

Rúbricas para la evaluación de proyectos de Aprendizaje-Servicio en los estudios universitarios de las artes¹

Rubrics for assessing service-learning projects in university arts courses

José Luis Menéndez Varela

Universitat de Barcelona

menendez@ub.edu

Fecha de recepción del artículo: mayo 2013

Fecha de publicación: noviembre 2013

Resumen

El artículo aborda el diseño de un sistema de rúbricas para evaluar proyectos de Aprendizaje-Servicio en asignaturas de formación básica de los grados de Bellas Artes, Diseño, Conservación-Restauración e Historia del Arte. En dicho proceso intervinieron los profesores responsables de la materia y tres estudiantes de máster y doctorado que tuvieron a su cargo un programa de tutoría entre iguales dirigido a estudiantes de reciente ingreso. Se describen y justifican los criterios utilizados y el proceso de elaboración. El artículo concluye con la presentación de las rúbricas.

Palabras clave: rúbricas, aprendizaje-servicio, educación superior, artes.

Abstract

The paper describes the design of a rubric scoring system to assess service-learning projects in the basic training subjects of several EHEA bachelor's degrees (Fine Arts, Design, Conservation-Restoration and History of Art). Participants in the design process included the lecturers who were responsible for the subject and three master's degree and doctoral students who were in charge of a peer tutoring programme for first-year students. The author describes and justifies the criteria that were used and the design process. The paper concludes with a presentation of the rubrics.

Keywords: rubrics, service-learning, higher education, arts.

¹ El presente trabajo forma parte de los resultados de los proyectos REDICE12-1622-01 y 2012PID-UB/067 financiados respectivamente por el ICE y el PMID de la Universitat de Barcelona.

Introducción

Las instituciones de educación superior se encuentran inmersas en una evaluación continuada de sus titulaciones. Mientras que este proceso de evaluación está en marcha desde los años 90 del siglo pasado en la universidad anglosajona, en el sistema universitario español ha cobrado importancia a partir de la implantación de las nuevas titulaciones de grado y máster, y se espera que alcance su pleno desarrollo en los próximos años. Así lo demuestra el hecho de que actualmente asistimos a una reorganización de los programas de doctorado y al establecimiento de sistemas de seguimiento, verificación y acreditación específicos. En esta coyuntura, las agencias de acreditación han implantado programas de evaluación estandarizados de amplia envergadura que son de obligado cumplimiento para el mantenimiento de las titulaciones. Aunque la necesidad de cumplir con los estándares de acreditación ha obligado a prestar más atención a la misión educativa de las universidades, no es menos cierto que dicha necesidad amenaza con desvirtuar el propósito educativo de los programas de verificación. La realidad es que la valoración de los entornos de enseñanza-aprendizaje es a menudo postergada por la cumplimentación de los informes oficiales correspondientes mediante estrategias más o menos sofisticadas.

En estos momentos, comienzan a observarse las repercusiones negativas del escaso reconocimiento de la docencia que ha caracterizado a la universidad hasta épocas demasiado recientes. Sin menoscabar otros aspectos concurrentes de contrastada importancia, no puede olvidarse que uno de los factores que explican el limitado papel desempeñado por la comunidad universitaria en el establecimiento de los programas de verificación ha sido la desconfianza de los beneficios de los sistemas de evaluación tradicionales sobre el aprendizaje de los estudiantes (Bahous, 2008; Cox, 1995; Klenowky, 2002). Ello ha traído como consecuencia que las instituciones de educación superior han recurrido mayormente a la compilación de datos en busca de un elevado impacto sobre la opinión pública, mientras se dejaba en un segundo término la explicación de los desafíos educativos que la sociedad les demanda (Rhodes, 2010) y de las actuaciones que se han emprendido con tal efecto. La desconfianza sobre la evaluación de los estudiantes está provocando también que el examen de los programas bascule hacia una evaluación del profesorado y se diluya así su objetivo principal de impulsar y valorar el aprendizaje de los estudiantes (Weld, 2002).

El actual modelo basado en competencias ha puesto definitivamente de manifiesto la complejidad del proceso educativo. No solo se trata de que los estudiantes demuestren lo que conocen y son capaces de hacer; además, estos resultados deben constituir aprendizajes que les permitan resolver los retos que deben afrontar como ciudadanos y profesionales. No se trata de acreditar que los estudiantes adquieren determinada información, habilidades o actitudes, sino de evidenciar una movilización integrada de estos aspectos que permita a los egresados una actuación competente en contextos diversos. La complejidad de este reto debería mover a que los sistemas de evaluación se centraran en la detección de una mejora continuada de la efectividad de entornos

didácticos y curricula; y con esta finalidad, promover una recopilación sistemática, un análisis y un uso de los datos para justificar que el desempeño de los estudiantes se ajusta a lo previsto. En una línea semejante a la señalada por Walvoord, Bardes, y Denton (1998), para que esto sea posible se precisa que la valoración continuada de los resultados de aprendizaje de los estudiantes se realice concertadamente a nivel del aula, del departamento y de la titulación, en observancia del marco general establecido a nivel institucional.

Revisión de la literatura especializada

Las rúbricas como recursos para la mejora de la enseñanza y el aprendizaje

Al hacerse explícito que el centro de la enseñanza y el aprendizaje es la movilización de los recursos de variado signo incorporados por los estudiantes, se entiende que el nuevo paradigma de evaluación se presente como el eje del proceso educativo y se conciba como un proceso continuado de seguimiento, impulso y valoración de los desempeños de los estudiantes que son significativos para la sociedad. En el contexto de la evaluación del desempeño o evaluación auténtica, las rúbricas han ido adquiriendo un papel destacado, y algunos autores han enfatizado particularmente el valor de las rúbricas en el aprendizaje basado en competencias (Arter y McTighe, 2001; Maxwell, 2010).

Muchos autores han definido las rúbricas como esquemas (Brookhart, 1999; Moskal, 2000), escalas (Baughin, Brod, y Page, 2002; Díaz Barriga, 2004) o guías (Mertler, 2001; Wolf y Goodwin, 2007) de puntuación vinculadas a una evaluación del desempeño de los estudiantes. En estos casos, en que se acostumbra a presentar las rúbricas como instrumento de puntuación, lo más importante es que se destaca su asociación con la evaluación del desempeño; una evaluación que se caracteriza por ser continuada, atender a aprendizajes social y profesionalmente relevantes, incluir procesos de realimentación e impulso de nuevo aprendizaje, explicitar anticipadamente los criterios y los niveles de desempeño, favorecer una responsabilidad compartida entre profesores y estudiantes, e integrar múltiples procedimientos de evaluación y evaluadores.

