

VOL. 19, Nº 2 (mayo-agosto 2015)

ISSN 1138-414X (edición papel)

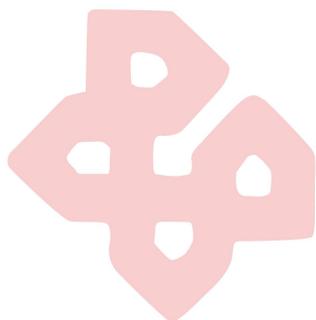
ISSN 1989-639X (edición electrónica)

Fecha de recepción 12/07/2014

Fecha de aceptación 7/01/2015

LAS RÚBRICAS COMO INSTRUMENTO DE EVALUACIÓN DE COMPETENCIAS EN EDUCACIÓN SUPERIOR: ¿USO O ABUSO?

The rubrics as an assessment tool of competency in higher education: use or abuse?



Elena Cano

Universidad de Barcelona

E-mail: ecano@ub.edu

Resumen:

Las rúbricas han emergido como un instrumento valioso para la evaluación de competencias. Pese a su interés, no pueden resolver todas las dificultades asociadas a los procesos de evaluación. En este artículo, se repasan algunos argumentos a favor y en contra del uso de rúbricas para realizar, finalmente, un análisis de las evidencias científicas derivadas de la investigación disponible en la actualidad sobre el uso de las rúbricas con el fin de saber si dichas evidencias ratifican o no los supuestos beneficios que suelen asociarse al empleo de las rúbricas. Los resultados muestran la utilidad de la rúbrica para promover la autorregulación del aprendizaje pero también arrojan la necesidad de acompañar el uso de las rúbricas con procesos de formación y con análisis de fiabilidad y validez del instrumento.

Palabras clave: Evaluación, competencias, rúbricas

Abstract:

The rubrics have emerged as a valuable tool for competency assessment. Despite their interest, they cannot solve all the difficulties associated with the assessment processes. In this paper, some arguments for and against the use of rubrics are reviewed. An analysis of the scientific evidence from the current available research on the use of rubrics is finally showed with the aim to know if these evidences confirm or contradict the supposed benefits usually associated with their use. The results show the usefulness of the rubric to promote self-regulated learning but also the need of training processes and of reliability and validity analysis.

Key words: Assessment, competencies, rubrics

1. Introducción

Hasta hace pocos años mucho se había escrito respecto a evaluación de los aprendizajes en educación superior pero pocas prácticas habían cambiado. Asistíamos a una separación entre los discursos teóricos y las prácticas de aula (Casanova, 1995; Mateo y Martínez, 2005; Prades, 2005). Sin embargo, desde la implantación del Espacio Europeo de Educación Superior y de las acciones estructurales y formativas asociadas a él, se han empezado a realizar algunos cambios sustantivos en los procesos de evaluación de los aprendizajes de los estudiantes universitarios (Dochy, Segers y Dierick, 2002; López Pastor, 2009).

El discurso de las competencias empieza a ir calando entre el profesorado. Ello lleva a pensar en términos de resultados de aprendizaje de los estudiantes (Kennedy, 2007; Paricio, 2010). Para verificar los mismos se instauran sistemas de evaluación alineados a dichos resultados y a las metodologías de aprendizaje que se proponen (Biggs, 2003), diversificando las estrategias e instrumentos de recogida de información. A la vez, otro de los elementos vinculados a la evaluación que ha variado ha sido la información acerca de los criterios de evaluación (Gibbs & Simpson, 2009; Stobart, 2010). Han pasado de ser algo reservado a considerar la conveniencia de que sean conocidos, transparentes y difundidos. En ese afán de dar a conocer los criterios y de mostrar también los niveles de alcance se han empezado a emplear usualmente las rúbricas (Sadler, 2009).

2. La rúbrica: un instrumento acorde con una visión de competencia

En primer lugar, parece necesario clarificar qué se entiende por rúbrica. En sentido amplio, se identifica con cualquier pauta de evaluación, preferentemente cerrada (tipo check-list o escala). En sentido estricto se asimila a una matriz de valoración que incorpora en un eje los criterios de ejecución de una tarea y en el otro eje una escala y cuyas casillas interiores están repletas de texto (no en blanco, como sucede con las escalas para que el evaluador señale el grado de adquisición de cada criterio). En cada casilla de la rúbrica se describe qué tipo de ejecución sería merecedora de ese grado de la escala.

Tabla 1
Ejemplo de rúbrica

Rúbrica de evaluación de las habilidades para avegar por Internet y hacer un buen trabajo de búsqueda				
Criterios de evaluación/ Niveles de expectativa	Pasable	Correcto	Bien	Excelente
Uso de internet	Necesita ayuda para usar los enlaces sugeridos o para navegar en una web	Usa ocasionalmente los enlaces sugeridos y navega con facilidad sin ayuda	Es capaz de utilizar los enlaces sugeridos para encontrar la información y navega fácilmente sin ayuda	Utiliza con éxito los enlaces sugeridos para encontrar la información y navega con facilidad sin ayuda
Informe/resumen	El informe es difícil de leer y cuenta con la información pedida	El informe incluye la mayor parte de información solicitada y cuesta comprenderlo	El resumen incluye la información solicitada y es comprensible	El resumen está claramente detallado e incluye toda la información solicitada
Cantidad de información	Uno o más de los temas no han sido tratados	Todos los temas han sido tratados y la mayor parte de las	Todos los temas han sido tratados y la mayor parte de las preguntas	Todos los temas han sido tratados y todas las preguntas

		preguntas han sido contestadas, como mínimo, con una frase	contestadas, como mínimo con dos frases, cada una	contestadas con tres o más frases cada una
Calidad de la información	La información tiene poco que ver con el tema principal	La información está relacionada con el tema principal pero no hay ejemplos	La información está claramente relacionada con el tema principal y se refuerza con algún ejemplo	La información está claramente relacionada con el tema principal y se refuerza con diversos ejemplos

Fuente: <http://www.xtec.cat/~cbarba1/webquest/avaluacio/Rnavegar.htm>

Sin embargo, no se sabe muy bien a qué obedece este término que muchos identificamos inicialmente con firmar o suscribir. El Diccionario de la Real Academia de la Lengua Española define rúbrica como “Rasgo o conjunto de rasgos de forma determinada, que como parte de la firma pone cada cual después de su nombre o título, y que a veces va sola, esto es, no precedida del nombre o título de la persona que rubrica”; como “Epígrafe o rótulo” (de hecho procede etimológicamente del latín ruber, rojo, porque en los libros antiguos solía escribirse con tinta roja) o como “Cada una de las reglas que enseñan la ejecución y práctica de las ceremonias y ritos de la Iglesia católica en los libros litúrgicos”. Por su parte, en el Collins Dictionary, las acepciones son las mismas pero se añade una cuarta definición: “instructions to a candidate at the head of the examination paper”. Probablemente de ahí derive su empleo.

