



# Las eRúbricas en la evaluación cooperativa del aprendizaje en la Universidad

## eRubrics in Cooperative Assessment of Learning at University

-  Dr. MANUEL CEBRIÁN-DE-LA-SERNA es Catedrático de Universidad en la Facultad de Ciencias de la Educación de la Universidad de Málaga (España) (mcebrian@uma.es).
-  Dr. JOSÉ SERRANO-ANGULO es Profesor Titular en la Facultad de Ciencias de la Educación de la Universidad de Málaga (España) (joseserrano@uma.es).
-  Lic. MAYERLY RUIZ-TORRES es Profesora de la Universidad Santo Tomas de Aquino de Bucaramanga (Colombia) (2009-2011) (mayerlyzulayr@hotmail.com).

### RESUMEN

El trabajo cooperativo mediante tareas y proyectos en equipo es una de las metodologías más generalizadas en educación para lograr las competencias de aprendizaje. Estas metodologías están teniendo cada vez mayor aceptación en la enseñanza universitaria, a pesar de la dificultad de conocer los aprendizajes individuales producidos en cada uno de sus miembros. Este artículo muestra los resultados de un proyecto de I+D+i cuyo objetivo general consiste en «Analizar el impacto de las eRúbricas –rúbricas electrónicas– en la evaluación de los aprendizajes universitarios en sus diferentes modalidades», y pretende mostrar el alcance de esta herramienta para mejorar el aprendizaje de las competencias o habilidades cooperativas, producto del trabajo en equipo y la evaluación cooperativa de tareas en el laboratorio. La experiencia se desarrolla con tres grupos elegidos (205 estudiantes) de un total de seis grupos de estudiantes de 1º del grado de Primaria en el curso 2011-12. De los tres grupos, uno actúa como grupo de control y los otros dos como grupos experimentales en los que se han utilizado eRúbricas. Se observan diferencias en las notas de una prueba escrita común a todos, con mejores resultados en los grupos con eRúbrica. Además, se realiza un análisis cualitativo categorizando las respuestas dadas por los estudiantes del grupo de control sobre los criterios de evaluación que éstos utilizan, para ver las coincidencias con los criterios de las eRúbricas usadas en los grupos experimentales.

### ABSTRACT

Teamwork is one of the most widespread teaching methods used to achieve learning skills. Despite the difficulty of finding out the degree of individual learning taking place in each member of the group, these methods are having an increasingly greater importance in university teaching. The present article shows the results of an R+D+i project aimed at «analysing the impact of eRubrics –electronic rubrics– on the assessment of university learning in various forms». Likewise, it aims to show the scope of eRubrics in improving cooperative skills, which are achieved through teamwork and cooperative assessment of tasks in the computer lab. The experiment takes place in three groups selected from a total of six groups of students from the First Year of Primary Education Teaching during the 2011-12 academic year. From the three groups, one acted as the control group and the other two as the experimental groups in which eRubrics were used. Differences were found in students' results in a written test taken by all the groups, as the group using eRubrics achieved better results than the other two. Additionally, a qualitative analysis was conducted, by classifying the students' answers in the control group with regard to the evaluation criteria they used, in order to check for coincidences with the eRubric criteria used by students in the experimental groups.

### PALABRAS CLAVE | KEYWORDS

Evaluación formativa, trabajo en equipo, aprendizaje cooperativo, evaluación cooperativa, formación inicial, enseñanza universitaria, competencias, eRúbricas.

Formative assessment, teamwork, cooperative learning, cooperative assessment, preservice teachers, university education, competences, erubrics.

## 1. Introducción

Entre las recientes modalidades didácticas desarrolladas en la educación, los modelos de colaboración mediante tecnologías (CSCL)<sup>1</sup> (Voogt & Knezek, 2008) representan una renovación profunda en la enseñanza. En las propuestas para el diseño y planificación de la docencia universitaria se insiste en la importancia de estas modalidades y el uso de las tecnologías para centrar la enseñanza en el aprendizaje de los estudiantes (Zabalza, 2010), a la vez que involucrarlos más en todo el proceso, especialmente en la evaluación (Falchikov & Goldfinch, 2000; Brown & Glasner, 2003; Falchikov, 2005; Blanco, 2009; López-Pastor, 2009), con lo que resultaría el planteamiento de diferentes modalidades metodológicas y distintas formas de organizar los procesos de enseñanza-aprendizaje según los diferentes contextos universitarios (DeMiguel, 2006: 31). Es aquí donde «el trabajo en equipo» aglutina como finalidad todos estos valores y principios pedagógicos: «que los estudiantes aprendan y evalúen en colaboración mediante la adopción de un papel más activo y comprometido con la enseñanza y el aprendizaje mediado con tecnologías».

El aprendizaje humano es fundamentalmente social, por lo que la construcción del conocimiento y el aprendizaje colaborativo es importante que se atienda como objetivo en todos los niveles educativos (Hargreaves, 2007). No obstante, y dado que los contextos y los estudiantes no siempre poseen los prerrequisitos mínimos para abordar un modelo de colaboración, suele emplearse con más frecuencia y como paso previo, un modelo de aprendizaje cooperativo, que resuelve en gran parte esta situación de partida al proporcionar mayor estructuración y guías a los estudiantes, a la vez que brinda mayor control a los docentes.