La aportación más relevante de las rúbricas a la evaluación del desempeño es que incorpora una descripción cualitativa de los criterios de evaluación y de los niveles de consecución que gradúan los resultados del objeto de evaluación (Andrade, 2000; Arter y Chappuis, 2007; Reddy y Andrade, 2010; Stiggins, 2001). Es fácil de comprender hasta qué punto las rúbricas responden a la necesidad de una evaluación continuada y de una realimentación que se orienta hacia la construcción de nuevos aprendizajes. Y es fácil comprender también que las rúbricas alcanza sus mejores beneficios educativos cuando son puestas en conocimiento de los estudiantes antes del inicio del proceso evaluador (Moskal y Leydens, 2001; Simon y Forgette-Giroux, 2001). Primero, estudiantes y profesores disponen de una referencia fundamental para enjuiciar los resultados obtenidos, co-

mentar los puntos débiles del trabajo realizado, establecer líneas futuras de mejora y valorar en consecuencia el progreso de los aprendizajes. Segundo, posibilitan procesos de auto-evaluación y evaluación entre iguales sobre la base de criterios reconocidos por el ámbito disciplinario, especialmente si los estudiantes han participado de alguna manera en el diseño y/o en la discusión del significado, la finalidad, la utilidad y el uso de las rúbricas. Tercero, favorece que el estudiante tome decisiones responsables acerca del valor que confiere a la asignatura y sienta las bases para un diálogo al respecto con el profesor. Cuarto, permiten a los estudiantes regular los esfuerzos, planificar las actividades a llevar a cabo y organizar todos los medios necesarios en cada fase del proceso de aprendizaje.

Puesto que la educación superior se orienta hacia resultados de alto nivel cognitivo, debe asumirse que la valoración de dichos resultados es un proceso de inferencia a partir de aspectos que no siempre son objeto de evidencia directa y que el proceso de aprendizaje no sigue una tendencia lineal y progresiva. En estas circunstancias, se hace todavía más evidente la importancia de disponer de información útil sobre los objetivos establecidos, sobre los aspectos fundamentales que deben ser considerados para su consecución, sobre los diferentes niveles en que pueden clasificarse los resultados y sobre modelos que ejemplifiquen dichos niveles. Y aquí, la información será útil cuando sea capaz de reflejar la complejidad de los aprendizajes, sea capaz de respetar los diferentes modos en que los objetivos pueden ser alcanzados y, al mismo tiempo, sea capaz de establecer pautas que ayuden a los estudiantes a desarrollar el proceso de aprendizaje adecuado.

Las rúbricas son auténticos recursos didácticos porque su utilidad como plantilla de puntuación debe ser sobrepasada por su valor como concreción del estrecho vínculo existente entre enseñanza, aprendizaje y evaluación. Las rúbricas contribuyen a alcanzar lo que Biggs (1996) denominó la alineación constructiva entre aprendizaje, enseñanza y evaluación. Sin embargo, Pintrich (2002) distinguió entre objetivos de consecución —*performance goals*—, en los que el fin es evitar los errores para alcanzar la mejor calificación posible, y objetivos de aprendizaje —*learning goals*—, en los que la finalidad es decididamente cognitiva. Las rúbricas pueden servir a ambos propósitos. De hecho, la descripción de los niveles de consecución entraña siempre el riesgo de que sean interpretadas por profesores y estudiantes de manera prescriptiva y, por consiguiente, se corre el peligro de que impulsen un enfoque superficial frente a un enfoque profundo del aprendizaje (Entwistle, 1997; Marton y Säljö, 1997). Para conjurar este peligro, es imprescindible que el diseño de las rúbricas venga precedido de una reflexión sobre con qué finalidad, cómo, por quién y en qué contexto van a ser utilizadas (Turley y Gallagher, 2008). Wolf y Stevens (2007) advirtieron hasta qué punto las rúbricas pueden actuar como una camisa de fuerza, y señalaron que el auténtico reto es diseñar una rúbrica que clarifique lo que es valorado de un producto o desempeño sin desvirtuar su naturaleza. Por ello, recomendaron la utilización de las rúbricas solo en el caso de las actividades de aprendizaje más importantes y complejas.

La utilidad de las rúbricas descansa sobre tres pilares. En primer lugar, depende de que ayuden a los estudiantes a comprender los objetivos de aprendizaje establecidos y los criterios que dan cuenta de las actividades que deberán realizar (Reddy y Andrade, 2010; Wolf y Goodwin, 2007). En segundo, deriva de su potencial diagnóstico para identificar el estado del estudiante en el proceso de aprendizaje que le ocupa. Esta función debe estar al alcance del estudiante para poder regular su progreso y determinar en cada momento el esfuerzo necesario. Se comprende entonces la importancia de prestar una atención escrupulosa a la redacción: claridad, precisión y adecuación del lenguaje a todos los participantes (Montgomery, 2002; Torrie, 2007). Las rúbricas no van dirigidas ni primero ni principalmente a los profesores sino a los estudiantes, y la elección del lenguaje más apropiado a estos últimos constituye uno de los elementos más difíciles del diseño de las rúbricas y una de las causas más habituales de los problemas de validez (Moni, Beswick y Moni, 2005; Tierney y Simon, 2004). Pero esa capacidad diagnóstica debe también asistir al profesor para entender los puntos fuertes y débiles del trabajo de los estudiantes, imprimir las modificaciones correspondientes a su planteamiento didáctico inicial y acompasar el avance del curso al progreso en los aprendizajes (Reynolds et al., 2009; Welch y James, 2007; Wolf y Goodwin, 2007). En tercer lugar, la utilidad de las rúbricas reside en que establecen las bases para una evaluación más transparente y justa (Andrade y Du, 2005; Reddy y Andrade, 2010; Wolf y Goodwin, 2007), en la medida en que la explicitud de criterios y niveles de desempeño reduce la interferencia de otro tipo de consideraciones más subjetivas y menos defendibles en relación con la actividad evaluada. En realidad, las tres dimensiones anteriores vienen a remarcar el elemento determinante de la enseñanza y el aprendizaje basado en competencias: la necesidad de la práctica de una reflexión sobre la práctica. Con la excepción de las rúbricas concebidas para evaluar la práctica reflexiva de los estudiantes (Parkes y Kajder, 2010; Welch y James, 2007), las rúbricas no dan noticia específica de este aspecto pero, si han sido bien diseñadas, sí que establecen fundamentos sólidos para que esta práctica tenga lugar al establecer el entorno de trabajo adecuado: unos objetivos, unas actividades y un sistema de evaluación orientados hacia aprendizajes de alto nivel cognitivo que son comprendidos y compartidos por los estudiantes.

