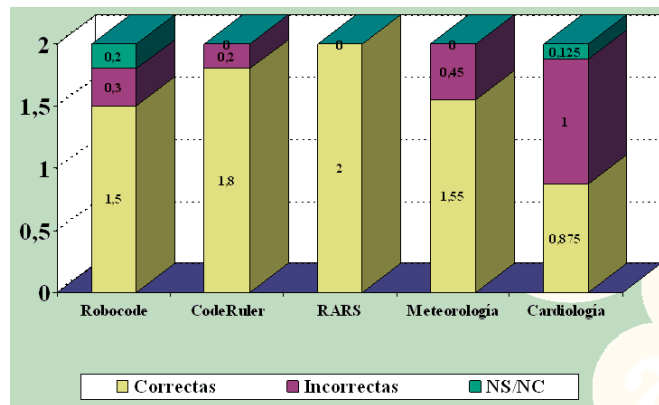


Figura 7 (a y b): Número de respuestas correctas/incorrectas/en blanco de autores (a) y revisores (b)

Se observa que en el caso de los autores, el número medio de respuestas en blanco es muy bajo (o nulo en algún caso –RARS–, esto es, que todos los alumnos contestaron a las dos preguntas), lo que indica que los alumnos pensaban que conocían la respuesta (independientemente de que luego acertaran o no). Sin embargo, los revisores muestran una mayor desconfianza en sus respuestas (el número de respuestas en blanco aumenta de forma notable), sobre todo en el caso de las prácticas percibidas como más difíciles, que son las de minería de datos.

4. Conclusiones

Respecto a los objetivos de aprendizaje, los formularios incluían una encuesta cualitativa para recoger la opinión subjetiva de los alumnos. Se confirman dos claras tendencias: aprendizaje “positivo”, de nuevos conceptos o procedimientos desconocidos a priori por el revisor, y aprendizaje “negativo”, a partir de errores ajenos. Al ser una medida subjetiva, se confirma también que los alumnos reconocen fácilmente el aprendizaje de primer tipo, pero es una minoría la que cuenta también con el segundo.

La iniciativa confirma carencias de los alumnos detectadas en otros cursos. Se han puesto de manifiesto dificultades de los alumnos con habilidades técnicas importantes, como la prueba de programas o resolución de errores³ (en OCA). Y, fundamentalmente, dificultades con habilidades no técnicas imprescindibles: la expresión escrita es en muchos casos muy deficiente, afectando incluso a la comprensión del texto⁴.

Es destacable la positiva respuesta de los alumnos. La iniciativa ha sido muy bien acogida por los alumnos y se han recibido valoraciones muy positivas. El esfuerzo dedicado a las revisiones por parte de los alumnos resulta de agradecer, especialmente cuando no se les exigía un determinado tiempo de dedicación a la revisión. En OCA, el grado de participación fue mucho mayor de lo previsto inicialmente, sobre todo teniendo en cuenta su carácter voluntario.

Por último, si bien se ha confirmado la utilidad de la metodología como instrumento de aprendizaje, los resultados obtenidos cuestionan su posible uso como herramienta de evaluación. Las calificaciones asignadas por los alumnos son excesivamente homogéneas y generosas. A la vista de los comentarios y las respuestas parciales de los cuestionarios, no se trata de que no sean capaces de discriminar el valor real de un trabajo, sino que se dejan influir por factores ajenos al proceso evaluativo en sí (por ejemplo, compañerismo). Así, a menudo las impresiones parciales que se les pedían reflejaban algo más correctamente el nivel real de la práctica examinada, pero eran maquilladas en la evaluación global que también se pedía.

La conclusión, a partir de los resultados obtenidos, es que la revisión entre pares debe complementarse con mecanismos de control de calidad estrictos para asegurar su fiabilidad como instrumento de evaluación, si bien ha demostrado su valor como instrumento de aprendizaje, que era el objetivo de esta experiencia.

Los principales problemas y dificultades encontrados han sido los habituales de la puesta en marcha de un proyecto de este estilo en un entorno docente real. Los problemas relacionados con la co-evaluación formativa derivan fundamentalmente de dificultades relativas a su aplicación en entornos reales:

- El problema típico de esta metodología es que algunos alumnos no entregan las revisiones asignadas, o no las completan a tiempo, con el consiguiente perjuicio para los compañeros autores destinatarios de la misma.
- Garantizar el anonimato en grupos reducidos es un problema no trivial y de muy difícil resolución. Frecuentemente los alumnos pueden reconocer los trabajos de sus compañeros, con lo cual la revisión no es anónima y puede inhibir la sinceridad de las revisiones.

³ Por ejemplo, en OCA, algunos revisores no conseguían visualizar el *applet* pedido simplemente porque el código no estaba compilado, tal como se pedía en el enunciado.

⁴ Se ha dado el caso, también en OCA, de tener que explicar el profesor (más acostumbrado por lo que parece a descifrar las expresiones de los alumnos que ellos mismos) a un grupo la revisión enviada por otro alumno.

Otra dificultad encontrada deriva de los reajustes realizados en la temporización inicial con el fin de lograr una mejor integración en la planificación del curso y con la agenda de los estudiantes. Una metodología secuencial como la propuesta, en la que cada fase depende de los resultados de la anterior⁵, es fuertemente sensible a los retrasos, que se van acumulando. En consecuencia, algunas revisiones (es el caso de IRC y del último ciclo de OCA) se realizaron en fechas próximas al final del curso, en las que los alumnos disponían de menos tiempo por la proximidad de los exámenes, lo cual se reflejó en la calidad de algunos trabajos.