Cuando pretendemos extender el aprendizaje colaborativo en todos los momentos del proceso de enseñanza, como en el caso de la evaluación, la necesidad de una orientación guiada se hace más patente y crucial, y la metodología del «trabajo en equipo» y la «evaluación cooperativa» resultan ser destacados recursos y técnicas pedagógicas como antesala a un modelo de evaluación colaborativa.

El aprendizaje entre iguales es especialmente beneficioso cuando se centra en los procesos de evaluación, donde toma en la literatura especializada con más frecuencia el nombre de «evaluación colaborativa» (Blanco, 2009: 115; Brown & Glasner, 2003: 31; López-Pastor, 2009: 94), utilizando otros términos sinónimos como: «co-evaluación», «evaluación compartida», «evaluación entre pares», etc. Surge la necesidad de una mayor definición conceptual y del uso de

los términos utilizados, dado que no siempre se diferencian —como sí ocurre entre el aprendizaje cooperativo vs. colaborativo—, entre los conceptos de la evaluación cooperativa frente a la evaluación colaborativa —más estructurada y guiada, en la primera frente a la segunda—.

Si bien, estas prácticas están generalizándose cada vez más, también existen voces críticas en cuanto a diferentes aspectos:

- La dificultad de un seguimiento y evaluación individualizada de las competencias adquiridas por los diferentes miembros de un equipo.
- Una revisión del impacto de dichas metodologías en los aprendizajes de los estudiantes según los nuevos contextos, especialmente por el uso de las tecnologías.
- Y por último, saber cómo abordar los prerrequisitos de los estudiantes para alcanzar una evaluación colaborativa que requiere mayor capacidad de reflexión y autocrítica.

Uno de los principios que sostiene la evaluación cooperativa consiste en hacer partícipes a todos sus miembros de los criterios por los cuales son evaluadas las evidencias de aprendizajes en las tareas y los proyectos en equipo. Este planteamiento más comunicativo y participativo sobre la evaluación comienza desde el intercambio y comprensión de objetivos, metas y procedimientos... hasta llegar a la evaluación de los procesos y sus resultados. En este proceso y en sus resultados es donde se aplican criterios e indicadores de calidad, y donde, a pesar de estar bien definidas y estructuradas las tareas, suelen existir dificultades de comunicación entre los docentes y los estudiantes, sobre todo cuando media Internet y la distancia, siendo en muchas ocasiones solo comprendida la evaluación al final del proceso si hay un análisis compartido.

Es obvio que esta situación problemática en la comunicación mediada por la tecnología se mejora con un diálogo permanente entre los docentes y los estudiantes sobre los indicadores de calidad, los criterios y cómo son aplicados éstos a las evidencias que plantean los trabajos en equipo. Unas de las técnicas y herramientas que facilita esta comunicación en la evaluación son las rúbricas (Osana & Seymour, 2004; Jonsson, & Svingby, 2007; Reddy & Andrade 2010; Rodríguez-Gómez & Ibarra-Sáiz 2011; Panadero & Jonsson, 2013), que en su versión digital se denominan «eRúbricas». Encontrando en su definición y entre sus funciones y ventajas, el permitir a los docentes y estudiantes compartir los indicadores, los criterios y las evidencias, así como el peso y la aplicación de las mismas en la evaluación de los objetos de aprendizaje

(Andrade, 2005). Al ser digital y federada –eRúbrica federada– es más interactiva, al tiempo que la federación proporciona el soporte ideal para la cooperación y la colaboración de los usuarios, salvando las dificultades de la interoperatividad entre herramientas, servicios, contextos y sistemas tecnológicos diferentes, ubicados tanto dentro como fuera de la propia institución educativa.

La eRúbrica federada representa un doble papel en la enseñanza: por un lado, como tecnología es un soporte ideal para mejorar la comunicación y comprensión sobre los procesos de evaluación, al tiempo que facilita los modelos de trabajo en equipo. Representa un instrumento imprescindible en los procesos de evaluación con eportafolios, dado los modelos de seguimiento y la calidad que requiere las interacciones entre el docente y los estudiantes para comprender los indicadores, criterios y evidencias de aprendizaje, sobre todo en la distancia y mediado por tecnologías, que en muchos casos se realizan entre instituciones con sistemas tecnológicos diferentes. Como el caso del practicum, cuando estos estudiantes están repartidos en distintas instituciones, cada una con sus herramientas y sistemas tecnológicos (Meeus, Petegem & Engels, 2009; Cebrián-de-la-Serna, 2011; Del-Pozo, 2012). Por otro lado, como técnica y metodología que facilita la evaluación formativa, dado que exige definir con claridad la escala de estándares de aprendizaje y la aplicación de criterios relacionados con una tarea. Existiendo una extensa literatura sobre su impacto con trabajos pioneros como Hafner y Hafner (2003) y Falchikov, (2005); pasando por el aprendizaje profundo y auténtico de Vickerman (2009); hasta llegar a las investigaciones sobre la evaluación entre pares en entornos de colaboración mediados por tecnologías (CSCL) (Prins, Sluijsmans, Kirschner & Strijbos, 2005); y los estudios algo más escasos dentro del contexto de la formación inicial de docentes y la adquisición de competencias profesionales (Osana & Seymour, 2004; Bartolomé, Martínez & Tellado, 2012, Gámiz-Sánchez, Gallego & Moya, 2012; Moril,

Ballester & Martínez, 2012; Martínez, Tellado & Raposo, 2013; Panadero, Alonso-Tapia & Reche, 2013).