El diseño de las rúbricas

No hay duda de que, en el tema de las rúbricas, el aspecto que mayor acuerdo suscita es el proceso de diseño. Es lógico que en todos los casos se aluda a la secuencia siguiente:

- Revisión sistemática de la literatura y búsqueda de modelos.
- Determinación de la finalidad de las rúbricas y su alineación con los objetivos y actividades que serán valoradas.
- Establecimiento de los criterios de evaluación y los niveles de consecución.
- Testeo.

Es también unánime el reconocimiento de que los puntos críticos son los mencionados en segundo y tercer lugar. A ellos debe supeditarse la revisión crítica de la literatura y los modelos ya existentes; y de ellos debe dar cuenta el análisis de los resultados obtenidos en la fase piloto de su aplicación. En lo concerniente a la alineación con los objetivos y las actividades, Beauchamp, McConaghy, Parsons, Sanford, y Ford (2000) subrayaron la importancia de decidir cuál es la finalidad de la evaluación, de determinar los objetivos de aprendizaje perseguidos y los aspectos observables de los mismos que serán considerados, y de desglosar dichas actividades en las tareas que comportan. Como advirtieron Moskal y Leydens (2001), si alguno de los objetivos no está representado, la rúbrica no proporciona la evidencia necesaria para examinar el objetivo dado; y si alguno de los criterios no está relacionado con los objetivos, la adecuación de las rúbricas queda en cuestión. En cuanto a los criterios y niveles de consecución, Wiggins (1998) estableció la importancia de responder a las preguntas de qué criterios darán cuenta de los logros de los estudiantes, qué evidencias hay que buscar y en qué tipo de actividades, cómo se presentan los diferentes niveles de calidad, cómo cada uno de estos niveles deben ser descritos y diferenciados, cómo determinar qué puntuación merece cada caso y cómo aclarar el significado de la puntuación. Por lo general, la literatura recomienda que la construcción de las rúbricas sea una empresa colectiva (De La Cruz Flores, Díaz-Barriga Arceo y Abreu Hernández, 2010; Mertler, 2001; Suskie, 2004; Wolf y Stevens, 2007) y que en ella participen los estudiantes. Su participación no solo es importante porque mejora la comprensión del sistema de evaluación y fomenta un mayor compromiso en el proceso de aprendizaje; también es útil para enriquecer el proceso de elaboración con aportaciones realizadas desde un enfoque que ya no está al alcance del profesor. Puesto que la mejor utilización de las rúbricas incluye su uso por los estudiantes, su participación es crucial en la redacción de las descripciones cualitativas a fin de adaptarlas a sus necesidades y registros idiomáticos (Turley y Gallagher, 2008). Por último, una elaboración colectiva en la que participen grupos heterogéneos en diferentes momentos de la creación de las rúbricas sostiene los procesos de validación de los resultados obtenidos.

En la elaboración de las rúbricas el número de criterios y niveles de consecución no es una cuestión menor, y algunos aspectos importantes deben ser atentamente considerados. Algunos autores refieren un total de entre tres a seis criterios y argumentan que estos no deben ser tantos que convierta la evaluación en una actividad abrumadora ni tan escasos que impidan valorar aspectos cardinales del fenómeno observado (Suskie, 2004; Wolf y Stevens, 2007). No puede olvidarse aquí que la causa más habitualmente esgrimida por los profesores renuentes a utilizar las rúbricas es el tiempo necesario en su diseño y empleo. Y no puede olvidarse tampoco hasta qué punto el número de criterios es un factor decisivo en la evaluación de un objeto complejo, especialmente si tiene una naturaleza fugaz como es el caso de una exposición oral. El número de los niveles de desempeño que da cuenta de cada criterio de evaluación depende de la finalidad de la rúbrica. Aunque un menor número de niveles facilita la tarea del evaluador y aumenta la fiabilidad de la evaluación, igualmente disminuye la calidad de la información que genera la utilización de la rúbrica.

ca (Wolf y Stevens, 2007). Esta última repercusión es capital porque, por un lado, puede llegar a simplificar en exceso los objetivos de aprendizaje, a desvirtuar las actividades programadas y, por consiguiente, a desorientar a los estudiantes; por otro lado, empobrece la calidad de la realimentación que obtiene el estudiante. La sugerencia habitual es que los niveles de consecución se sitúen entre tres y cinco (Suskie, 2004).

Respecto a las descripciones cualitativas que dan noticia de cada uno de los niveles de desempeño, son comunes las advertencias sobre la claridad del lenguaje utilizado, la necesidad de remarcar los rasgos observables del objeto evaluado —particularmente si las rúbricas son utilizadas por los estudiantes en procesos de autoevaluación y evaluación entre iguales— y la necesidad de valorar atentamente la extensión de las descripciones. La dificultad aquí, una vez más, es alcanzar el justo término: escribir con un lenguaje sencillo pero que no hipoteque la precisión; destacar lo observable y lo medible sin que ello menoscabe los procesos de inferencia que exigen los resultados de aprendizaje complejos, redactar descripciones operativas que no sobrecarguen el proceso evaluador pero que tampoco apoquen o mutilen el fenómeno evaluado. Wolf y Stevens (2007) avanzaron una estrategia que puede ser de utilidad ante estos retos: definir los niveles de desempeño que despliegan cada criterio utilizando unos mismos atributos y un mismo esquema formal de redacción. Una variante es construir cada nivel de desempeño sobre el cumplimiento de lo establecido en los anteriores para añadir a continuación algún aspecto nuevo que suponga una mejora de los aprendizajes. Cualquier deficiencia en la claridad y especificidad de las descripciones cualitativas es una fuente probable de discrepancias en la interpretación de los resultados que se trasladan a las puntuaciones de los evaluadores; de ahí la importancia de que tras la fase de testeo se negocien los ajustes necesarios en este terreno (Moskal y Leydens, 2001). Por último, y ante la existencia de diferentes modelos de cada una de las principales actividades de aprendizaje en la literatura especializada, es indispensable proceder a una escrupulosa adaptación de la rúbrica al contexto educativo particular en que serán utilizadas. Simon y Forgette-Giroux (2001), Baughin, Brod y Page (2002) y Wolf y Stevens (2007) remarcaron que las rúbricas deben responder a la contextualización que caracteriza el proceso de enseñanza y aprendizaje. Solo de este modo se evita el riesgo de que las afirmaciones que constituyen las rúbricas sean demasiado generales o inadecuadas hasta el punto de impedir una interpretación coherente de los datos y/o abandonar el proceso de evaluación en manos de prejuicios no declarados.

