



Received: 14-08-2013
Reviewed: 17-09-2013
Accepted: 04-02-2014

RECYT Code: 22770
Preprint: 15-05-2014
Final Publication: 01-07-2014

DOI: 10.3916/C43-2014-15

Manuel Cebrián, José Serrano and Mayerly Ruiz
Malaga / Bucaramanga (Spain / Colombia)

eRubrics in Cooperative Assessment of Learning at University

Las e-rúbricas en la evaluación cooperativa del aprendizaje en la universidad

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+i¹ 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 answers of students in the control group with regard to the evaluation criteria that they used, in order to check for coincidences with the eRubric criteria used by students in the experimental groups.

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+i¹ cuyo objetivo general consiste en «Analizar el impacto de las e-rúbricas –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 ha utilizado e-rúbricas. Se observan diferencias en las notas de una prueba escrita común a todos, con mejores resultados en los grupos con e-rúbrica. 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 e-rúbricas usadas en los grupos experimentales.

Keywords / Palabras clave

Formative assessment, teamwork, cooperative learning, cooperative assessment, initial training, university education, skills, eRubrics.

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

Dr. Manuel Cebrian-de-la-Serna is Full Professor in the Faculty of Educational Sciences at the University of Malaga (Spain) (mcebrian@uma.es).

Dr. José Serrano-Angulo is Lecturer in the Faculty of Educational Sciences at the University of Malaga (Spain) (joseserrano@uma.es).

Mayerly Ruiz-Torres is Lecturer at the University of Santo Tomás de Aquino in Bucaramanga (Colombia) (2009-2011) and a PhD from the Faculty of Educational Sciences at the University of Malaga (mayerlyzulayr@hotmail.com).

1. Introduction

Among the different teaching methods developed in recent years, cooperative models that use technologies –CSCL–² (Voogt & Knezek, 2008) represent a deep renewal in education. In the field of university teaching design and planning, these methods (together with the use of technologies) have become increasingly important when it comes to centring teaching on student learning (Zabalza, 2010), while engaging students in their own learning process, especially when it comes to evaluation (Falchikov & Goldfinch, 2000; Brown & Glasner, 2003; Falchikov, 2005; Blanco, 2009, López-Pastor, 2009). As a result, different methods and ways of organising the teaching-learning process are planned based on the context of different universities (De-Miguel, 2006: 31). This is where the «teamwork» purpose brings all values and pedagogical principles together: «Students learn and assess collaboratively, by playing a more active and committed role in teaching and learning through technologies».

Human learning is fundamentally social; hence the construction of knowledge and collaborative learning should be a priority at all levels of education (Hargreaves, 2007). Nevertheless, considering that students and educational contexts do not always count on the necessary requirements to implement a collaborative model, a model of cooperative learning is most frequently used as a first step, resolving much of this starting situation, as it provides students with structure and guidance, while providing teachers with control.

When trying to extend collaborative learning to all stages of the teaching process, as in the case of evaluation, the need for guidance becomes more evident and crucial, and methods such as «teamwork» and «cooperative assessment» become important resources and techniques as a prelude to a model of collaborative assessment. This might be the reason why cooperative learning through teamwork is one of the most used methods in the promotion of skills development in all educational stages.

In any case, peer learning is especially beneficial when focused on the assessment process, where it is more often referred to as «collaborative assessment» by

academic literature (Blanco, 2009: 115; Brown & Glasner, 2003: 31; López-Pastor, 2009: 94), and also known as «co-assessment», «shared assessment», «peer assessment», etc. A more accurate conceptual definition is needed, as the terms used do not always differ from each other - as it is the case for cooperative vs. collaborative learning -. Cooperative assessment is more structured and guided than collaborative assessment.

While these practices are becoming increasingly widespread, criticisms of certain aspects are raised, including the following:

- The difficulty of carrying out an individualized follow-up and assessment of the skills acquired by the different team members.
- A review of the impact these methods have on student learning, in relation to new contexts and given the use of technology.
- How to approach what students need in order to achieve a collaborative assessment, which requires greater reflection and self-criticism.

One of the principles supporting collaborative assessment consists of involving all team members in defining the criteria by which proof of learning in the team projects will be evaluated. This is a rather communicative and participatory approach to evaluation, starting with the exchange and understanding of goals, objectives and procedures, and ending with the evaluation of processes and outcomes. Quality criteria and indicators are applied on results and on the process. While learning tasks are well-defined and structured, there are often difficulties in communication between teachers and students, especially when online teaching is involved. The final evaluation is often the only aspect that is understood, provided there has been a shared analysis.