No obstante, y a pesar de estos resultados, debemos considerar tales trabajos con prudencia y tener una visión más de conjunto con meta-análisis como los de Jonsson y Svingby (2007), y Reddy y Andrade (2010), donde nos ofrecen una visión general de las rúbricas en la enseñanza universitaria, destacando, por un lado, que existe una percepción positiva por parte de los estudiantes en la utilización de los programas,

**La eRúbrica federada representa un doble papel en la enseñanza: por un lado, como tecnología es un soporte ideal para mejorar la comunicación y comprensión sobre los procesos de evaluación, al tiempo que facilita los modelos de trabajo en equipo. Representa un instrumento imprescindible en los procesos de evaluación con eportafolios, dado los modelos de seguimiento y la calidad que requiere las interacciones entre el docente y los estudiantes para comprender los indicadores, criterios y evidencias de aprendizaje, sobre todo en la distancia y mediado por tecnologías, que en muchos casos se realizan entre instituciones con sistemas tecnológicos diferentes.**

mientras que hay estudios que señalan la resistencia de ciertos grupos de docentes a utilizarlas. Y por otro lado, hay investigaciones sobre el impacto positivo de las rúbricas sobre los rendimientos académicos, mientras que otros estudios no encuentran este impacto.

Sin duda, necesitamos más estudios sobre el impacto de las eRúbricas a pesar de la existencia de esta literatura amplia y extensa. En especial sobre la evaluación cooperativa y colaborativa, que si bien, ya fueron estudiadas desde esta perspectiva –«collaborative assessment»– (Falchikov, 2005: 125), no ha sido así con el auge reciente de las nuevas tecnologías, especialmente en el estudio del impacto de las «eRúbricas federadas», que a diferencia de las rúbricas de papel son más interactivas, facilitando la comunicación, cooperación y colaboración entre estudiantes y docentes

de diferentes instituciones. Queda por tanto conocer su alcance y su impacto en los modelos de enseñanza y aprendizaje cooperativo y colaborativo. Por lo que es necesario realizar investigaciones donde puedan analizarse las nuevas funciones interactivas y comunicativas que ofrecen estas tecnologías y las redes sociales (Bartolomé, 2012); especialmente, como propone Reddy y Andrade (2010) con metodologías más rigurosas, con mayor fiabilidad y validez en sus procedimientos, desde perspectivas geográficas y culturales más amplias.

Siguiendo en esta línea de estudios, los resultados que a continuación se presentan son parte de un proyecto de investigación, donde se experimentan las tecnologías de federación en general y las «eRúbricas federadas» en particular para la educación y la colaboración intra e interinstitucional. Siendo el tema que nos ocupa aquí, la evaluación cooperativa entre los estudiantes y las tareas de equipos desarrolladas en el laboratorio.

Utilizamos para dicha cooperación dentro de la misma institución, la interoperatividad que permiten las tecnologías de federación, donde los estudiantes solo necesitan autenticarse –log in– y –log out– una sola vez para utilizar las diferentes herramientas y servicios federados puestos a disposición en esta experiencia e investigación, como fueron las siguientes: la plataforma institucional donde se subieron y se compartieron los recursos de las tareas; el servicio de eRúbrica federada para la evaluación cooperativa; la herramienta «Consigna federada» para subir y compartir ficheros de gran tamaño; el servicio de «Webquest federado» para la elaboración de materiales didácticos; y por último, el servicio de Limesurvey federado para recoger las evaluaciones abiertas de los grupos de control, con el fin de comparar sus resultados<sup>2</sup>.

## 2. Metodología

La utilización de las rúbricas en la evaluación de los aprendizajes se ha introducido en diferentes asignaturas y en diferentes titulaciones universitarias, pero en su formato digital –eRúbrica federada– no es muy corriente su uso. La innovación de este proyecto radica precisamente en la poca existencia de experiencias con estas tecnologías, al tiempo que, al investigar su impacto, parte de un marco conceptual más amplio hasta la presente, al introducir una nueva variable que puede representar diferentes papeles según sea una evaluación cooperativa y/o colaborativa (en el primer caso, eRúbricas dadas por el docente, y en el segundo, eRúbricas negociadas).

Si bien nuestra investigación no aborda todas las posibilidades del gráfico 1, sí plantea la necesidad de

responder a las siguientes preguntas: ¿Mejoran los aprendizajes académicos de los estudiantes cuando se utiliza la evaluación cooperativa con eRúbricas en el trabajo en equipo?, ¿qué criterios de evaluación utilizan los estudiantes en una evaluación entre pares sin la estructuración y guía de las eRúbricas? Tomando como referencia estas preguntas, los objetivos específicos del proyecto son los siguientes:

1) Analizar el impacto de la eRúbrica en los aprendizajes académicos mediante el desarrollo de metodología de la evaluación cooperativa y el trabajo en equipo (evaluación cooperativa con eRúbricas).