Una rúbrica es, como se ha indicado, un registro evaluativo que posee ciertos criterios o dimensiones a evaluar y lo hace siguiendo unos niveles o gradaciones de calidad y tipificando los estándares de desempeño.

La primera autora en emplear en los 90’s el término en el sentido en que se conoce actualmente fue probablemente Goodrich (s/f) y después se ha venido popularizando hasta disponer en la actualidad de numerosa literatura sobre el tema. Sin embargo, la rúbrica es un instrumento acorde con una visión de competencia que no todos los autores comparten.

Muchos han sido los autores que han puesto de manifiesto las profundas divergencias en el modo de concebir las competencias (Rey, 1999; Díaz Barriga, 2006). Existen visiones recelosas de la lógica misma de las competencias aplicada a los estudios universitarios, como la de Barnett (2001), Bolívar (2008) o Carabaña (2011). Existen también posiciones que cuestionan la viabilidad de los enfoques basados en competencias en el marco de estructuras aún disciplinares (Escudero, 2009). Frente a ello se encuentran propuestas (Tourón, 2009; Jornet, González, Suárez y Perales, 2011; ANECA, 2014) que defienden el desarrollo operativo de este enfoque trasladando las competencias seleccionadas a resultados de aprendizaje de las asignaturas y dichos resultados a conductas observables. Y es en este marco donde la rúbrica parece haber tomado mayor sentido.

Probablemente es necesario admitir que las competencias, si se hallan integradas en las realizaciones que se piden a los estudiantes, deben de evaluarse de algún modo (Villardón, 2006; Tejada, 2011). Y para ello hay que revisar las estrategias e instrumentos de evaluación, así como los criterios que se emplean para valorar los diversos procesos de aprendizaje y los productos que generan. Sin embargo, es necesario establecer una frontera entre una visión de la competencia amplia, holística e integrada (Perrenoud, 2004) y la pedagogía por objetivos, a la que se puede acercar peligrosamente, como alertan Gimeno (2008), Cano (2008) o Carabaña (2011) y desestimar una visión conductista del enfoque de aprendizaje por competencias.

3. Cinco razones en contra de las rúbricas

Parece evidente que no existen instrumentos buenos o malos. Existen instrumentos coherentes o no con los resultados de aprendizaje de los que desean informar y, claro está, bien o mal elaborados desde el punto de vista técnico (Escudero, 2010). Pero ningún instrumento es bueno o malo en sí mismo. Por ello parece claro que no todas las dificultades asociadas a los procesos de evaluación pueden resolverse con las rúbricas y que las esperanzas idealizadas que aspiran a que la rúbrica aporte objetividad y rigor a la evaluación de ciertos aprendizajes que, por su naturaleza, son menos “objetivables” o más cualitativos, pueden generar frustración al no hallar respuesta a las inquietudes que las han generado.

No se puede estar taxativamente en contra de la rúbrica pero sí en contra de un mal uso y de un abuso de dicho instrumento. Por el excesivo protagonismo que está tomando la rúbrica en la actualidad, ofrecemos cinco motivos para desestimar las rúbricas (o cuanto menos, para dudar de ellas):

a) Porque no todo se puede encapsular

En primer lugar, la rúbrica puede utilizarse para comportamientos observables (en simulaciones, en prácticas de laboratorio, en exposiciones,...) o para evaluar trabajos (proyectos, monografías, ensayos,...). Puede utilizarse, como indica Moskal (2000) en muy diversas disciplinas escogiendo entre rúbricas analíticas u holísticas, entre generales o para una tarea puntual, puesto que el tema es poco relevante. Lo esencial es el tipo de aprendizaje que quiere contribuir a fomentar. Pese a ello, los aspectos actitudinales van a resultar difíciles de incorporar a esas tablas. Y, aunque existen propuestas, principalmente para la educación básica (ver, por ejemplo, http://www.fcaq.k12.ec/index.php?option=com_content&view=article&id=291&Itemid=396&lang=en o <http://www.getworksheets.com/samples/rubrics/>), en educación superior, si traducimos la actitud a términos observables pueden quizá darse situaciones donde se simule cierta actitud sin que se asuma como propia ni se incorpore a la práctica profesional futura. Las actitudes, por su naturaleza, requieren de tiempos largos y quizás de una recogida de información de carácter más cualitativo.

b) Porque hacer una buena rúbrica es difícil

En segundo lugar, la rúbrica resulta difícil de hacer (Moskal & Leydens, 2000). Es cierto que disponemos de numerosos ejemplos de rúbricas ya elaboradas pero quizás no se ajusten a nuestro contexto o nuestra asignatura. El hecho de seleccionar los criterios pero, sobre todo, de redactar el texto de las casillas, que debe de ser descriptivo y en ningún caso valorativo (“lo hace bien”, “excelente”, “adecuado”,...) resulta muy difícil. Sólo si los textos describen un comportamiento o una realización se reduce el margen de subjetividad de los evaluadores, que es una de las ventajas que tradicionalmente se asigna a la rúbrica (sobre todo en procesos de evaluación entre iguales). Pero, como hemos indicado, ese tipo de texto es difícil de conseguir sin acabar en cuantificaciones, que bien pudieran llegar a pervertir el sentido de la rúbrica.

c) Porque hacerla para una actividad es un coste excesivo en términos de tiempo

En tercer lugar, el coste de elaboración de la rúbrica en términos de tiempo, por lo que se ha expuesto, es muy alto. Por ello elaborarla para una actividad puntual de una asignatura parece claramente ineficiente.

d) Porque se hace para los alumnos y no con los alumnos

En cuarto lugar, el instrumento cambia pero la dinámica de evaluación unidireccional parece no variar. El profesorado construye o selecciona la rúbrica, la muestra (en algunas ocasiones) a los estudiantes y la aplica o la hace aplicar a los estudiantes (Cebrián, 2010; Ibarra, Rodríguez & Gómez, 2012). La participación de los estudiantes se limita, como mucho, a esta última actividad, lo que puede llevar a un conocimiento superficial de los criterios.

e) Porque para su misma finalidad existen alternativas eficaces

Básicamente las finalidades de la rúbrica son dar a conocer los criterios y disminuir la subjetividad de ciertas evaluaciones, en especial asistiendo a los procesos de evaluación entre iguales o de autoevaluación, para garantizar una cierta homogeneidad en los criterios que se emplean (Valle y Manso, 2013; Panadero y Jonsson, 2013).