Rúbricas para la valoración de proyectos en los estudios universitarios de las artes

Contexto educativo

El Observatorio sobre la Didáctica de las Artes (ODAS) de la Universidad de Barcelona implantó, en el curso académico 2012-2013, un sistema de rúbricas destinado a la evaluación de la exposición y defensa pública de proyectos de Aprendizaje-Servicio en las facultades de Bellas Artes

y de Geografía e Historia. La propuesta se introdujo en los Grados de Bellas Artes, Diseño, Conservación-Restauración e Historia del Arte. Concretamente, afectó a estudiantes de reciente ingreso que estaban matriculados en la asignatura de formación básica Conceptos del Arte Moderno impartida en los tres primeros grados citados, así como también a estudiantes que cursaban la asignatura —también de formación básica— Teoría del Arte y la obligatoria Historia de la estética en época moderna y contemporánea; ambas asignaturas del Grado de Historia del Arte programadas en tercer y sexto semestre respectivamente.

Las competencias transversales comunes de la Universidad de Barcelona a las que se orientaron principalmente los objetivos y actividades de aprendizaje fueron dos: la de formular, diseñar y gestionar proyectos, y la de comprender y expresarse oralmente. De manera subsidiaria, se establecieron la de colaborar con otros y contribuir a un proyecto común, y la de valorar el impacto social de las actuaciones. Estas competencias se desplegaron en los siguientes objetivos de aprendizaje:

- Reflexionar sobre la función social del arte y el perfil del profesional propio de la titulación, y aplicar en la práctica dichas reflexiones.
- Identificar y familiarizarse con la estructura básica de un proyecto en el ámbito de las artes.
- Identificar los aspectos básicos de un proyecto de Aprendizaje-Servicio.
- Identificar y organizar las fases principales de un proyecto. Establecer un calendario de actuación.
- Identificar y familiarizarse con los aspectos básicos de la exposición y defensa pública de un proyecto.
- Identificar los puntos críticos de una exposición oral y establecer estrategias adecuadas para su resolución.
- Planificar la exposición oral colectiva de un proyecto.
- Identificar y potenciar los puntos fuertes del trabajo en grupo.
- Identificar, prever y solucionar los problemas habituales en el trabajo en grupo.
- Equilibrar las acciones de cooperación, negociación y liderazgo en el trabajo en grupo.

En cuanto a las actividades de aprendizaje, los estudiantes realizaron un proyecto de Aprendizaje-Servicio en un contexto y para un grupo humano de su elección, con las limitaciones de que los proyectos debían estar orientados hacia un colectivo en riesgo de exclusión social, de que el arte debía desempeñar un papel crucial, y de que el enfoque se alineara con la titulación en la que estaban matriculados. Por tratarse de estudiantes que emprendían por primera vez esta actividad, los proyectos fueron diseñados y planificados pero no llevados a la práctica. Asimismo, fueron desarrollados en tres fases para propiciar que los estudiantes pudieran reflexionar

sobre el trabajo realizado, y revisar y mejorar sus resultados. La envergadura de los mismos promovió que fueran realizados en grupo cooperativos de cinco miembros.

Para el seguimiento de los grupos de trabajo por parte de los profesores y el aprendizaje de trabajar de manera colectiva se prestó especial atención a la gestión de su dinámica interna. Los grupos se reunieron fuera de las horas de clase y con la periodicidad que cada grupo entendiese más oportuna. En cada sesión, se distinguieron dos planos de trabajo. Un plano de discusión de las ideas, en el que participaron todos los miembros en condiciones de igualdad, y un plano de gestión del grupo. Para este segundo plano los estudiantes se distribuyeron de manera rotativa a lo largo de todo el semestre cinco roles —organizador/a, secretario/a, moderador/a, verificador/a y escéptico/a— que suponían la asunción de responsabilidades específicas en la organización del trabajo colectivo. El/la organizador/a fue el encargado de enviar las convocatorias de las reuniones, establecer el plan de trabajo con su respectiva distribución de tareas, y de que en las sesiones se trataran todos los puntos del orden del día. El/la secretario/a se responsabilizó de resumir los temas tratados y los acuerdos alcanzados y de enviarlo a todos los miembros. Para facilitar esta tarea, el equipo de ODAS estableció un formulario de acta que se puso a disposición de los estudiantes desde el comienzo del curso académico. El/la moderador/a gestionó que en cada sesión todos los miembros tomaran la palabra en iguales condiciones y veló para resolver las posibles fricciones de tipo personal. El verificador se cercioró de que cada miembro hubiese entendido las ideas tratadas y los acuerdos alcanzados, de que cada uno de los miembros recibiera la información necesaria, y de que hubiera concluido en el tiempo establecido las tareas que le habían sido asignadas. Por último, el/la escéptico encarnó la figura crítica del grupo y su cometido consistió en incidir en los puntos débiles del trabajo, prevenir posibles fuentes de error y procurar por la mejora general de los resultados.

Este seguimiento interno del funcionamiento de los grupos de trabajo se completó por un doble seguimiento externo. Por un lado, las tutorías con los profesores. Por el otro, el programa de tutoría entre iguales que los estudiantes tuvieron a su disposición a lo largo del todo el semestre. En dicho programa participaron un equipo de tres estudiantes tutores que estaban matriculados en estudios de máster o de doctorado y, como estudiantes tutorizados, los estudiantes matriculados en las titulaciones de grado. El programa fue totalmente voluntario y gratuito y su utilización no supuso ningún incremento en las calificaciones. Consistió en un encuentro semanal fuera de las horas de clase, en un espacio reservado para los estudiantes y con una duración no superior a 60 minutos. La utilización del programa fue totalmente flexible por parte de los estudiantes tutorizados. Pudieron hacer uso de él según su propio criterio, inscribiéndose al número de sesiones programadas y en las fases de realización del proyecto que creyesen más acorde con sus intereses. Una vez incorporados en el programa pudieron negociar con los estudiantes tutores otras fechas y horas para nuevas sesiones, e incluso pudieron elegir el tutor con el que trabajarían a lo largo del semestre.

El sistema de evaluación se alineó con los objetivos de aprendizaje y se ciñó a la realización de los proyectos. Para la valoración de los mismos se utilizó un sistema de rúbricas que los estudiantes conocían desde el comienzo del curso. El sistema de evaluación comprendió tres rúbricas analíticas constituidas por componentes generales y otros componentes específicos de la actividad realizada; cada rúbrica fue asociada a una de las tres fases de valoración de los proyectos. En cada rúbrica, los criterios de evaluación incluyeron siempre cuatro niveles de desempeño con sus definiciones cualitativas a las que se imprimió un carácter descriptivo con el fin de facilitar su uso por los estudiantes. Cada rúbrica mantuvo los criterios de evaluación de la fase anterior y añadió alguno nuevo; del mismo modo, y en lo referente a los niveles de desempeño, los niveles superiores recogieron las definiciones cualitativas de los inferiores y añadieron además un nuevo aspecto que especificaba una reflexión más amplia y profunda sobre el proyecto.