2) Analizar los criterios y valoraciones utilizados por los estudiantes en las evaluaciones de los trabajos de sus compañeros sin orientaciones y guías para la evaluación (evaluación cooperativas sin eRúbricas).

Con la respuesta a estas preguntas y el desarrollo de los dos objetivos de investigación pretendemos mostrar la utilidad y la eficacia del empleo de la eRúbrica como herramienta y metodología para la evaluación formativa, permitiendo mejorar los aprendizajes de los estudiantes; al tiempo que la comprensión e interiorización de los criterios de evaluación y su aplicación.

La investigación estuvo planificada en dos fases; en la primera, se consensaron y diseñaron los contenidos de las eRúbricas y su funcionalidad, experimentando por primera vez los contenidos con «limesurvey». En la segunda fase, una vez evaluado los contenidos de las rúbricas, así como creada la herramienta eRúbrica propia, se aplicó el diseño de investigación previsto. Para esta segunda fase se ha empleado un enfoque plurimetodológico, tanto por las características de los objetivos como por la naturaleza de los datos a recoger, donde se pudo observar la realidad a nivel cualitativo y cuantitativo. Para conseguir el primer objetivo específico se ha diseñado una metodología cuasi-experimental, donde un grupo no utiliza la eRúbrica (grupo de control) y sus resultados se comparan con otros dos grupos de clases que sí lo utilizan (grupos experimentales). Para el logro del segundo objetivo se ha planteado una metodología cualitativa mediante el análisis de contenido, cuyas valoraciones son obtenidas de las evaluaciones sin eRúbricas de los trabajos del grupo de control.

La muestra estudiada está constituida por tres grupos seleccionados al azar, de seis grupos que constituyen la población de la titulación del grado de Primaria que cursan la asignatura «Tecnologías de la comunicación y la información aplicadas a la educación» de la Facultad de Ciencias de la Educación de la Universidad de Málaga en el periodo académico 2011-12.

Estos tres grupos contaban con 75 estudiantes cada uno, y el contexto del diseño de investigación para ambos objetivos se realizaron con la división de cada grupo en dos subgrupos de 37 estudiantes aproximadamente en los laboratorios de informática (6 subgrupos en total) con dos horas de clase para la realización de las tareas y la evaluación entre iguales. Por tanto, la investigación se realizó con el 50% de la población alcanzando un total de 225 estudiantes, con 75 para el grupo de control y 150 para los grupos experimentales. Los contenidos de las eRúbricas pueden encontrarse en la base de datos pública de la propia herramienta introduciendo la descripción de la asignatura mencionada.

La elección de la muestra se realizó con la técnica probabilística aleatoria por conglomerados, ya que en este muestreo la unidad muestral es el grupo, minimizándose las diferencias que pudiesen existir entre el grupo control y el grupo experimental, asignando al azar los grupos que recibirían el tratamiento y el grupo que haría de control, con el fin de lograr una equivalencia entre los dos, y así evitar como manifiestan Colás, Buendía y Hernández (2009), problemas de validez interna y externa, determinando a los grupos B y C como el grupo experimental y el grupo A como el grupo de control.

Para llevar a cabo la investigación se realizaron cuatro actividades durante el curso, cada una de ellas con la misma metodología de evaluación en los tres grupos de clase.

La metodología de clase fue dirigida por el mismo docente en los tres grupos y siguiendo estas fases:

- Dos horas: Exposición de la actividad en todo el grupo (75 estudiantes), realización de la tarea y evaluación por todos los estudiantes coordinada por el docente.

- Dos horas: Subdivisión de la mitad de cada grupo (37 aproximadamente) en el laboratorio donde en equipo se realizaba la misma tarea anterior pero con otros materiales o ejemplos. Para la tarea se utilizaba una ficha ubicada en una plataforma, se bajaba para su realización y se subían los resultados del equipo a un espacio compartido. Después de que todos hubieran subido la tarea se abría el acceso a la plataforma para que el resto de los equipos la bajase y la evaluaran de forma individual. Al final de todo el proceso el docente cerraba la posibilidad de evaluación y de subir o bajar los trabajos. Realizando éste la evaluación de todos los equipos y subiendo los resultados de la tarea a la plataforma.

Las cuatro tareas analizadas y sus objetivos eran de diferente naturaleza, ofreciendo a los grupos A, B y C

las mismas tareas como sus explicaciones. La evaluación entre los equipos se realizaba con una combinación al azar dada por el docente, evitando la evaluación entre equipos iguales, a pesar de que en todos los casos las evaluaciones eran anónimas. La tarea se realizaba en equipo de tres a cinco estudiantes, pero las evaluaciones eran individuales, cada miembro del mismo equipo evaluaba el trabajo del equipo asignado por el docente.