Por último, también han existido problemas de carácter tecnológico:

- Aunque el sistema de gestión de revisiones utilizado había sido exhaustivamente probado antes de su uso, adolece de la falta de una interfaz de usuario amigable. Ello redundaba en un mayor consumo de recursos para su uso por parte del profesor. E introduce la posibilidad de confusiones, especialmente por parte de los alumnos, menos habituados a su uso.
- Las limitaciones en los recursos también han supuesto dificultades en algunos casos. El ancho de banda de las conexiones de los alumnos ha originado problemas en algunas entregas de ficheros de gran tamaño. La cuenta de correo de los alumnos tiene una capacidad limitada y los alumnos no siempre liberan el espacio necesario, por lo que no ha sido necesario recurrir a mecanismos alternativos de intercambio de datos en algunas ocasiones.

Referencias y bibliografía

- [1] K. Topping, "Peer Assessment Between Students in Colleges and Universities", *Review of Educational Research*, vol. 68, no 3, pp. 249-276, Fall 1998.
- [2] B. W. Ford, "The effects of peer editing/grading on the grammar-usage and theme-composition ability of college freshmen", in *Dissertation Abstracts International*, 33, 6687, 1973.
- [3] "Daedalus integrated writing environment", The Daedalus Group, 2003. [Online]. Available: www.daedalus.com
- [4] C: Helfers, S. Duerden, J. Garland, and D. L. Evans, "An effective peer revision method for engineering students in first-year English courses", in *Frontiers in Education Conference*, November 1999, pp. 7-12.
- [5] S. Nelson, "Teaching collaborative writing and peer review techniques to engineering and technology undergraduates", in *Frontiers in Education Conference*, 1998, pp. 14-18.
- [6] "Calibrated peer review", 2004. [Online]. Available: cpr.molsci.ucla.edu
- [7] J. McGourty, P. Dominick and R. R. Reilly, "Incorporating student peer review and feedback into the assessment process", in *1998 FIE Conference*, 1998, pp. 14-18.
- [8] J. McGourty, "Using multisource feedback in the classroom: A computer-based approach", *IEEE Transactions on Education*, vol. 43, no.2, pp. 120-124, May 2000.
- [9] E. Z.-F. Liu, S. Lin, C.-H. Chiu and S.-M. Yuan, "Web-based peer review: the learner as both adapter and reviewer", *IEEE Transactions on Education*, vol. 44, no. 3, pp. 246-251, August 2001.
- [10] E. F. Gehringer, "Strategies and mechanisms for electronic peer review", in *Frontiers in Education Conference*, October 2000.
- [11] E. F. Gehringer, "Electronic peer review and peer grading in computer-science courses", in *Proc. Of the Technical Symposium on Computer Science Education*, 2001, pp. 139-143.
- [12] A. Ward, J. Sitthiworachart and M. Joy, "Aspects of web-based peer assessment systems for teaching and learning computer programming", in *IASTED International Conference on Web-based Education*, February 2004, pp. 292-297.
- [13] "Curso: Proyecto docente", Plan de formación de profesorado curso 2001/2002, Gesem – Universidad Carlos III de Madrid, 2002.
- [14] R. Lister, "Objectives and objective assessment in CS1", in *Proc. Of the Technical Symposium on Computer Science Education*, pp. 292-296, 2001.

⁵ Obviamente, no se puede revisar una práctica antes de entregarla, por ejemplo.

- [15] www.it.uc3m.es/tsioca/curso0304/PracticasPuntuables/pr1/
- [16] www.it.uc3m.es/tsioca/curso0304/PracticasPuntuables/pr2/
- [17] www.it.uc3m.es/tsioca/curso0304/PracticasPuntuables/pr3/
- [18] R. M. Crespo García, "Cuestionario de revisión", Organización de Contenidos Audiovisuales, 2004. [Online]. Available: www.it.uc3m.es/tsioca/curso0304/pr/PracticasPuntuables/revision1.html
- [19] R. M. Crespo García, "Cuestionario de valoración", Organización de Contenidos Audiovisuales, 2004. [Online]. Available: www.it.uc3m.es/tsioca/curso0304/pr/PracticasPuntuables/cuestionario1.html
- [20] Universidad Carlos III de Madrid, Encuestas de evaluación de la docencia (asignatura Inteligencia en Redes de Comunicaciones), 2005.
- [21] Julio Villena Román, "Inteligencia en Redes de Ordenadores" (web de la asignatura). [Online]. Available: www.it.uc3m.es/jvillena/irc/indice.html
- [22] R.Mª Crespo, A. Pardo, C. Delgado Kloos, "An Adaptive Strategy for Peer Review", in *Frontiers in Education Conference, October 2004*. Savannah, GA, USA.

Agradecimientos

Este proyecto ha sido financiado por la Universidad Carlos III de Madrid, en el marco de la 2ª Convocatoria de apoyo a Experiencias de Innovación Docente, organizada por el Vicerrectorado de Ordenación Académica. Agradecemos especialmente a Patricia López Navarro, directora del Programa de Mejora de la Docencia y el Aprendizaje, su inestimable ayuda y dedicación.

Asimismo, agradecemos su colaboración al resto de profesores de la asignatura OCA, José Jesús García Rueda y Francisco Gómez Molinero, por su ayuda en el desarrollo de esta experiencia, y a Abelardo Pardo por el sistema de entregas y la colaboración prestados.

Y, cómo no, a los alumnos de ambas asignaturas por su participación e ilusión.