En cuanto a la estructura de las rúbricas, los criterios de evaluación se organizaron en dos grandes bloques. En el primero, se incluyó todo aquello directamente relacionado con el contenido del proyecto: título, finalidad y objetivos (presente en las tres rúbricas), documentación (en las tres rúbricas), antecedentes (rúbricas segunda y tercera), y descripción detallada y discusión (rúbrica tercera). El segundo bloque recogió aspectos principales de la competencia comunicativa oral tales como adecuación al auditorio, adecuación al tiempo disponible, elocución, gramática y contacto visual; todos estos criterios formaron parte de las tres rúbricas. El nivel de desempeño más elevado de los criterios de evaluación de contenido comprendía una descripción del apartado correspondiente, una justificación del orden seguido en la exposición y una reflexión sobre la relevancia de lo expuesto para el desarrollo del proyecto. Por su parte, el nivel más alto de los criterios relativos a las competencia comunicativa venía determinado por la interacción entre el ponente y el auditorio. Por ejemplo, el criterio de elocución señalaba una forma de hablar entusiasta y vívida; el criterio de gramática refería un lenguaje efectivo para enfatizar el significado y atraer la atención del público, el criterio de contacto visual valoraba la capacidad de particularizar en algunos de los asistentes; y el criterio de adecuación al auditorio incidía en un lenguaje y un comportamiento respetuoso hacia el auditorio y los compañeros del propio grupo de trabajo. El único criterio de evaluación de carácter estrictamente cuantitativo fue, por razones obvias, el de adecuación al tiempo disponible, y fue también el único que se añadió al bloque de los criterios de la exposición oral en la tercera fase.

Las rúbricas fueron utilizadas por estudiantes, estudiantes tutores y profesores como recursos didácticos en el desarrollo y seguimiento de los proyectos, y también como instrumentos de valoración de los resultados obtenidos a lo largo del semestre. Para ello, se implementó una aplicación en línea que fue usada por todos los participantes en idénticas condiciones. La aplicación formaba parte de la plataforma electrónica vinculada a las asignaturas, y fue diseñada de manera que las rúbricas se activaron automáticamente en coincidencia con la apertura de cada una de las tres fases de evaluación. Luego, cada evaluador introducía sus valoraciones en la aplicación: se

leccionaba el nivel de desempeño alcanzado en cada criterio de evaluación y, a continuación, le asignaba una puntuación numérica entre un rango de puntuaciones que por defecto tenía asignados cada uno de los niveles. Todos los participantes puntuaron individualmente con estas rúbricas las presentaciones orales de los proyectos realizados por los diferentes grupos de trabajo; así, junto con la valoración de los profesores se puso en práctica un proceso de auto-evaluación y evaluación entre iguales. Sin embargo, a la espera de un estudio sobre su uso por los estudiantes, solo las puntuaciones de los profesores computaron en la determinación de las calificaciones de los estudiantes.

El proceso de diseño de las rúbricas

En el diseño de las rúbricas trabajaron cinco investigadores del equipo de ODAS: dos profesores y tres estudiantes que participaban en el programa de tutorías como tutores. El proceso comenzó por una revisión en profundidad de la literatura existente con el fin de tomar las decisiones oportunas acerca de las ventajas y desventajas del tipo de rúbricas, del número y tipo de criterios de evaluación, del número y tipo de niveles de consecución, y de los problemas habituales que afectan al diseño. El examen de la literatura concluyó con una revisión de los modelos existentes en el ámbito universitario para la valoración de proyectos de variado signo y de exposiciones orales.

En una segunda fase, los dos profesores plantearon y discutieron los resultados de aprendizaje esperados al final del semestre. Ello se hizo tras considerar tres cuestiones: las propiedades de un proyecto de Aprendizaje-Servicio y de una exposición oral en el campo profesional de las artes, la experiencia previa de los estudiantes en este tipo de actividades, y su pericia en otras competencias requeridas —principalmente la de búsqueda, selección e integración de información y la de trabajo en grupo. A partir de aquí, se establecieron las características principales de los proyectos y de las exposiciones orales que deberían orientar las actividades que los estudiantes realizarían en el curso de las asignaturas participantes. Por último, se acordó los aspectos centrales del sistema de evaluación continuada y un primer esquema de la distribución de las puntuaciones. Fue en este momento cuando se propuso una evaluación de los proyectos en tres fases y el peso específico de cada fase en la calificación final. La razón principal fue lograr un equilibrio entre dos aspectos. Por un lado, llevar a cabo un seguimiento de los proyectos que permitiera a los profesores tener un conocimiento profundo de su evolución. Este punto resultaba esencial a fin de proceder a un comentario de los resultados que fuera útil para que los estudiantes pudiesen entender en qué estadio se encontraban respecto de los objetivos previstos y a fin de acordar con los estudiantes líneas futuras de mejora. Por otro lado, que el proceso de evaluación fuese lo suficientemente espaciado para que los estudiantes tuvieran tiempo de reflexionar sobre los resultados obtenidos y planificar, organizar y desarrollar en consecuencia la fase siguiente. Recuérdese que cada fase debería suponer una revisión y mejora del trabajo realizado así como la consideración de dimensiones nuevas del proyecto.

La tercera fase fue cuando se procedió al diseño propiamente dicho de las rúbricas. La primera decisión fue la de elegir un sistema de rúbricas analíticas en lugar de holísticas. La literatura especializada reconoce que la rúbrica holística es más fácil de elaborar y de utilizar como instrumento de evaluación; además, la evaluación basada en estas rúbricas suele venir acompañada por unos coeficientes de fiabilidad superiores. Sin embargo, se optó por una rúbrica analítica por su mayor capacidad informativa, lo que resultaba determinante para favorecer una mayor comprensión de los aprendizajes por parte de los estudiantes. Asimismo, al parecer de ambos profesores, resultaba la alternativa obligada si se pretendía la participación de los estudiantes en procesos de auto-evaluación y evaluación entre iguales. Se desecharon también los modelos consultados de rúbricas de proyectos y de exposiciones orales. En el caso de las exposiciones orales, los modelos acostumbraban a ser demasiado genéricos, mientras que los de proyectos pecaban justamente de lo contrario. Otro problema hallado fue que las definiciones cualitativas de las rúbricas analíticas apuntaban en ocasiones hacia aspectos de bajo nivel cognitivo para los estudios universitarios. Un tercer inconveniente fue que el objeto de evaluación que los dos profesores se habían propuesto entrañaba tal complejidad que una valoración mediante el uso de dos rúbricas —una para los proyectos y otra para la exposición oral— resultaba impracticable; máxime en el caso de los estudiantes. En efecto, el hecho de que debería evaluarse la calidad de los proyectos y de las exposiciones orales en actos que duraban 20 minutos en la primera fase, 30 en la segunda y 45 en la tercera, obligaba a limitar el número de criterios de evaluación y a seleccionarlos cuidadosamente. Con el agravante de que la evaluación de ambas actividades debía realizarse de manera simultánea y que, en lo referente a la exposición oral, los aspectos de la misma discurren indisolublemente unidos. Todo ello hizo imprescindible inclinarse por un sistema de rúbricas analíticas y mixtas, en el que se conjugaban componentes genéricos y otros específicos de las actividades.