Para recoger los datos cuantitativos se diseñaron cuatro eRúbricas diferentes, una para cada tarea, desde donde los estudiantes de los grupos experimentales realizaban la evaluación cooperativa, exportándose al final del proceso todas las evaluaciones de éstos y del docente a una hoja Excel. Todos los estudiantes estaban identificados con un número a fin de poder comparar su calificación individual en la prueba final del curso, con sus evaluaciones realizadas y recibidas entre sus iguales durante las prácticas de laboratorio. En cambio, para el grupo de control no se utilizó ninguna eRúbrica para evaluar los trabajos de sus compañeros, tan solo un cuestionario –Limesurvey federado– con una sencilla pregunta abierta: «¿Qué te ha parecido la actividad que ha realizado este equipo?», este enfoque para la evaluación entre iguales y sin criterios dados por el docente pretendía sin lograrlo del todo un modelo de evaluación colaborativa (por su falta de guías, orientación y estructuración); no obstante, este planteamiento nos permitía conocer sus argumentos, los criterios y las reflexiones utilizadas en las evaluaciones de sus compañeros en el grupo A, como un medio quizás –con más estudios sin duda– para conocer los prerrequisitos que requiere una evaluación colaborativa.

### 3. Análisis y resultados

Dado que todos los estudiantes estaban identificados (grupo experimental y de control) nos permitía un análisis de sus resultados con otras variables como las calificaciones finales de la asignatura y las evaluaciones específicas que sobre las prácticas de laboratorio se realizó al final del curso. Esta prueba de las prácticas consistía en un examen individual sobre un ejemplo elegido al azar entre las cuatro tareas diseñadas para mostrar las mismas competencias de las prácticas del laboratorio, pero con materiales diferentes a los utilizados durante el curso.

Entre los diferentes análisis, la metodología nos permitió para el primer objetivo específico, analizar las calificaciones individuales y grupales obtenidas en el examen práctico y las evaluaciones de las cuatro actividades realizadas en el laboratorio, tanto en el grupo

A (grupo de control) como los grupos B y C (grupos experimentales).

Para el segundo objetivo específico, nos permitió comparar las categorías halladas en el análisis de contenidos de las valoraciones y evaluaciones del grupo A (grupo de control) con las evidencias y criterios de la eRúbrica utilizadas para los grupos B y C.

### 3.1. Análisis cuantitativo

Para el análisis cuantitativo se consideró la variable independiente como el uso, o no, de la eRúbricas por los estudiantes y el docente. Como se dijo antes, al final del curso se realizó una misma prueba escrita a todos los grupos, la nota obtenida en esta prueba se consideró como la variable dependiente, determinando si había diferencias entre las notas de los estudiantes en los distintos grupos de clase.

Para comparar las notas de la prueba escrita, se ha realizado un análisis de varianza con los tres grupos, ya que los dos grupos experimentales han mostrado evoluciones distintas como grupo. Esto pudo ser debido a que durante el primer curso del grado, los estudiantes forman equipos de trabajo con más, o menos, aciertos, ya que no suelen conocerse entre sí, y que con el transcurso del tiempo se van consolidando y rehaciéndose en otros equipos de trabajo con diferentes componentes. Este fenómeno común para los primeros cursos de los grados tuvo mayor incidencia y problemática en el grupo C, donde se produjo este fenómeno en mayor número de ocasiones. Los tamaños de los grupos eran iguales, aunque al final hay una ligera diferencia de 2 en el número de estudiantes del grupo de control, como se muestra en la tabla 1.

#### 3.1.1. Comparación de las notas del examen en los tres grupos de Clase A, B, y C

En la tabla 2 se muestran los resultados de la comparación de medias de cada clase, obteniéndose que hay diferencias significativas entre las medias. En la prueba de comparación múltiple de Scheffé (tabla 3), se obtiene que las diferencias se dan entre todos los grupos de clase. Se puede decir, dado la significación y el signo de las diferencias de medias en la prueba Scheffé, que las medias de las notas de clases se encuentran en el siguiente orden  $A < C < B$ . Es decir, la media de la clase A es significativamente menor que la media de la clase C y que la media de la clase B y la media de la clase C es significativamente menor que la media de la clase B.

Tabla 1: Descriptivos de las notas de cada grupo en el examen

	N	Media	Desviación típica	Error típico	Intervalo de confianza media (95%)	
					Límite inferior	Límite superior
A	67	5,150448	1,8785967	0,2295071	4,692222	5,608674
B	69	7,656667	1,0151292	0,1222072	7,412806	7,900527
C	69	6,977101	1,2659129	0,1523980	6,672996	7,281207
Total	205	6,608829	1,7706570	0,1236680	6,364998	6,852661

En la gráfica 1 se muestra el diagrama de caja de las notas de cada uno de los grupos, en la que se puede observar como las notas en el grupo A tienen una mayor dispersión, mientras que las de los grupos B, y C están más próximas entre sí a la vez que son más altas, siendo este hecho más acentuado en el grupo B. La mayor homogeneidad de las notas en el grupo B hace que resulten, estadísticamente, valores extremadamente altos y bajos respecto a los de este grupo, quedando señalados por los círculos.