Para ello existen alternativas más eficientes, como check-list (algunos aprendizajes básicos debe asegurarse que se dan, independientemente de la frecuencia o intensidad con que se produzcan), anotaciones cualitativas, feedback interactivo (durante la ejecución de una tarea) o elaboración de portafolios (con selección argumentada de evidencias de aprendizaje), entre otros.

4. Cinco motivos para apostar por las rúbricas

Frente a las críticas vertidas, numerosos autores defienden la utilización de las rúbricas y sus beneficios (Mertler, 2001; Alsina, 2010; García y Terrón, 2010; Del Pozo, 2012). Si se rescatan los principales argumentos a favor del empleo de las rúbricas, se pueden hallar los siguientes:

a) Por el valor formativo (versus sumativo) y formador

Aunque cada vez más las rúbricas se están empleando para calificar, las escalas inicialmente asociadas a las rúbricas eran de naturaleza cualitativa y permitían dar un feedback al estudiante, de modo que pudiese mejorar sus procesos en el futuro. Es necesario reivindicar ese potencial formativo y rescatar las posibilidades que la rúbrica ofrece para orientar a los estudiantes hacia qué se espera de ellos, cómo hacer un buen trabajo, qué es lo más relevante y, por tanto, dónde hay que poner el énfasis (Blanco, 2008).

En esta línea es interesante rescatar el estudio de Reddy & Andrade (2010), quienes comparan el uso formativo de las rúbricas con su uso para la clasificación y la rendición de cuentas. Tras una extensa revisión, no hallaron resultados concluyentes, puesto que, en efecto, los estudios realizados por Reitmeier, Svendsen & Vrchota (2004) y por Petkov y Petkova (2006) sugirieron que la participación de los estudiantes en el desarrollo y uso de rúbricas antes de una tarea iba asociada con mejoras en el rendimiento académico, pero Green & Bowser (2006), por el contrario, no hallaron diferencias en la calidad de los trabajos realizados por los alumnos con y sin rúbricas. La implicación parece ser que simplemente la entrega de una rúbrica no puede llevar a esperar ilusoriamente que tenga un impacto en el trabajo de los estudiantes: Los estudiantes deben aprender a usar activamente una rúbrica y sólo si lo hacen sus resultados pueden mejorar sustancialmente.

A este efecto positivo para los estudiantes hay que añadir que también resulta beneficiosa para el propio docente al facilitarle la tarea orientadora:

“La rúbrica tiene un doble valor en el uso que le damos cuando trabajamos con ella en nuestra práctica educativa. De una parte, es una herramienta de evaluación que debe entenderse en un contexto diferente al de la evaluación convencional. La rúbrica no sólo pretende evaluar los conocimientos del alumnado, sino que, además, debe servir como herramienta de reflexión que le permita tomar conciencia de lo aprendido. De otra parte, también sirve al alumnado como guía para cumplimentar las partes en las que se estructura una actividad. Precisamente, esta última función apoya la acción tutorial del docente” (Torres & Perera, 2010: 148).

b) Por la posibilidad de guiar un proceso (versus valorar un producto)

Uno de los primeros campos en los que aparecieron las rúbricas fue aplicadas a las webquest (véase, por ejemplo, el material disponible en http://www.ite.educacion.es/formacion/materiales/59/cd/pdf/m5_ceg.pdf o <http://platea.pntic.mec.es/~erodri1/TALLER.htm#rubrica>). Quizá es uno de los ejemplos más claros respecto a su uso durante el proceso, como guía (Picket & Dodge, 2007; Quintana & Higuera, 2009), frente al uso para la evaluación de productos.

c) Por el valor de construirla (versus consumirla)

La rúbrica puede constituir la base para la participación de los estudiantes. En lugar de ser aplicadores “pasivos” de los criterios establecidos por el profesorado, pueden construir colaborativamente los criterios que son relevantes para valorar la calidad de un trabajo. Alguna experiencia al respecto muestra el interés y los beneficios de esta práctica en términos de apropiación de los criterios por parte del alumnado y de mejora del rendimiento (Puigdelívol y Cano, 2011; Puigdelívol, García y Benedito, 2012).

d) Por el valor de la rúbrica de titulación (versus la rúbrica de asignatura)

Si bien una rúbrica para una asignatura puede constituir, como ya se ha señalado, un coste excesivo, una rúbrica de titulación puede ser un buen modo de graduar a lo largo de los cursos de una titulación, las consecuciones progresivas que se exigen respecto a ciertas competencias, especialmente transversales, que involucran a gran parte de las materias del grado (Solé, 2011).

e) Por la existencia de algunos ejemplos valiosos (versus la necesidad de construirlas)

Destacamos en último lugar la existencia de propuestas bien elaboradas por algunos profesores y profesoras, que pueden ser recuperadas y adaptadas, minimizando el coste de la rúbrica. Destacamos los ejemplos de Martínez et al. (2009) o el grupo de competencias de la URV (2009).

5. Evidencias científicas respecto a los beneficios y validez del uso de rúbricas

Existen numerosas aportaciones acerca de cómo una mayor comprensión de los criterios de evaluación genera un mejor rendimiento académico tanto en educación básica (Fuchs & Fuchs, 1986) como en educación superior (O'Donovan, Princes & Rust, 2001; Rust, Princes & O'Donovan, 2003; O'Donovan, Princes & Rust, 2001; 2004).