This communication issue caused by technology is likely to be solved if teachers and students keep a permanent dialogue on quality indicators, criteria and how criteria can apply to the proof of learning of teamwork. In this type of evaluation, rubrics are one of the techniques or tools that can facilitate communication (Osana & Seymour, 2004; Jonsson & Svingby, 2007; Reddy & Andrade, 2010; Rodríguez Gómez & Ibarra Sáiz, 2011; Panadero & Jonsson, 2013), and they are called «eRubrics» in their digital version. One of the advantages of eRubrics is that, they allow teachers and students to share quality indicators, criteria and proof of learning when evaluating learning objectives (Andrade, 2005). Federated eRubrics are even more interactive, as they are federated as well as digital. Federation provides the ideal support for cooperation and collaboration among users, thus overcoming the difficulties of interoperability among tools, services, contexts and technological systems, located both inside and outside the educational institution itself.

Federated eRubrics play a double role in teaching. On the one hand, as a technological system, they represent an ideal support for improving communication and understanding of the assessment process, while facilitating teamwork. They are an essential tool in assessing e-Portfolios, considering the monitoring process required by the teacher-student interaction through which students are enabled to understand the quality indicators, criteria and proof of learning. This is especially true when distance and technology are involved, considering that institutions often have different technological systems. An example is found in the Practicum,

when students are distributed across different educational institutions, each with their own tools and technological systems (Meeusa, Petegema & Engelsb, 2009; Cebrián-de-la-Serna, 2011; Del-Pozo, 2012). On the other hand, as a technique and as a methodology, federated eRubrics facilitate formative assessment, because they require a clear definition of the level of learning standards and the implementation of task-related criteria. There is extensive literature on the impact of federated eRubrics, such as research conducted by Hafner & Hafner (2003) and Falchikov (2005), the so-called «deep and authentic learning» explained by Vickerman (2009), research on peer-assessment in technology-mediated collaboration environments -CSCL- (Prins, Sluijsmans, Kirschner & Strijbos, 2005), and a few studies on initial teacher training and acquisition of professional skills (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).

However, despite the results, such research must be cautiously considered. We should aim for a much bigger picture with meta-analysis, such as the one offered by Svingby & Jonsson (2007) or Reddy & Andrade (2010), where a general view of rubrics in university education is offered, emphasizing the positive perception of students towards the use of programmes, taken in conjunction with research showing the resistance of certain groups of teachers to use them. Additionally, there is research on the positive impact of rubrics on academic performance, despite other studies finding no such impact.

Certainly, more studies on the impact of rubrics are needed, despite this broad and extensive literature. Research is especially required in the field of cooperative and collaborative assessment, since, although eRubrics have already been studied from a collaborative assessment approach (Falchikov, 2005: 125), this has not been the case with all the products of the recent boom in new technologies. This is important in studying the impact of «federated eRubrics», as they are more interactive than paper rubrics, facilitating communication, cooperation and collaboration between students and teachers of different institutions. Therefore, the scope and impact of federated eRubrics on cooperative and collaborative teaching and learning models is still unknown. Thus, new research is needed to analyse the interactive and communicative functions offered by technologies and social networks (Bartolomé, 2012). In particular there is a requirement for more rigorous methods, for greater reliability and for checking the validity of the procedures from broader geographic and cultural perspectives, as suggested by Reddy & Andrade (2010).

In pursuit of this aim, the results presented below are part of a research project in which federation technologies in general and federated eRubrics in particular are used for educational purposes and intra- and inter-institutional collaboration. The latter is precisely the topic of the present research: cooperative peer-assessment and teamwork developed in the lab. The interoperability enabled by federation technologies was used for cooperation within the same institution. Students only needed to log in and out to access the tools and federated services available, namely: an institutional platform where task resources were uploaded and shared, a federated eRubrics service for cooperative assessment, a «federated

key» tool to upload and share large files, a «federated webquest» service to elaborate teaching materials, and a «federated Limesurvey» service to collect open assessments from the control group in order to contrast their results³.

2. Methodology

The use of rubrics to assess learning has been introduced in different subjects and university degrees, but their digital version -federated eRubric- is rarely used. Indeed, the innovation in this project lies in the lack of experience with these technologies. Likewise, a broad conceptual framework has been used to examine their impact, following the introduction of a new variable, which can play different roles according to whether the assessment is cooperative and/or collaborative (if it is cooperative, eRubrics are given by the teacher; if it is collaborative, eRubrics are negotiated).