Tabla 2: Análisis de Varianza de las notas del examen

	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Inter-grupos	227,618	2	113,809	55,804	0,000
Intra-grupos	411,968	202	2,039		
Total	639,586	204			

### 3.2. Análisis cualitativo

Al analizar los criterios de evaluación que describen los estudiantes del grupo A, se obtiene que existe una mayor coincidencia entre las categorías de las evaluaciones de los estudiantes del grupo de control con las evidencias de las eRúbricas, cuando las tareas ofrecen mayor número de evidencias. La tabla 4 muestra los porcentajes de coincidencias entre categorías que los estudiantes del grupo de control expresan y las evidencias de esta, especialmente para las de las actividades 2 y 4 con 16 evidencias cada una, hay mayor porcentaje que en las eRúbricas para la actividad 1 que con 5, y la actividad 3 con 6 evidencias.

Si nos centramos exclusivamente en el análisis de las categorías que tienen una coincidencia con las evidencias de la eRúbrica, se observó que un mayor o menor número de equipos evaluados entre sí para cada actividad dentro del grupo A, no asegura mayor coincidencia entre las categorías encontradas y los criterios y evidencias de las eRúbricas utilizadas en el grupo experimental. Se tiene un porcentaje del 100%

Tabla 3: Prueba de comparación múltiple de Scheffé

(I) Clase	(J) Clase	Diferencia de medias (I-J)	Error típico	Sig.
A	B	-2,5062189	0,2449421	0,000
	C	-1,8266537	0,2449421	0,000
B	C	0,6795652	0,2431344	0,022

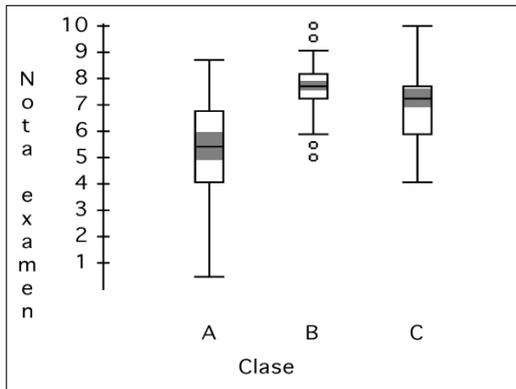


Gráfico 1: Diagrama de caja de las notas de cada grupo.

tanto en la actividad 1 con 15 equipos evaluados, como en la actividad 2 con la mitad de equipos evaluados. Este resultado también se produce en la actividad 3 con un mayor porcentaje (7 equipos y 85,67%) y un menor número de equipos frente a la actividad 4 (9 equipos y 66,66%). Es decir, el número de equipos evaluados en cada actividad no asegura el surgimiento espontáneo en el grupo A de valoraciones y criterios

**Tabla 4. Relación de categorías con las evidencias de las eRúbricas**

Número de Actividades	Número de evidencias de la eRúbrica	Número de categorías relacionadas con las evidencias	% de la Relación entre categorías de los estudiantes y las evidencias de la eRúbrica
1	5	2	40,00%
2	16	10	62,50%
3	6	3	50,00%
4	16	9	56,25%

más próximos o coincidentes con las evidencias de la eRúbrica.

Siguiendo con este análisis, podemos entender que el entrenamiento con una evaluación sin criterios y orientaciones del profesor de ningún tipo (como pudieran ser las eRúbricas) no asegura la capacidad necesaria en los estudiantes para evaluar de forma más objetiva y específica con el tiempo. Este hecho se observa igualmente en la tabla 5, donde se muestra la relación de estudiantes evaluadores identificados por su número de lista cuyas valoraciones, o criterios, coin-

**Tabla 5. Relación de estudiantes evaluadores dentro de las cuatro actividades**

Actividades	Relación de estudiantes evaluadores	Estudiantes evaluadores	%
En 1 Actividad	3-4-5-8-10-13-19-24-26-28-29-30-34-35-37-41-42-44-46-48-49-50-53-56-61-62-67-68-69-73-74-75-76-79-80	35	50,72%
En 2 Actividades	1-7-12-15-16-18-20-23-32-33-36-38-40-43-45-51-52-58-60-65-66-71-72	23	33,33%
En 3 Actividades	9-11-14-25-27-31-39-57-63-64-70	11	15,95%
En 4 Actividades	-	0	0%
Total		69	100%

ciden con las evidencias de las eRúbricas para las actividades 1, 2, 3 o 4. Observándose que no existe coincidencia de valoraciones por parte de los evaluadores pares en las 4 actividades de la asignatura, mientras que el porcentaje más alto es de 50,72% y se registra en las evaluaciones realizadas en la primera actividad.

#### 4. Discusión y conclusión

Existe un uso generalizado de las rúbricas como instrumentos para la evaluación de los resultados y calificaciones, en comparación con un uso más próximo a la evaluación formativa en sus diferentes modalidades. Este artículo está orientado a exponer los resultados desde un enfoque formativo de la evaluación, especialmente para la evaluación cooperativa en el trabajo en equipo. Aborda también una práctica poco difundida aún, como es el uso de las «eRúbricas federadas», que permiten estudiar mejor las variables que confluyen en el trabajo en equipo, por la facilidad de crear y exportar los datos digitales y por las tecnologías de federación que la sustentan, que facilitan la interoperatividad de diferentes herramientas. Igualmente, la investigación pretende buscar mayor fiabilidad y validez a estas prácticas, en consonancia con algunas

de las críticas encontradas en las revisiones realizadas (Reddy & Andrade, 2010), a la vez que, abrir nuevas líneas de investigación al destacar la existencia y posibilidad de estudiar en el futuro desde un marco conceptual más amplio, con los diferentes usos de la eRúbricas según las modalidades de aprendizaje / evaluación cooperativas / colaborativas.