Especialmente ilustrativos acerca de la necesidad de conocer y aplicar los criterios son los trabajos de Boud (Boud, 1995; Boud, 1998; Boud, Lawson & Thomson, 2013; Boud & Molloy, 2013). La capacidad de que cada estudiante sea juez de la calidad de su propio trabajo ha de constituir, según estos autores, una competencia a promover en la educación superior por considerarse la base para aprendizajes futuros. Ello ha lograrse intentando que cada estudiante sea capaz de reorientar sus procesos de aprendizaje a partir de analizar sistemáticamente lo que hace y cómo lo hace, con la ayuda de sus colegas y del establecimiento de criterios claros, conocidos, transparentes y, en la medida de lo posible, negociados.

Sin embargo, las investigaciones que justifiquen los beneficios de las rúbricas en términos de motivación o de aprendizaje no son tan frecuentes (ausencia ya constatada por Reddy & Andrade en 2010). Si bien hasta el momento en este artículo se han ido desgranando algunas de las valoraciones que se suelen realizar más frecuentemente sobre las rúbricas, se hace necesario acudir a las aportaciones de la investigación reciente sobre rúbricas para comprobar si existen evidencias científicas acerca de los efectos de la rúbrica sobre el aprendizaje. Un análisis interesante sobre dichos efectos en la educación básica es el realizado por Marzano & Haystead (2008) sobre 14 estudios que emplean rúbricas. En educación superior los estudios más rigurosos sobre los efectos de las rúbricas son los realizados por Roblyer & Wiencke (2004); Jonsson y Svingby (2007); Serrano & Cebrián (2011); Panadero, Alonso-Tapia y Huertas (2012); Panadero, Romero y Strijbos (2013); Panadero, Alonso-Tapia y Reche (2013), Panadero y Jonsson (2013) o más recientemente por Valverde y Ciudad (2014).

En un primer estudio, Roblyer & Wiencke (2004) hallan buenos índices de validez de las rúbricas en los cursos en línea. Éstas guardan correlaciones relativamente altas con las evaluaciones post-curso. Resulta interesante el análisis que realizan de los diversos criterios (la participación de los estudiantes posee alta relación con el rendimiento, mientras que la tecnología empleada posee el índice de correlación más bajo, hallazgos que enfatizan la importancia de la dinamización que realizan los tutores y de la interacción que generan). Hallan también una alta fiabilidad entre evaluadores en los puntajes de la rúbrica, puesto que en todos los cursos los resultados son consistentes. Un tercer resultado que merece la pena resaltar es la necesidad de formar a los estudiantes en el uso de la rúbrica, explicando claramente los criterios con el fin de reducir las variaciones en la comprensión de los mismos por parte de los estudiantes.

Tres años después Jonsson & Svingby (2007) consideran que usualmente se señala que la rúbrica sirve para juzgar competencias complejas y que aumenta la consistencia de las puntuaciones pero para asegurar si dichas aseveraciones se desprenden de la literatura científica analizan 75 estudios. Tras considerar las dificultades en la definición de la rúbrica (aludiendo a los estudios de Hafter & Hafner, 2003) y tras caracterizar las rúbricas como instrumento vinculado a la evaluación auténtica, realizan un análisis de los 75 artículos (en su mayoría posteriores al año 2000, pertenecientes a diversas revistas y relativos a todas las etapas educativas) y estudian la fiabilidad y la validez de los hallazgos. De su estudio se desprende (Tabla 2) que con la rúbrica sí se da fiabilidad intra-calificadores (básicamente demostrados con alpha de Cronbach), fiabilidad entre calificadores (estudiando los porcentajes de concordancia total y parcial y el coeficiente Kappa de Cohen) -aunque este grado depende del nivel de explicitación de la rúbrica- y aumento de la consistencia en las puntuaciones, aunque ello depende de si es para un uso individual o del conjunto de la clase. Estas mismas conclusiones son reforzadas por el reciente estudio de Valverde y Ciudad (2014)

quienes obtienen un índice alpha de Cronbach aceptable pero sólo cuando todos los estudiantes hacen la misma tarea.

Tabla 2

Panorama de los estudios que informan sobre las medidas de fiabilidad de las rúbricas analizadas en los estudios consultados por Jonsson & Svingby

<i>Método</i>	<i>Número de estudios</i>
<i>Concordancia</i>	
<i>Porcentaje de concordancia total</i>	18
<i>Porcentaje de concordancia parcial</i>	14
<i>Kappa de Cohen</i>	4
<i>Otros</i>	7
<i>Total^a</i>	27
<i>Medidas de consistencia</i>	
<i>.....Correlación de Person</i>	4
<i>.....Alpha de Cronbach</i>	8
<i>.....Correlación de Spearman</i>	6
<i>.....Otras</i>	9
<i>Total^b</i>	24
<i>Estimación de la medición</i>	
<i>.....Generalización de la teoría</i>	15
<i>.....Modelo de Rasch</i>	3
<i>.....Otros</i>	1
<i>Total</i>	19
TOTAL	46

Fuente: Jonsson & Svingby (2007: 134)

Notas de los autores:

(a) Algunos artículos informan sobre más de un método, por ejemplo, de la concordancia total y parcial. Esta categoría resume el número total de artículos que informan sobre cada medida de fiabilidad, sin contar el artículo dos veces.

(b) Algunos artículos informan sobre más de una medida de fiabilidad, por ejemplo, las estimaciones tanto de consistencia como de medida. Esta categoría resume el número total de artículos que informan sobre la consistencia sin contar el artículo dos veces.

Sin embargo, se demuestra también que no es más válido puntuar con una rúbrica que sin ella. Jonsson & Svingby (2007) estudian varios tipos de validez (Tabla 3). En primer lugar, la validez de contenido (grado en que éste es representativo y relevante del conocimiento o habilidades que desea evaluar); de generalización (grado en que las interpretaciones son coincidentes entre grupos, tareas, ocasiones,...) y externa (relación de la puntuación que la rúbrica arroja con otras medidas relevantes relativas al constructo analizado, como las evaluaciones al final de curso o las pruebas extras estandarizadas). Estudian, por otra parte, otros aspectos de la validez de la rúbrica, analizando, entre otras cosas, si es empleada por los profesionales o si ha sido adoptada por otras instituciones. Entre los aspectos que analizan dentro de esta validez, con la ayuda de jueces expertos, está la propia estructura (la alineación de las dimensiones y escala y la rúbrica), la sustantividad de la rúbrica (la validez del “modelo teórico” que hay tras la rúbrica -en este extremo los autores alertan de la dificultad de la rúbrica para procesos de pensamiento en lugar de aplicarlas siempre a productos, como señalan Osama y Seymour, 2004- y la validez “consecuente”, que incluye evidencias de las implicaciones de la interpretación de la puntuación a corto y largo plazo sobre futuras rúbricas, algo así como el impacto en siguientes instrumentos.