While our research does not address all the possibilities in Chart 1, it does raise the need to answer the following questions: Does student academic learning improve when using cooperative assessment with eRubrics in teamwork? Which evaluation criteria are used by students in peer-assessment without the structure and guidance of eRubrics?

Drawing on these questions, the specific objectives of this project are as follows:

- To analyse the impact of eRubrics in academic learning by developing collaborative assessment methods and teamwork (cooperative assessment with eRubrics).
- To analyse the criteria and rating used by students in peer-assessment without guidance (cooperative assessment without eRubrics).

By answering these questions and by developing the two research objectives set, the researchers wish to show the usefulness and effectiveness of eRubrics as a tool and as a method for formative assessment, thus allowing for the improvement of student learning, the internalization of evaluation criteria and the application of these criteria.

The research was planned in two stages: in the first stage, the contents and functionality of the eRubrics were agreed upon and designed, using Limesurvey to experiment with the contents for the first time. In the second stage, after evaluating the contents of the rubrics and creating our own eRubric tool, the planned research design was applied. For this second stage a multi-method approach was used, due to both the characteristics of the objectives and the nature of the data to be collected. The reality could then be seen at a qualitative and quantitative level. To achieve the first specific objective, a quasi-experimental methodology was designed: one class group would not use eRubrics (control group) and their results would be compared with the other two groups that would (experimental groups). To achieve the second objective, a qualitative methodology was applied through content analysis, where assessments were extracted from the control group that did not use eRubrics.

The sample consisted of three randomly selected groups out of six class groups studying the subject «Information and Communication Technologies Applied to Education», from the Primary Education Degree in the Faculty of Educational Sciences at the University of Malaga during the 2011/12 academic year. The

three groups had 75 students each, and the context of the research design for both objectives was achieved by dividing each group into two sub-groups of 37 students each (six sub-groups in total), which were given two class hours to perform tasks and carry out peer-assessment in the computer labs. Therefore, research was conducted with 50% of the student population, i.e. 225 students: 75 students for the control group and 150 for the experimental groups. The contents of the eRubrics can be found in the public database of the tool by typing in the aforementioned course description.

The sample was selected by using the cluster sampling technique, where the sample unit was the class group. Differences between the control and experimental groups were minimized, randomly assigning the groups that would receive instruction and the group that would act as the control group, in order to achieve equality between the two, thus avoiding problems of internal and external validity (Colás, Buendía & Hernández, 2009). Groups B and C were the experimental groups and Group A was the control group.

In order to carry out the research, four tasks were conducted during the academic year, each with the same assessment methodology in the three class groups.

The methodology was directed by the same teacher in the three groups, following these steps:

- Two hours. Presentation of the task to the whole group (75 students), task completion and peer-assessment coordinated by the teacher.
- Two hours. Dividing each group into two smaller sub-groups (37 students approx) in the same computer lab, performing the same task but using different materials and examples. In order to perform the task, a file had to be downloaded from an online platform, completed and the team results uploaded so as to be shared. Once all teams had uploaded their tasks, access to the platform was open for all teams to download and assess tasks individually. At the end, the teacher closed the possibility of assessing, uploading and downloading tasks. He assessed all the teams and uploaded their task results to the platform.

The four tasks and their objectives were different. Group A and B had the same tasks. Team assessment was conducted by using a random formula provided by the teacher, although assessment between the same teams was avoided, despite the fact that assessment was anonymous in all cases. Tasks were performed in teams of 3-5 students, but assessment was individual, that is, each member of the team assessed the work of another team assigned by the teacher.

To collect quantitative data, four eRubrics were designed, one for each task. Students from the experimental groups carried out cooperative peer-assessment, which, together with the teacher's assessment, was included in an excel spreadsheet. Students were identified with a number in order to be able to compare their individual scores in the final test with peer-assessments made and received during the lab practice. In contrast, the control group did not use any eRubric to assess the work of their peers, only a questionnaire - federated Limesurvey - with a simple open question: «What did you think of the task this team has performed?». This approach to peer-assessment and criteria given by the teacher

sought a model of collaborative assessment, but was unsuccessful due to the lack of guidance, counselling and structure. Nevertheless, it has enabled us to get to know the arguments, criteria and thoughts of students in Group A, as a means (with further research) to pinpoint the requirements for collaborative assessment.