En virtud de todo lo dicho, se convino restringir a diez el número de criterios de evaluación. El número de niveles de consecución por cada criterio quedó fijado en cuatro, de acuerdo con la horquilla recomendada en la literatura. Un número menor —tres, por ejemplo— fija adecuadamente los niveles inferior y superior, pero no discrimina los niveles intermedios de una manera adecuada; y son en estos niveles donde se sitúa la mayor parte de los estudiantes, sobre todo en las primeras fases. Por el contrario, un número superior despliega la parte central de la escala pero presenta dos inconvenientes: primero, la rúbrica es mucho más difícil de diseñar por la especificidad y precisión que exige la necesaria diferenciación de estos niveles intermedios; segundo, la observación inequívoca de los matices que distinguen los niveles en el momento de la evaluación es una actividad extremadamente compleja, incluso en el caso de los evaluadores expertos, en una actividad efímera como lo es la exposición oral. Por consiguiente, un número mayor de niveles intermedios comprometería la fiabilidad de la evaluación y el grado de acuerdo entre los evaluadores. Además, la proliferación de estos niveles intermedios disminuiría la utilidad de las rúbricas al hacer más complicado para los estudiantes el distinguir con claridad entre niveles más próximos.

Hay que advertir aquí que los estudiantes tenían una experiencia muy limitada en las exposiciones orales, y no tenían ninguna en la realización de proyectos de Aprendizaje-Servicio y en la utilización de rúbricas. En lo concerniente al peso de cada criterio de evaluación en la puntuación final de cada fase, se acordó que los criterios relativos al proyecto que en cada fase aparecieran por primera vez recibirían el 20% de la puntuación total de la rúbrica y que ese porcentaje disminuiría hasta el 10% en las fases sucesivas. Los criterios de la exposición oral se mantuvieron estables en las tres fases evaluadoras. En la primera, se les asignó un 15% del valor total de la puntuación y, en las sucesivas, un 10%.

Con posterioridad, los profesores trabajaron por separado en la selección de los criterios de evaluación y en la redacción de las descripciones cualitativas de los niveles de desempeño. Para, a continuación, discutir cada propuesta, negociar y acordar el sistema de rúbricas definitivo. Fue en este momento en que comenzó la colaboración de los estudiantes tutores. Las rúbricas propuestas fueron enviadas a cada uno de ellos para que realizaran un análisis crítico e incorporaran las modificaciones que creyeran oportunas. Transcurrido el tiempo suficiente, la fase de diseño se cerró con dos reuniones de todos los miembros participantes en las que se explicaron las razones que motivaron cada propuesta de modificación y se examinaron sus pros y contras. Aunque se trataron todos los aspectos de las rúbricas y en todos ellos los estudiantes tutores contribuyeron con sus comentarios, su aportación más importante se produjo en la redacción de las descripciones cualitativas. Más concretamente, desempeñaron un papel inestimable en la depuración del lenguaje con el fin de aproximarlos a los estudiantes y realizaron observaciones de máxima importancia que ayudaron a precisar más lo señalado en cada nivel y equilibrar el grado de diferencia que mediaba entre ellos.

Por último, el sistema de rúbricas entró en su fase de pruebas, que se prolongó durante todo un semestre, y en la que las rúbricas fueron utilizadas en la práctica evaluadora por los estudiantes y los miembros de ODAS que participaron en su diseño. Tras este ensayo, el sistema de rúbricas se mejoró en dos facetas. En primer lugar, se mejoraron los descriptores cualitativos para señalar con la máxima claridad y concreción posibles qué indicadores eran determinantes en cada nivel. De este modo se intentó que todos los evaluadores se concentraran en los mismos aspectos del fenómeno observado y así se alcanzara un mayor acuerdo. En segundo, se prescindió del criterio destinado a valorar los recursos audiovisuales utilizados en las exposiciones orales. Esta supresión estaba justificada por dos razones. La primera era que una exposición oral que se situase en el máximo nivel no requería necesariamente la utilización de dichos recursos. Pero, justamente el hecho de no utilizarlos impedía de por sí que una exposición brillante alcanzase la máxima puntuación debido a la redacción original del criterio de evaluación. La segunda razón era que la valoración del correcto funcionamiento de los medios audiovisuales fluctuó en ocasiones entre el auditorio —estudiantes o miembros de ODAS— por motivos espurios como fueron, por ejemplo, la calidad de los sistemas de video y audio que los estudiantes tenían a su disposición en el aula o la

imposibilidad de distribuir los asientos de manera que todos los asistentes pudieran observar en las mismas condiciones la calidad de los recursos audiovisuales utilizados. Dado que la solución del problema no estaba en manos de los profesores, se estimó conveniente eliminar el criterio de las rúbricas.

El sistema de rúbricas

La Tabla 1 presenta el sistema de rúbricas completo. La tabla incluye los criterios y los niveles de desempeño con sus descripciones cualitativas, los porcentajes de la nota final asignados a cada criterio en cada una de las fases y, como nota a pie de tabla, los umbrales de puntuación que se asociaron con cada uno de los niveles de desempeño. En este caso, se utilizó una escala del cero al diez por ser la habitual en el sistema universitario español. Se podrá observar que, en general, los criterios de evaluación ven decrecido su porcentaje de la nota final a medida que se van repitiendo en fases sucesivas. Sin embargo, los criterios de evaluación «Título, finalidad y objetivos», «Documentación» y «Adecuación al tiempo disponible» representaron tres excepciones. Los dos primeros fueron los criterios incorporados desde la primera fase, constituyeron el fundamento de los proyectos sobre los que se apoyó su desarrollo ulterior y supusieron los aspectos que más dedicación exigieron a los estudiantes. Por este motivo, mantuvieron el porcentaje inicial del 20% también en la segunda fase. Con el criterio «Adecuación al tiempo disponible» ocurrió justamente lo contrario. A diferencia de los demás criterios referidos a la exposición oral, fue el único que no se valoró desde la primera fase y, cuando se incluyó en las rúbricas (tercera fase), se le atribuyó un porcentaje del 10% y no del 15% como ocurrió con los restantes.

Tabla 1.