Tabla 3

Panorama de los estudios que informan sobre las medidas de validez de las rúbricas analizadas en los estudios consultados por Jonsson & Svingby

Aspectos de la validez ^a	Número de estudios
Validez de contenido	10
Validez de generalización	3
Validez externa	15
Validez de estructura	7
Validez sustantiva	1
Validez consecuencial	2
TOTAL^b	25

Fuente: Jonsson & Svingby (2007: 136).

Notas de los autores:

(a) Adaptado de Messick (1996).

(b) Algunos artículos informan sobre más de un aspecto. Esta categoría resume el número total de artículos que informan sobre cada aspecto sin contar el artículo dos veces.

Justamente la despreocupación por la validez y la fiabilidad es la principal crítica que realizan Reddy & Andrade (2010) al hallar que, en los estudios que revisan, no se describe el proceso de desarrollo de las rúbricas para establecer su calidad. Los pocos estudios que informan de la validez (Moni, Beswick & Moni, 2005; Green & Browser, 2006; Lapsley & Moody, 2007) no abordan la validez de contenido ni de constructo ni de criterio. Por otra parte también son escasos los estudios de fiabilidad entre evaluadores. Los artículos que analizan (Simon & Forgette-Giroux, 2001; Hafner & Hafner, 2003; Dunbar, Brooks & Kubicka-Miller 2006) tienden a mostrar que las rúbricas pueden conducir a una interpretación relativamente común del rendimiento de los estudiantes siempre que los evaluadores posean formación ad hoc.

Otros hallazgos vienen dados por la aportación de Blanco (2011) o por el estudio de Panadero, Alonso-Tapia y Huertas (2012) que comparan, mediante ANOVA, el uso de rúbricas y guiones en los procesos de autorregulación del aprendizaje y la autoeficacia con estudiantes de secundaria. Llegan a la conclusión de que los guiones permiten más la autorregulación que las rúbricas, de que las rúbricas facilitan el feedback (porque los estudiantes tienen dificultades en entender el lenguaje académico) y de que el uso de ambos instrumentos aumenta significativamente el aprendizaje. Estos autores llegan a concluir que aumenta la fiabilidad de las notas (sobre todo si las rúbricas son analíticas) y que proporcionan juicios válidos. Sin embargo, por sí mismas no promueven el aprendizaje.

Otros estudios interesantes son los que focalizan el análisis en el empleo de la rúbrica para el desarrollo de la autorregulación, que parece ser uno de los usos más adecuados de la rúbrica, ya sea a partir de la autoevaluación (Panadero, Alonso-Tapia y Reche, 2013) o de la evaluación entre iguales (Serrano & Cebrián, 2011; Panadero, Romero y Strijbos, 2013) puesto que estimula el aprender a aprender. En el estudio de Serrano & Cebrián (2011), 70 estudiantes emplean eRubricas para evaluar el trabajo en equipo de sus compañeros en otras universidades. Los autores insisten en la importancia de intercambiar criterios de calidad entre estudiantes de diversas localizaciones por considerar que discutir los estándares en un mundo global debe ser una práctica profesional a la que nos aproximemos desde la educación superior. Los resultados del empleo de las rúbricas durante 3 años, comparando además los resultados de la evaluación entre iguales con los de la evaluación por parte del profesorado en el último curso, muestran una mejora paulatina, con lo que se evidencia la progresiva internalización o interiorización de los criterios de las rúbricas por parte del alumnado. Pese

a ello, se detectan dificultades para la autoevaluación que debieran conducir al empleo de otras estrategias, como los diarios.

Por su parte, en el estudio de Panadero, Romero & Strijbos (2013) la mitad de un grupo de 209 estudiantes evalúa a sus compañeros mediante rúbrica y la otra mitad sin ella. Se comprueba que, pese a que todos los estudiantes sobrevaloraron a sus compañeros, los que usaron la rúbrica consiguieron mayor validez. Se recoge la percepción de los estudiantes por el uso de la rúbrica constatando cómo aumenta su sensación de justicia y de objetividad entre quienes la aplican. Este extremo ya había sido hallado parcialmente por O'Donovan, Price & Rust (2001) y es, de nuevo confirmado por la investigación de Vickerman (2009) y de de Van de Watering, Gijbels, Dochy & Van der Rijt, (2008), quienes insisten en la importancia de considerar la percepción de los estudiantes acerca de los procesos de evaluación y hallan la vinculación entre las preferencias de los estudiantes respecto a los instrumentos y las sensaciones de rigor u objetividad que aportan (frente, por ejemplo, a los portafolios o la evaluación entre iguales, que son rechazados por la falta de claridad en sus criterios).

Finalmente, otra revisión interesante es la realizada por Panadero y Jonsson (2013), quienes realizan un análisis de 21 estudios sobre rubricas llegando a determinar:

The use of rubrics may mediate improved performance through (a) providing transparency to the assessment, which in turn may (b) reduce student anxiety. The use of rubrics may also (c) aid the feed-back process, (d) improve student self-efficacy, and (e) support student self-regulation; all of which may indirectly facilitate improved student performance. Consequently, there are a number of different, but not necessarily unconnected, ways in which the use of rubrics may support student learning” (Panadero & Jonsson, 2013: 140).

Los autores estudian las posibles correlaciones entre las características de las rúbricas y el rendimiento académico o los procesos de autorregulación, llegando a descubrir que las rúbricas poseen un gran potencial para influir sobre el aprendizaje de los estudiantes positivamente pero que existen múltiples modos de emplearlas y que hay numerosos factores que moderan el efecto de las rúbricas, factores sobre los que habrá que seguir investigando, con lo que las evidencias acerca de los beneficios de las rúbricas tampoco resultan concluyentes.