Entre los resultados más destacados en este estudio se puede concluir que los grupos que han utilizado la eRúbrica para la evaluación cooperativa sobre el trabajo en equipo, cuando se enfrentan a las pruebas de examen obtienen resultados más altos y más homogéneos en sus calificaciones individuales, mientras que en el grupo de control las notas muestran mayor dispersión y hay calificaciones muy por debajo del aprobado. Lo que significa que los estudiantes del grupo de control al no disponer de los criterios tan concretos de la eRúbrica, disponen de menos elementos para la comprensión de las tareas y se ven más limitados para enfrentarse a las mismas. Esto se refuerza con los resultados del análisis cualitativo, donde el grupo de control a pesar de que aplicaba sus propios criterios que coincidían con más de un 40% con las evidencias de las eRúbricas, obtuvo peo-

res resultados en la prueba y mayor dispersión en las calificaciones. Con respecto al análisis de los criterios y valoraciones utilizados por los estudiantes del grupo de control, se concluye también que existe un mayor acercamiento a las evidencias de las eRúbricas, cuando se diseña un mayor número de evidencias en las tareas, por lo que se puede concluir que el diseño de tareas con más evidencias facilita el mejor resultado en las evaluaciones de los aprendizajes.

Desde ambos análisis —cuantitativo y cualitativo— se concluye que la eRúbrica ofrece un impacto positivo en la obtención de mejores logros de aprendizaje individual, debido principalmente a la concreción de los criterios y el aprendizaje de los mismos para realizar evaluaciones cooperativas de las tareas en equipo.

El estudio destaca el análisis de unas prácticas poco generalizadas, encontrando cómo las eRúbricas junto con la evaluación cooperativa facilitan las competencias que tendrán que desarrollar los estudiantes en algún momento de su vida profesional, donde tendrán que evaluar los trabajos de otros compañeros, aplicar criterios de calidad a los procesos y productos...

En suma, con estas metodologías y tecnologías nos adelantamos a esta realidad profesional desde la misma formación inicial. Quedan aún muchas más experiencias e investigaciones que validen los datos conseguidos en este trabajo y nos proporcionen mayor comprensión sobre estas metodologías y usos tecnológicos de la eRúbricas para la evaluación formativa.

El trabajo de Falchikov (2005) estudia la evaluación colaborativa y atiende a uno de los problemas de esta evaluación cuando no tenemos en cuenta los aspectos individuales de los equipos, variables que de alguna forma están influyendo en los resultados, tales como el género, la etnicidad, el nivel del curso y la edad, las experiencias previas... entre otras. En dicha obra recoge diversas fórmulas de diferentes autores para mitigar estas variables contaminantes, como «ponderar el factor individual igual a la calificación del esfuerzo individual dividido por la media de los esfuerzos de las calificaciones». En nuestro proyecto no hemos podido controlar el factor individual en el diseño. Esperamos para futuras ocasiones poder considerarlas, procurando en lo posible respetar la naturalidad de los grupos con diseños cuasi-experimentales y cualitativos.

### Apoyos

Proyecto titulado «Servicio federado de eRúbrica para la evaluación de aprendizajes universitarios». Plan Nacional I+D+i EDU2010-15432. (<http://eRubrica.org>). En la investigación se empleó la eRúbrica federada de Gtea (<http://gteavirtual.org/rubric>).

### Notas

<sup>1</sup> CSCL (Computer Supported Collaborative Learning).

<sup>2</sup> Entorno federado de Gtea (<http://gteavirtual.org>).