3. Analysis and results

Given that all students were identified (experimental and control), the results could be compared with other variables such as final scores and specific assessments of lab practice carried out at the end of the year. The practice test consisted of an individual test on a randomly chosen example from the four tasks designed to show the same skills worked in the lab practice, but with different materials from those used during the year. With regard to the first specific objective, the methodology used allowed us to analyse individual and group scores from the test and assessments of the four tasks conducted in the lab, both in Group A (control) as well as Groups B and C (experimental).

As for the second specific objective, the methodology used allowed us to compare the categories found in the content analysis of assessments from Group A (control) with eRubric criteria used by Groups B and C.

3.1. Quantitative analysis

For the quantitative analysis, the use or not of eRubrics by students and teacher was considered as the independent variable. As mentioned above, all groups took a final test at the end of the year. The scores from this test were considered as the dependent variable, thus examining possible differences among students' scores in the different class groups.

To contrast scores from the final test, an analysis of variance has been conducted in the three groups, as the two experimental groups showed different trends. This may be because students do not usually know each other in the first year of their Degree, so the groups they form are more or less successful. Over time groups consolidate and reshape in different work teams. This common phenomenon - which takes place in the first year of any degree - had a higher incidence and caused more problems in Group C, where it happened most frequently. The group sizes were the same, although in the end there is a slight difference of two in the number of students in the control group, as shown in Table 1.

	N	Mean	Standard Deviation	Standard Error	Confidence interval of 95% for the mean	
					Lower Limit	Upper Limit
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

3.1.1. Test Scores Comparison in the Three Class Groups: A, B and C

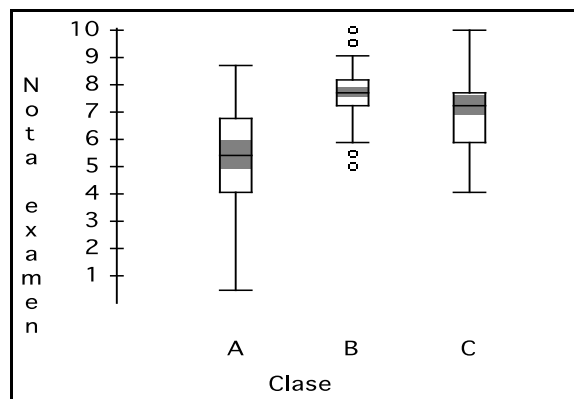
Table 2 shows the results of the comparison of each group's mean, showing significant differences. Scheffé's multiple comparison test has shown differences

among all class groups. In other words, there are differences between Group A's scores and Group B's scores, and between Group A's scores and Group C's scores. There are also differences between Group B's mean and Group C's mean. Given the significance and sign of mean differences in Scheffé's test, the means of the groups' scores are in the following order: $A < C < B$. That is, Group A's mean is significantly lower than Group C's mean and Group B's mean, while Group C's mean is significantly lower than Group B's mean.

	Sum of Squares	df	Root Mean Square	F	Sig.
Inter-groups	227.618	2	113.809	55.804	0.000
Intra-groups	411.968	202	2.039		
Total	639.586	204			

(I) Group	(J) Group	Mean Difference (I-J)	Standard Error	Sig.
A	B	-2.5062189*	0.2449421	0.000
	C	-1.8266537*	0.2449421	0.000
B	C	0.6795652*	0.2431344	0.022

Graph 1 shows the box plot of the scores in each group. As can be seen, Group A's scores show greater dispersion, while Groups B and C's scores are closer together and also higher, especially in Group B. The greater homogeneity of scores in Group B causes extremely high and low values in this group, marked below by the circles.



Graph 1. Box Plot of Scores in Each Group.

3.2. Qualitative analysis

When analysing the evaluation criteria described by students in Group A, a greater overlap is found between the categories of 1) control group's students' assessments and 2) eRubric responses, whenever tasks include a high number of responses. Table 4 shows the overlap percentage between categories expressed by students in the control group and eRubric responses. eRubrics of Tasks 2 and 4 (with 16 responses each) show a higher percentage than eRubrics for Task 1 (with 5) and Task 3 (with 6).

Table 4. Categories with eRubric responses

Number of Tasks	Number of eRubric responses	Number of categories related to responses	% of the relationship between students' categories and eRubric responses
1	5	2	40%
2	16	10	62.5%
3	6	3	50%
4	16	9	56.25%

When focusing exclusively on the analysis of the categories that match eRubric responses, it can be seen that a greater (or lesser) number of teams evaluated by each other for each task within the group A does not ensure a greater overlap between the categories found and the eRubric responses used in the experimental group. In other words, Table 5 shows a relationship and an equivalent rate of 100% in Activity 1 (with 15 assessed teams) and Activity 2 (with half of the teams assessed). This result also occurs in Activity 3, which has a higher percentage and a fewer number of teams in comparison to Activity 4. That is, the number of teams evaluated in each task does not ensure the spontaneous emergence in Group A of closer or coincidental criteria with eRubric responses.