6. Conclusiones

Una de las dificultades tradicionales asociadas a los procesos de evaluación ha sido el desconocimiento de los criterios, que ha llevado a los estudiantes a percibir la evaluación como un suceso arbitrario e incluso azaroso (Salinas, 2002). Ello ha generado que los estudiantes prefieran actividades de evaluación “tradicionales”, con las que se hallan más familiarizados y cuyos criterios, aunque sea implícitamente, conocen mejor (Van De Watering, Gijbels, Dochy y Van Der Rijt, 2008). Por ello la evaluación alternativa apuesta firmemente por la difusión y publicidad de los criterios de evaluación. Sin embargo, ambos polos pueden acarrear dificultades. Pese a asumir que el conocimiento de los criterios puede llevar a resultados más ajustados a procesos y productos de calidad, hay que reflexionar sobre el grado de explicitación de tales criterios, de modo que no anulen la autonomía ni la creatividad (Stobart, 2010). En los últimos años se ha progresado en la necesidad de establecer criterios llegando en ocasiones a un grado de concreción excesivo e innecesario.

Algunos autores (Jornet, González, Suárez y Perales, 2011) se muestran a favor de la estandarización para facilitar los procesos de certificación de las competencias, con la finalidad de que, de este modo sean algo relevante y no complementario o accesorio a los conocimientos disciplinares. Sin embargo, otros autores (Bolívar, 2008) consideran que la estandarización conduce a un encorsetamiento que poco tiene que ver con la naturaleza de los aprendizajes universitarios, que debieran llevar a desarrollar capacidades como profesional pero también como ciudadano, ni con la dinámica de la enseñanza universitaria, que debiera caracterizarse por la flexibilidad y, sobre todo, por la coherencia entre las propuestas metodológicas, los resultados de aprendizaje que se persiguen y el modelo de profesor que se es.

En cualquier caso, acerca de lo que sí hay unanimidad es acerca de la necesidad de aumentar la investigación acerca de las rúbricas. Reddy & Andrade (2010) sugieren cuatro líneas de actuación. En primer lugar, proponen realizar investigaciones con metodologías rigurosas (sugieren investigaciones cuasi-experimentales) y medidas de análisis robustas acerca de la fiabilidad y la validez de las rúbricas. Muchos de los estudios que hallan son realizados por los autores en sus propias aulas, por lo que aconsejan repeticiones de dichas prácticas en otros contextos (o en cursos subsiguientes, como es el caso de Valverde y ciudad, 2014) para fortalecer la credibilidad de los resultados. En esta línea de aumentar el rigor, en segundo lugar, proponen además realizar nuevos estudios sobre la validez y la fiabilidad de las rúbricas (aspecto en el que insisten también Panadero & Jonsson, 2013), informando sobre los resultados de las pruebas piloto, sobre los procedimientos de creación, sobre la fiabilidad de la puntuación, sobre la formación del evaluador (y el grado de confiabilidad entre evaluadores) y sobre posibles correlaciones entre las rúbricas y otras medidas de rendimiento. En tercer lugar, Reddy & Andrade (2010) sugieren también ampliar las perspectivas geográficas y culturales, puesto que la mayor parte de investigación se sitúa en Estados Unidos y consideran que estudios internacionales podrían contemplar las diferencias culturales y sobre teorías de la educación existentes en cada contexto. En cuarto y último lugar, se sugiere el empleo de la rúbrica no tanto para la calificación como para el aprendizaje y la autorregulación. En esta línea, aportaciones de la investigación reciente sobre rúbricas desarrolladas, sobre todo por Panadero (Panadero, Alonso y Huertas, 2012; Panadero y Jonsson, 2013) demuestran los efectos de la rúbrica sobre los procesos de autorregulación de los aprendizajes pero alertan de las limitaciones de las investigaciones. Estas limitaciones pasan por la multiplicidad de elementos que concurren en los procesos de evaluación y en el empleo de las rúbricas y que pueden afectar los resultados que éstas arrojan sobre el aprendizaje. Por ello no se puede afirmar que el uso de las rúbricas en sí mismo genera mayor rendimiento. En cambio sí hay un aspecto especialmente positivo vinculado a las rúbricas: quizás sea el empleo para el desarrollo de la autorregulación uno de los usos más adecuados de la rúbrica, ya sea a partir de la autoevaluación (Panadero, Alonso-Tapia y Reche, 2013) o de la evaluación entre iguales (Panadero, Romero y Strijbos, 2013) pero, en cualquier caso, con una dimensión formativa y que oriente a los estudiantes hacia el desarrollo de la capacidad autónoma y sostenible de aprendizaje a lo largo de la vida.

Referencias bibliográficas

Alsina, J. (Coord) (2010). *La evaluación por competencias en la universidad*. Barcelona: ICE de la UB (Colección Cuadernos de Docencia Universitaria). Recuperado de <http://diposit.ub.edu/dspace/bitstream/2445/14922/6/qdu18.pdf>