### Referencias

- ANDRADE, H. (2005). Teaching with Rubrics: The Good, the Bad, and the Ugly. *College Teaching*, 53 (1), 27-31. (DOI: <http://dx.doi.org/10.3200/CTCH.53.1.27-31>).
- BARTOLOMÉ, A. (2012). De la Web 2.0 al e-learning 2.0. *Perspectiva*, 30 (1), 131-153. (DOI: <http://dx.doi.org/10.5007/2175795X-2012v30n1p131>).
- BARTOLOMÉ, A., MARTÍNEZ, E. & TELLADO, F. (2012). Análisis comparativo de metodologías de evaluación formativa: diarios personales mediante blogs y autoevaluación mediante e-rúbricas. In C. Leite & M. Zabalza (Coords.), *Enseño superior. Inovação e qualidade na docencia*. (pp. 417-429). CIIIE: Porto.
- BLANCO, A. (2009). *Desarrollo y evaluación de competencias en educación superior*. Madrid: Narcea.
- BROWN, S. & GLASNER, A. (2003). *Evaluar en la Universidad*. Madrid: Narcea.
- CEBRIÁN-DE-LA-SERNA, M. (2011). Supervisión con e-portafolios y su impacto en las reflexiones de los estudiantes en el Practicum. Estudio de caso. *Revista de Educación*, 354, 183-208.
- COLÁS, P., BUENDÍA, L. & HERNÁNDEZ, F. (2009). *Competencias científicas para la realización de una tesis doctoral*. Madrid: Davinci.
- DE-MIGUEL, M. (2006). *Modalidades de enseñanza centradas en el desarrollo de competencias*. Oviedo: Universidad de Oviedo. (<http://goo.gl/1X1ESu>) (07-07-2013).
- DEL-POZO, J. (2012). *Competencias profesionales. Herramientas de evaluación y portafolio, la rubrica y las pruebas situacionales*. Madrid: Narcea.
- FALCHIKOV, N. (2005). *Improving Assessment Through Student Involvement*. New York (USA): Routledge.
- FALCHIKOV, N. & GOLDFINCH, J. (2000). Student Peer Assessment in Higher Education: A Meta-Analysis Comparing Peer & Teacher Marks. *Review of Educational Research*, 70, 3, 287-322. (DOI: 10.3102/00346543070003287).
- GÁMIZ-SÁNCHEZ, V., GALLEGO-ARRUFAT, M.J. & MOYA, E. (2012). Experiencias docentes de evaluación en metodologías activas con TIC: Análisis de casos en la Universidad de Granada. *II Congreso Internacional sobre evaluación por competencias mediante e-rúbricas*. Universidad de Málaga, Octubre 2012. (<http://goo.gl/fUY4Jt>) (16-08-2013).
- HAFNER, J.C. & HAFNER, P.H. (2003). Quantitative Analysis of the Rubric as an Assessment Tool: An Empirical Study of Student Peer-Group Rating. *International Journal of Science Education*, 25, 12, 1.509-1.528. (DOI: <http://dxdoi.org/10.1080/0950069022000038268>).
- HARGREAVES, E. (2007). The Validity of Collaborative Assessment for Learning. *Assessment in Education*, 14, 2, 185-199. (DOI: <http://dx.doi.org/10.1080/0950069022000038268>).
- JONSSON, A. & SVINGBY, G. (2007). The Use of Scoring Rubrics: Reliability, Validity and Educational Consequences. *Educational Research Review* 2, 130-144. (DOI: <http://dx.doi.org/10.1016/j.edurev.2007.05.002>).
- LÓPEZ-PASTOR, V. (2009). *Evaluación formativa y compartida en educación superior*. Madrid: Narcea.
- MARTÍNEZ, M.E., TELLADO, F. & RAPOSO, M. (2013). La rúbrica como instrumento para la autoevaluación: un estudio piloto. *Revista de Docencia Universitaria*, 11 (2), 373-390.
- MEEUS, W., PETEGEM, P. & ENGELS, A. (2009). Validity and Relia-

- bility of Portfolio Assessment in Pre-Service Teacher Education. *Assessment & Evaluation in Higher Education*, 34 (4), 401-413. (DOI: <http://dx.doi.org/10.1080/02602930802062659>).
- MORIL, R., BALLESTER, L. & MARTÍNEZ, J. (2012). Introducción de las matrices de valoración analítica en el proceso de evaluación del Practicum de los Grados de Infantil y de Primaria. *Revista de Docencia Universitaria*, 10 (2), 251-271.
- OSANA, H. & SEYMOUR, J. (2004). Critical Thinking in Preservice Teachers: A Rubric for Evaluating Argumentation and Statistical Reasoning. *Educational Research and Evaluation*, 10, 4-6, 473-498. (DOI: <http://dx.doi.org/10.1080/13803610512331383529>).
- PANADERO, E. & JONSSON, A. (2013). The Use of Scoring Rubrics for Formative Assessment Purposes Revisited: A Review. *Educational Research Review*, 9, 129-144. (DOI: <http://dx.doi.org/10.1016/j.edurev.2013.01.002>).
- PANADERO, E., ALONSO-TAPIA, J. & RECHE, E. (2013). Rubrics vs. self-assessment scripts effect on self-regulation, performance and self-efficacy in pre-service teachers. *Studies in Educational Evaluation*, 39 (3), 125-132. (DOI: <http://dx.doi.org/10.1016/j.stueduc.2013.04.001>).
- PRINS, F.J., SLUIJSMANS, D.M.A., KIRSCHNER, P.A. & STRIJBOS, J.W. (2005). Formative Peer Assessment in a CSDL Environment: A Case Study. *Assessment & Evaluation in Higher Education*, 30, 417-444. (<http://goo.gl/u2VOGY>). (DOI: 10.1080/02602930500099219).
- REDDY, Y. & ANDRADE, H. (2010). A Review of Rubric Use in Higher Education. *Assessment & Evaluation In Higher Education*, 35, 4, 435-448. (DOI: <http://dx.doi.org/10.1080/02602930902862859>).
- RODRÍGUEZ-GÓMEZ, G. & IBARRA-SÁIZ, M.S. (2011). *E-Evaluación orientada al e-aprendizaje estratégico en educación superior*. Madrid: Narcea.
- VICKERMAN, PH. (2009). Student Perspectives on Formative Peer Assessment: an Attempt to Deepen Learning. *Assessment & Evaluation in Higher Education*, 34, 221-230.
- VOOGT, J. & KNEZEK, G. (2008). *International Handbook of Information Technology in Primary and Secondary Education*. New York: Springer.
- ZABALZA, M. (2010). *Planificación de la docencia en la universidad*. Madrid: Narcea.