Sistema de rúbricas para la exposición oral de proyectos artísticos de Aprendizaje-Servicio

Criterios y niveles de desempeño ^a	% por fase		
	1 ^a	2 ^a	3 ^a
Título, finalidad y objetivos	20%	20%	10%
A. No hay título. O no hay finalidad o el auditorio no puede hacerse una idea del contexto en el que se desarrolla el proyecto. O no hay objetivos o la mayoría de estos no se relacionan con la finalidad o se confunden con ella.			
B. Hay título. El auditorio puede hacerse una idea del contexto en el que se desarrolla el proyecto y del colectivo al que se dirige. Los objetivos concretan la finalidad en líneas de trabajo.			
C. Hay título. El auditorio puede hacerse una idea del contexto en el que se desarrolla el proyecto y del colectivo al que se dirige. Los objetivos concretan la finalidad en líneas de trabajo y son realizables por el equipo.			
D. El auditorio puede hacerse una idea general del objeto del proyecto a			

través del título. El auditorio puede hacerse una idea del contexto en el que se desarrolla el proyecto y del colectivo al que se dirige. Los objetivos concretan la finalidad en líneas de trabajo y son realizables por el equipo. La participación del grupo es indispensable para lograr los objetivos.

Documentación 20% 20% 10%

A. No se presenta un listado de documentos accesible al auditorio. O la mayor parte de los documentos no son adecuados a los estudios universitarios.

B. Se presenta un listado de documentos accesible al auditorio. La mayor parte de los documentos no son adecuados a los estudios universitarios. Se resume la información extraída de la documentación.

C. Se presenta un listado de documentos accesible al auditorio. La mayor parte de los documentos son adecuados a los estudios universitarios. Se resume la información extraída de cada documento.

D. Se presenta un listado de documentos accesible al auditorio. La mayor parte de los documentos son adecuados a los estudios universitarios. Se resume la información extraída de cada documento y se justifica su utilidad para el proyecto. Se utilizan estándares de referencia en la documentación utilizada.

Antecedentes 20% 10%

A. No se presentan antecedentes. O no se justifica la relación con el proyecto de la mayor parte de los antecedentes.

B. Se presentan antecedentes. No se justifica la relación con el proyecto de la mayor parte de los antecedentes. Se explica la manera de organizar la exposición de los antecedentes.

C. Se presentan antecedentes. Se justifica la relación con el proyecto de la mayor parte de los antecedentes. Se explica la manera de organizar la exposición de los antecedentes. La mayor parte de los antecedentes son diversos según el ámbito geográfico (local, nacional, internacional) o según el período histórico o según las diferentes actividades que los componen.

D. Se presentan antecedentes. Se justifica la relación con el proyecto de la mayor parte de los antecedentes. Se explica la manera de organizar la exposición de los antecedentes. La mayor parte de los antecedentes son diversos según el ámbito geográfico (local, nacional, internacional) o según el período histórico o según las diferentes actividades que los componen. Se comentan las similitudes y diferencias de los antecedentes con el proyecto.

Descripción y discusión 20%

A. La mayor parte de las actividades que componen el proyecto no están descritas con detalle. O no están previstos el tipo de materiales, de equipamientos y de actuaciones necesarias.

B. La mayor parte de las actividades que componen el proyecto están descritas con detalle. Están previstos el tipo de materiales, de equipamientos y de actuaciones necesarias.

C. La mayor parte de las actividades que componen el proyecto están descritas con detalle. Están previstos el tipo de materiales, de equipamientos y de actuaciones necesarias. Se explica la evolución del proyecto en sus diferentes fases.

D. La mayor parte de las actividades que componen el proyecto están descritas con detalle. Están previstos el tipo de materiales, de equipamientos y de actuaciones necesarias. Se explica la evolución del proyecto en sus diferentes fases. Se han anticipado los posibles problemas y las soluciones adecuadas. Se explica cómo los participantes en el proyecto intervienen en el seguimiento y revisión del proyecto.

Adecuación al acto público

15% 10% 10%

A. Se utilizan términos o expresiones inapropiadas al tipo de acto público.

B. Se usan términos o expresiones apropiadas al tipo de acto público.

C. Se usan términos o expresiones apropiadas al tipo de acto público. Se mantiene un comportamiento respetuoso con el auditorio.

D. Se usan términos o expresiones apropiadas al tipo de acto público. Se mantiene un comportamiento respetuoso con el auditorio. No hay interrupciones u otras faltas de respeto entre los miembros del grupo de trabajo.

Elocución

15% 10% 10%

A. El volumen es tan bajo que el auditorio no puede oír lo que se dice. O el ritmo es tan rápido que no se puede seguir. O la pronunciación es tan confusa que no se entienden las palabras o frases.

B. El volumen es demasiado bajo o demasiado alto pero se puede oír. O el ritmo es demasiado lento o demasiado rápido, pero se puede seguir. O la pronunciación no es clara. O hay demasiados murmullos, balbuceos y muletillas.

C. El volumen, ritmo y las pausas del discurso son apropiados. La pronunciación es clara. El ponente exhibe pocos murmullos, balbuceos y muletillas.

D. El ponente utiliza el volumen, ritmo y las pausas para subrayar el significado del mensaje y suscitar el interés en el auditorio. El ponente habla de forma entusiasta y vívida. No hay prácticamente murmullos, balbuceos y muletillas.

Gramática

15% 10% 10%

A. Hay tantos errores que no se puede entender la mayor parte de la exposición.

B. Hay muchos errores gramaticales pero no dificultan la comprensión del

discurso.

C. Hay pocos errores gramaticales.

D. Hay pocos errores gramaticales. Y el hablante usa un lenguaje altamente efectivo para enfatizar el significado y atraer la atención del auditorio.

Contacto visual

15% 10% 10%

A. No hay contacto visual con el auditorio. El ponente lee.

B. Contacto visual ocasional. El ponente se apoya mucho en sus notas.

C. Contacto visual mantenido. El ponente se apoya poco en sus notas.

D. Contacto visual permanente. Capacidad de particularizar en algunos de los asistentes. Raro recurso a las notas.

Adecuación al tiempo disponible

10%

A. Defecto o exceso de 16 o más minutos.

B. Defecto o exceso de 11 a 15 minutos.

C. Defecto o exceso de 6 a 10 minutos.

D. Defecto o exceso de 5 minutos como máximo.

^aUmbral de las puntuaciones de los niveles de desempeño: A: 0-2, B: 3-4, C: 5-7, D: 8-10.

Referencias

Andrade, H. (2000). Using rubrics to promote thinking and learning. *Educational Leadership*, 57(5), 13-18.

Andrade, H., y Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research & Evaluation*, 10(5), 1-11.

Arter, J., y Chappuis, J. (2007). *Creating and recognizing quality rubrics*. Upper Saddle River, NJ: Pearson/Merrill Prentice Hall.