- ANECA (2014). *Guía de apoyo para la redacción, puesta en práctica y evaluación de los resultados del aprendizaje*. Madrid: ANECA.
- Barnett, R. (2001). *Los límites de la competencia: el conocimiento, la educación superior y la universidad*. Madrid: Gedisa.
- Biggs, J. (2003). *Calidad del aprendizaje universitario. Cómo aprenden los estudiantes*. Madrid: Narcea.
- Blanco, A. (2008). Las rúbricas: un instrumento útil para la evaluación de competencias. En L. Prieto (Coord.). *La enseñanza universitaria centrada en el aprendizaje* (171-188). Barcelona: Octaedro/ICE Universidad de Barcelona.
- Blanco, A. (2011). Tendencias actuales de la investigación educativa sobre las rúbricas. En K. Buján; I. Rekalde y P. Aramendi (coords.). *La evaluación de competencias en la educación superior. Las rúbricas como instrumento de evaluación*. Sevilla: Eduforma, pp. 59-74
- Bolívar, A. (2008). El discurso de las competencias en España: educación básica y educación superior. *Revista de Docencia Universitaria RED-U*. número monográfico 2. Recuperado de <http://red-u.net/redu/index.php/REDU/article/view/70/pdf>
- Boud, D. (ed.) (1995). *Enhancing learning through self-assessment*. London: Routledge Falmer.
- Boud, D. (1998). *Developing student autonomy in learning*. London: Kogan Page.
- Boud, D. y Molloy, E. (Eds.) (2013). *Feedback in Higher and Professional Education: Understanding it and doing it well*. London: Routledge.
- Boud, D.; Lawson, R. y Thomson, D. (2013). Does student engagement in self-assessment calibrate their judgement over time? *Assessment & Evaluation in Higher Education*, 38 (8), 941-956.
- Cano, E. (2008). La evaluación por competencias en la educación superior. *Profesorado. Revista de currículum y formación del profesorado*, 12 (3) (Monográfico Las reformas educativas basadas en el enfoque por competencias: Una visión comparada). Disponible en: <http://www.ugr.es/~recfpro/rev123COL1.pdf>
- Carabaña, J. (2011). Competencias y universidad, o un desajuste por mutua ignorancia. *Revista de Orientación Pedagógica*, 63(1), 15-31.
- Casanova, M^a. A. (1995). *Manual de evaluación educativa*. Madrid: La Muralla.
- Cebrián, M. (2010). *La evaluación formativa con el e-portafolios y la e-rúbrica*. Recuperado de http://vicadc.uvigo.es/opencms/export/sites/vicadc/vicadc_gl/documentos/ciclos_conferencias/Material.ePor_eRubric.pdf
- Del Pozo, J.A. (2012). *Competencias profesionales: herramientas de evaluación: el portafolios, la rúbrica y las pruebas situacionales*. Madrid: Narcea.
- Díaz Barriga, A. (2006). El enfoque de competencias en la educación. ¿Una alternativa o un disfraz de cambio? *Perfiles educativos*, vol. XXVIII, núm. 111, pp. 7-36.
- Dochy, F., Segers, M. y Dierick, S. (2002). Nuevas vías de aprendizaje y enseñanza y sus consecuencias: una nueva era de evaluación. *Boletín de la Red Estatal de Docencia Universitaria* 2(2): 13-29.
- Dunbar, N.E., Brooks, C.F. y Kubicka-Miller, T. (2006). Oral communication skills in higher education: Using a performance-based evaluation rubric to assess communication skills. *Innovative Higher Education*, 31, (2), 115-28.
- Escudero, J.M. (2009). Las competencias profesionales y la formación universitaria: posibilidades,

- riesgos. *Revista de docencia universitaria*, 2, 7-26. Recuperado de <http://revistas.um.es/redu/article/view/35231/33751>
- Escudero, T. (2010). *Sin tópicos ni malentendidos: fundamentos y pautas para una práctica evaluadora de calidad en la enseñanza universitaria*. Zaragoza: ICE de la Universidad de Zaragoza.
- Fuchs, L. S. y Fuchs, D. (1986). Effects of systematic formative evaluation: A metaanalysis. *Exceptional Children*, 53(3), 199-208.
- García, M^a J. y Terrón, M^a J. (2010). Desarrollo de Recursos Docentes para la Evaluación de Competencias Genéricas. *ReVisión*, Vol 3, No 2. Recuperado de <http://www.aenui.net/ojs/index.php?journal=revisión&page=article&op=view&path%5B%5D=70&path%5B%5D=113>
- Gibbs, G. y Simpson, C. (2009). *Condiciones para una evaluación continuada favorecedora del aprendizaje*. Barcelona: ICE y Ediciones Octaedro.
- Gimeno, J. (2008). *Enseñar por competencias. ¿Qué hay de nuevo?* Madrid: Morata.
- Goodrich, H. (s/f). Understanding rubrics. Recuperado de <http://learnweb.harvard.edu/alps/thinking/docs/rubricar.htm>
- Green, R. y Bowser, M. (2006). Observations from the field: Sharing a literature review rubric. *Journal of Library Administration*, 45, (1-2), 185-202.
- Grupo de Competencias de la URV (2009). *Competències transversals (s/f). Guia per treballar i avaluar les competències* Tarragona: URV. Recuperado de http://www.urv.cat/media/upload/arxius/SRE/AVA.%20X%20COMPETENCIES/guia_compet_transversals_1_06_09.pdf
- Hafner, J. C., y Hafner, P. M. (2003). Quantitative analysis of the rubric as an assessment tool: An empirical study of student peer-group rating. *International journal of science education*, 25, 1509-1528.
- Ibarra, M.S.; Rodríguez, G. y Gómez, M.A. (2012). La evaluación entre iguales: beneficios y estrategias para su práctica en la universidad. *Revista de Educación*, 359, pp. 206-231.
- Jonsson, A.; Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2, 130-144.
- Jornet, J.M.; González, J.; Suárez, J.M. y Perales, M^a J. (2011). Diseño de procesos de evaluación de competencias: consideraciones acerca de los estándares en el dominio de las competencias. *Bordón* 63 (1), 125-145.
- Kennedy, D. (2007). *Redactar y utilizar resultados de aprendizaje*. Cork: CCU. Recuperado de <http://www2.udla.edu.ec/archivos/MANUAL%20para%20REDACTAR%20Y%20UTILIZAR%20RESULTADOS%20DE%20APRENDIZAJE.pdf>
- Lapsley, R., y R. Moody. 2007. Teaching tip: Structuring a rubric for online course discussions to assess both traditional and non-traditional students. *Journal of American Academy of Business*, 12, (1), 167-72.
- López Pastor, V. (2009). *Evaluación formativa y compartida en educación superior*. Madrid: Narcea.
- Martínez, M. [et al.]. (2009). *Guia per a l'avaluació de les competències en els laboratoris de ciències i tecnologia*. Barcelona AQU Recuperado de <http://hdl.handle.net/2117/9421> y en http://upcommons.upc.edu/e-prints/bitstream/2117/9421/1/doc_62206423_1.pdf