The above analysis proves that teacher training with no evaluation criteria and no guidance from the teacher (as is the case of eRubrics) does not guarantee the necessary skills for students to evaluate in a more objective and specific manner over time. This fact can also be seen in Table 6, which shows the ratio of students-evaluators identified by their list number, whose assessments or criteria coincide with the eRubric assessments for tasks 1, 2, 3 or 4. There are no coincidences in the assessments of peer evaluators in all 4 tasks of the subject, while the highest percentage is 50.72% in the assessments of the first task.

Table 5. Ratio of Students-Evaluators in All 4 Tasks

Tasks	Ratio of Students-Evaluators	Number of Students-Evaluators	%
In 1 Task	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%
In 2 Tasks	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%
In 3 Tasks	9-11-14-25-27-31-39-57-63-64-70	11	15.95%
In 4 Tasks	-	0	0%
Total		69	100%

4. Discussion and conclusion

There is widespread use of the rubrics as a tool for evaluating results and scores, instead of for formative assessment in its various forms. The present paper seeks to present the results from a formative assessment approach, especially with regard to cooperative assessment in teamwork. It also addresses a practice, which is not yet well known: the use of «federated eRubrics», which enables researchers to better study the variables that come together in teamwork, due to the ease of

creating and exporting digital data and the federation technologies that support them, hence facilitating interoperability among different tools. Overall, research aims at developing a greater reliability and validity for these practices, in line with some of the reviews (Reddy & Andrade, 2010), while opening up new lines of research to highlight the possibility of studying from a broader conceptual framework in the future, by using eRubrics according to cooperative/collaborative learning/assessment modalities.

Among the most important results of this study, it is worth highlighting that the groups using eRubrics for cooperative assessment of teamwork have scored better and more homogeneous results than the control group in their individual marks when faced with the written test, whereas scores there were more dispersed. There were even scores well below the pass grade in the control group. This means that, in the absence of eRubric specific criteria, students in the control group had fewer elements with which to understand the tasks and more difficulties in facing them. This is reinforced by the results of the qualitative analysis, where the control group scored worse in the test and scores were more dispersed, even though they applied their own criteria, which matched the eRubric responses at over 40%. Regarding the analysis of the criteria used by students in the control group, we may also conclude that the higher the number of responses in the task, the higher approach to eRubric responses. As a consequence, the design of tasks with a high number of responses facilitates good results in learning assessment.

From both - quantitative and qualitative - analysis, eRubrics have proven to have a positive impact on achieving good individual learning results, mainly due to the specification of criteria for carrying out cooperative assessment of teamwork.

The present study highlights the analysis of uncommon practices. eRubrics, together with cooperative assessment, elicit skills that students will have to develop at some point in their career, as they will have to evaluate colleagues' work and apply quality criteria to processes and products. In short, these teaching methods and technologies anticipate the professional realities students will face from an educational perspective. There are many other experiments and much research that together serve to validate the results of this study and provide a greater insight into eRubric methods and their technological use for formative assessment.

Falchikov (2005) studies collaborative assessment and tackles some of the problems, when individual differences are not taken into account in the team. These variables - gender, ethnicity, educational level, age, previous experiences, and so on - somehow influence results. To minimize these confounders, he gathers various formulae used by different authors, such as «weighting the individual factor equal to the rating of the individual effort divided by the average of the efforts in the scores». In this project we were unable to control the individual differences in the design. However we hope to consider them on future occasions, while trying as much as possible to respect the naturalness of groups with quasi-experimental and qualitative designs.

References

- 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/10.5007/2175-795X.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.), *Ensino superior. Inovação e qualidade na docencia* (417-429). CIEE: 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*. Ediciones Universidad de Oviedo. (<http://goo.gl/1X1ESu>) (07-07-2013).
- Del-Pozo, J. (2012). *Competencias profesionales. Herramientas de evaluación y portafolio, la rúbrica 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.
- 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 Reliability 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 CSCL Environment: A Case Study, *Assessment and Evaluation in Higher Education*, 30, 417-444. (<http://goo.gl/u2VOGY>).
- 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.