Arter, J., y McTighe, J. (2001). *Scoring rubrics in the classroom: Using performance criteria for assessing and improving student performance*. Thousand Oaks, CA: Corwin.

Bahous, R. (2008). The Self-Assessed Portfolio: A Case Study. *Assessment and Evaluation in Higher Education*, 33(4), 381-393.

Baughin, J. A., Brod, E. F., y Page, D. L. (2002). Primary Trait Analysis: A Tool for Classroom-Based Assessment. *College Teaching*, 50(2), 75-80.

Beauchamp, L., McConaghy, G., Parsons, J., Sanford, K., y Ford, D. (2000). *Teaching From the Outside In*. Edmonton: Duval.

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32(3), 347-364.

Brookhart, S. M. (1999). *The Art and Science of Classroom Assessment: The Missing Pedagogy*. ASHE-ERIC Higher Education Report (Vol. 27, No. 1). Washington DC: The George Washington University, Graduate School of Education and Human Development.

Cox, J. W. (1995). Assessment from the inside out. *Issues and Inquiry in College Learning and Teaching*, 17/18(2), 19-29.

De La Cruz Flores, G., Díaz-Barriga Arceo, F., y Abreu Hernández, F. (2010). La labor tutorial en los estudios de posgrado. Rúbricas para guiar su desempeño y evaluación. *Perfiles Educativos*, 32(130), 83-102.

Díaz Barriga, F. (2004). Las rúbricas: su potencial como estrategia para una enseñanza situada y una evaluación auténtica del aprendizaje. *Perspectiva Educativa*, 43, 51-62.

Entwistle, N. (1997). Contrasting perspectives on learning. In F. Marton, D. Hounsell, y N. Entwistle (Eds.), *The experience of learning. Implications for teaching and studying in higher education* (2nd ed., pp. 3–22). Edinburgh: Scottish Academic Press.

Marton, F., y Säljö, R. (1997). Approaches to learning. In F. Marton, D. Hounsell, & N. Entwistle (Eds.), *The experience of learning. Implications for teaching and studying in higher education* (2nd ed., pp. 39–58). Edinburgh: Scottish Academic Press.

Maxwell, S. (2010). *Using Rubrics to Support Graded Assessment in a Competency Based Environment*. Occasional Paper. Adelaide: National Centre for Vocational Education Research.

Mertler, C. A. (2001). Designing scoring rubrics for your classroom. In L. M. Rudner, y W. D. Schafer. *Practical Assessment, Research & Evaluation*, 2000-2001 (pp. 189-197). Consultado en <http://www.eric.ed.gov/PDFS/ED458254.pdf>

Moni, R. W., Beswick, E., y Moni, K. B. (2005). Using student feedback to construct an assessment rubric for a concept map in physiology. *Advances in Physiology Education*, 29, 197-203.

Moskal, B. M. (2000). Scoring Rubrics: What, When, and How? In L. M. Rudner, y W. D. Schafer. *Practical Assessment, Research & Evaluation*, 2000-2001 (pp. 22-29). Consultado en <http://www.eric.ed.gov/PDFS/ED458254.pdf>

- Moskal, B., y Leydens, J. (2001). Scoring Rubric Development: Validity and Reliability. In L. M. Rudner, y W. D. Schafer. *Practical Assessment, Research & Evaluation, 2000-2001* (pp. 71-81). Consultado en <http://www.eric.ed.gov/PDFS/ED458254.pdf>
- Montgomery, K. (2002). Authentic Tasks and Rubrics: Going beyond Traditional Assessments in College Teaching. *College Teaching, 50*(1), 34-39.
- Parkes, K. A., y Kajder, S. (2010). Eliciting and Assessing Reflective Practice: A Case Study in Web 2.0 Technologies. *International Journal of Teaching and Learning in Higher Education, 22*(2), 218-228.
- Pintrich, P. R. (2002). The role of metacognitive knowledge in learning, teaching and assessing, *Theory into Practice, 41*(4), 219-225.
- Reddy, Y. M., y Andrade, H. (2010). A Review of Rubric Use in Higher Education. *Assessment & Evaluation in Higher Education, 35*(4), 435-448.
- Reynolds, J. et al. (2009). BioTAP: A Systematic Approach to Teaching Scientific Writing and Evaluating Undergraduate Theses. *BioScience, 59*(10), 896-903.
- Rhodes, T. L. (2010). Since We Seem to Agree, Why Are the Outcomes so Difficult to Achieve? *New Directions for Teaching and Learning, 121*, 13-21.
- Simon, M. y Forgette-Giroux, R. (2001). A rubric for scoring postsecondary academic skills. In L. M. Rudner, y W. D. Schafer. *Practical Assessment, Research & Evaluation, 2000-2001* (pp. 139-145). Consultado en <http://www.eric.ed.gov/PDFS/ED458254.pdf>
- Stiggins, R. J. (2001). *Student-involved classroom assessment* (3rd ed.). Upper Saddle River, NJ: Prentice-Hall.
- Suskie, L. (2004). *Assessing student learning: A common sense guide*. Bolton, MA: Anker.
- Tierney, R., y Simon, M. (2004). What's still wrong with rubrics: focusing on the consistency of performance criteria across scale levels'. *Practical Assessment, Research & Evaluation, 9*(2), Consultado en <http://PAREonline.net/getvn.asp?v=9&n=2>.
- Torrie, H. (2007). *A Web-Based Tool for Oral Practice and Assessment of Grammatical Structures*. Online submission (ED508831). Utah: Department of Linguistics and English Language. Brigham Young University. Consultado en <http://www.eric.ed.gov/PDFS/ED508831.pdf>

Turley, E. D., y Gallagher, C. W. (2008). On the "Uses" of Rubrics: Reframing the Great Rubric Debate. *English Journal*, 97(4), 87-92.

Walvoord, B. E., Bardes, B., y Denton, J. (1998). Closing the feedback loop in classroom-based assessment. *Assessment Update*, 10(5), 1-4.

Welch, M., y James, R. C. (2007). An Investigation on the Impact of a Guided Reflection Technique in Service-Learning Courses to Prepare Special Educators. *Teacher Education and Special Education*, 30 (4), 276-285.

Weld, J. (2002). Making the Grade: Student Design and Defense of a Seminar Course Assessment. *Journal of College Science Teaching*, 31(6), 394-398.

Wiggins, G. (1998). *Educative Assessment: Designing Assessment to Inform and Improve Student Performance*. San Francisco: Jossey-Bass.

Wolf, K., y Goodwin, L. (2007). Evaluating and Enhancing Outcomes Assessment Quality in Higher Education Programs. *Metropolitan Universities*, 18(2), 42-56.

Wolf, K., y Stevens, E. (2007). The Role of Rubrics in Advancing and Assessing Student Learning. *The Journal of Effective Teaching*, 7(1), 3-14.