- Martínez Rojas, J.G. (2008). Las rúbricas en la evaluación escolar: su construcción y su uso. *Avances en Medición*, 6, 129-138. Recuperado de http://www.freewebs.com/cesarmerino/Store%20of%20pub/TPVNM_rev.pdf
- Marzano, R.J. y Haystead, M.W. (2008). *Making Standards Useful in the Classroom*. Denver: ASCD.
- Mateo, J. y Martínez, F. (2005). *La evaluación alternativa de los aprendizajes*. Barcelona: ICE de la UB. Disponible en: <http://www.ub.edu/ice/sites/default/files//docs/qdu/3cuaderno.pdf>
- Mertler, C. A. (2001). Designing scoring rubrics for your classroom. *Practical Assessment, Research & Evaluation*, 7(25). Recuperado de <http://pareonline.net/getvn.asp?v=7&n=25>
- Messick, S. (1996). Validity of performance assessments. In G. Phillips (Ed.), *Technical issues in large-scale performance assessment* (pp. 1-18). Washington, DC: National Center for Education Statistics.
- Moni, R.W., Beswick, E. & Moni, K.B. (2005). Using student feedback to construct an assessment rubric for a concept map in physiology. *Advances in Physiology Education*, 2, 197-203.
- Moskal, B.M. (2000). Scoring rubrics: what, when and how? *Practical Assessment, Research & Evaluation*, 7 (3). Disponible en: <http://pareonline.net/getvn.asp?v=7&n=3>
- Moskal, B. M. y Leydens, J.A. (2000). Scoring rubric development: validity and reliability. *Practical Assessment, Research & Evaluation*, 7(10). Recuperado de <http://PAREonline.net/getvn.asp?v=7&n=10>
- O'Donovan, B., Price, M y Rust, C. (2001). The student experience of criterion-referenced assessment (through the introduction of a common criteria assessment grid). *Innovations in education and Teaching international*, 38 (1), 74-85.
- O'Donovan, B., Price, M. y Rust, C. (2004). Know what I mean? Enhancing student understanding of assessment standards and criteria. *Teaching in Higher Education*, 9 (3), 325-35.
- Osana, H. P., y Seymour, J. R. (2004). Critical thinking in preservice teachers: A rubric for evaluating argumentation and statistical reasoning. *Educational Research and Evaluation*, 10, 473-498.
- Panadero, E., Alonso-Tapia, J. y Huertas, J.A. (2012). Rubrics and self-assessment scripts effects on self-regulation, learning and self-efficacy in secondary education. *Learning and Individual Differences*. 22, 806-813.
- Panadero, E., Alonso-Tapia y 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*. 30, pp. 125-132.
- Panadero, E. y Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, 9, 129-144.
- Panadero, E.; Romero, M. y Strijbos, J.W. (2013). The impact of a rubric and friendship on peer assessment: Effects on construct validity, performance, and perceptions of fairness and comfort. *Studies in Educational Evaluation*, 39, 195-203.
- Paricio, J. (2010). *Un modelo de guía docente desde los resultados de aprendizaje y su evaluación*. Zaragoza: ICE de la Universidad de Zaragoza. Recuperado de <http://www.unizar.es/ice/images/stories/calidad/Modelo%20gu%C3%ADa%20docente.pdf>
- Perrenoud, P. (2004). *Diez nuevas competencias para enseñar*. Barcelona: Graó.

- Petkov, D. y Petkova, O. (2006). Development of Scoring Rubrics for IS Projects as an Assessment Tool. *Issues in Informing Science and Information Technology*, 3, 499-510.
- Pickett, N., & Dodge, B. (2007). *Rubrics for Web lessons*. Recuperado de <http://webquest.sdsu.edu/rubrics/weblessons.htm>
- Prades, A. (2005). *Competències i EEES, aprenentatge de competències i avaluació de les competències*. Barcelona: Universitat de Barcelona.
- Puigdellívol, I. y Cano, E. (2011). La rúbrica en los estudios de educación. En K. Buján, I. Rekalde y P. Aramendi (Coord.) *La evaluación de competencias en la Educación Superior: Las rúbricas como instrumento de evaluación* (pp. 131-156). Madrid: Mad Eduforma.
- Puigdellívol, I.; García, N. y Benedito, V. (2012). Rúbricas, más que un instrumento de evaluación. En E. Cano (Coord.), *Aprobar o aprender. Estrategias de evaluación en la sociedad red* (pp. 67-94). Barcelona: Transmedia XXI.
- Quintana A. J. y Higuera A. E. (2009). *Las webquests, una metodología de aprendizaje cooperativo, basada en el acceso, el manejo y el uso de información de la red*. Barcelona: Octaedro. Recuperado de <http://www.octaedro.com/ice/pdf/11CUADERNO.pdf>
- Raposo, M. y Martínez, E. (2011). La Rúbrica en la Enseñanza Universitaria: Un Recurso Para la Tutoría de Grupos de Estudiantes. *Formación universitaria*, 4(4), 19-28. Recuperado en 23 de febrero de 2014, de [http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-50062011000400004](http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-50062011000400004&lng=es&tlng=es).
- Reddy, M. y Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education*, 35/4, 435-448.
- Reitmeier, C.A., Svendsen, L.K. & Vrchota, D.A. (2004). Improving oral communication skills of students in food science courses. *Journal of Food Science Education*, 3, 15-20.
- Rey, B. (1999). *Las competencias transversales en cuestión*. Santiago de Chile: Documento de la Escuela de Filosofía Universidad ARCIS.
- Rey, B. et al. (2012). *Les compétences à l'école. Apprentissage et évaluation*. Bruxelles (Belgium): De Boeck.
- Roblyer, M.D. & Wiencke, W.R. (2004). Exploring the interaction equation: validating a rubric to assess and encourage interaction in distance courses. *Journal of Asynchronous Learning Networks*, 8 (4).
- Rodríguez, G.; Ibarra, M.S. y Gómez, M.A. (2011). e-Autoevaluación en la universidad: Un reto para profesores y estudiantes. *Revista de Educación*, 356. Septiembre-Diciembre, pp. 401-430.
- Rust, C.; Price, M. & O'Donovan, B. (2003). Improving Students' Learning by Developing their Understanding of Assessment Criteria and Processes. *Assessment & Evaluation in Higher Education*, Vol. 28, No. 2, 147-164.
- Sadler, D.R. (2009). Indeterminacy in the Use of Preset Criteria for Assessment and Grading in Higher Education. *Assessment & Evaluation in Higher Education*, 30, 175-194.
- Salinas, D. (2002). *¡Mañana examen! La evaluación: entre la teoría y la realidad*. Barcelona: Graó.
- Sánchez-Elvira, A. (Coord.) (2010). *Del diseño a la evaluación en competencias genéricas: Análisis empíricos e intervención mediante rúbricas*. Proyecto 2009-0102 del programa Estudios y Análisis. Madrid. Recuperado de http://portal.uned.es/pls/portal/docs/PAGE/UNED_MAIN/LAUNIVERSIDAD/VICERRECTORADOS/

