



IMPACTOS DE LA INNOVACIÓN EN LA DOCENCIA Y EL APRENDIZAJE

La incorporación de estudiantes de posgrado en equipos de evaluación

Análisis de la aportación de las rúbricas al desarrollo del juicio profesional

Gregori Giralt, Eva

Universitat de Barcelona

Departamento de Historia del Arte / Facultad de Bellas Artes

Adolf Florensa, s/n. 08028 Barcelona (España)

gregori@ub.edu

Menéndez Varela, José Luis

Universitat de Barcelona

Departamento de Historia del Arte / Facultad de Bellas Artes

Adolf Florensa, s/n. 08028 Barcelona (España)

menendez@ub.edu

1. **RESUMEN:** Se analiza el grado de acuerdo de tres estudiantes de posgrado y dos profesores en una evaluación basada en rúbricas de proyectos de estudiantes. Los estudiantes de posgrado habían participado en un programa de formación, en el rediseño de las rúbricas y en las sesiones de moderación con los profesores. Los resultados señalan que la mejor contribución de las rúbricas es que concretan los procesos de negociación, y que su utilización debe venir acompañada de otros recursos de aprendizaje.
2. **ABSTRACT:** This paper analyze the degree of agreement among three graduate students and two teachers in a rubrics-based performance assessment of students' projects. The graduate students had participated in a training program, in the redesign of the rubrics and in moderation sessions with teachers. The results show that the best contribution of rubrics is to materialize the negotiation processes, and that their use should be accompanied by other learning resources.
3. **PALABRAS CLAVE:** rúbrica, evaluación, juicio profesional, comunidades de práctica.

KEYWORDS: rubric, assessment, professional judgement, communities of practice.



IMPACTOS DE LA INNOVACIÓN EN LA DOCENCIA Y EL APRENDIZAJE

4. DESARROLLO:

Introducción

El experto es aquél que está familiarizado con los paradigmas, metodologías y objetos de estudio propios de su disciplina, conoce la organización social del trabajo y participa de las representaciones y prácticas cotidianas compartidas con sus colegas. Por ello, formar al experto requiere introducirlo en la práctica del juicio profesional integrándole en contextos reales en los que pueda compartir su responsabilidad y dialogar sobre los criterios y procedimientos de evaluación con otros evaluadores más experimentados (Beck, Skinner, y Schwabrow, 2013). Esto es, la formación del experto requiere su familiarización con el juicio en contextos profesionales reales o en entornos didácticos que los simulen. Pero también que la acción educativa se centre en la metacognición y en el desarrollo de herramientas de aprendizaje autónomo.

Vista así, la idea de disciplina deja de ser un constructo epistemológico para devenir auténtica cultura profesional. Y ello tiene vínculos directos con lo que Polanyi (1958; 1962) denominaba conocimiento tácito, es decir, esa dimensión del conocimiento que puede ser progresivamente formalizada hasta transformarse en conocimiento explícito (Duguid, 2005; Klein, 2008; Tsoukas, 2002). La consideración de una modalidad de conocimiento poco articulada trae consigo y determina una idea del juicio del experto como un acto basado en procesos metacognitivos y en prácticas de interpretación socialmente situadas. Algunos autores han establecido una conexión entre el concepto de evaluación sostenible y el de universidad entendida como comunidad de práctica (Carroll, 2005; O'Malley, 2010; Ward y Selvester, 2012). Si con la evaluación sostenible se pretende introducir al estudiante en el juicio profesional para que participe de manera más relevante en la disciplina, con la noción de comunidad de práctica se persigue crear grupos de trabajo informales y heterogéneos que dialogan y opinan libremente sobre las prácticas profesionales. Varias investigaciones han probado los beneficios de equipos de profesores que funcionan como comunidades de práctica (Bird y Yucel, 2013; Brooks, 2012; Price, 2005). Sin embargo, está todavía por ver qué recursos y cómo deben ser utilizados para incorporar también a los estudiantes en dichos equipos de evaluación.

Las rúbricas han sido planteadas como un recurso útil para esta incorporación del estudiante en equipos de evaluadores y hacerle comprensible el juicio del experto. Sin embargo, la claridad y transparencia generadas por las rúbricas no son suficientes para asegurar una aplicación consistente de los criterios porque estos están sujetos a procesos sociales de construcción de sentido (Knight, 2006). Las rúbricas no pueden recoger todo aquello implícito en el juicio que también interviene en la evaluación (Moskal y Leydens, 2001), pero sí que



IMPACTOS DE LA INNOVACIÓN EN LA DOCENCIA Y EL APRENDIZAJE

pueden ayudar al estudiante a afrontar dificultades inherentes a la formación de un juicio experto: por ejemplo, prevenir la dispersión de los debates, identificar los consensos de una cultura profesional e incluso mejorar la comprensión sobre cómo operan los juicios.

Pero para que ello ocurra es imprescindible discutir las rúbricas con los estudiantes (Oakleaf, 2008; Sadler, 2009) y que estas discusiones sean periódicas (Beck, Skinner, y Schwabrow, 2013). Varias investigaciones tratan la participación de los estudiantes en el diseño de las rúbricas (Hughes y Cappa, 2007; Wolf y Stevens, 2007), los programas de formación dirigidos a los estudiantes (Crotwell Timmerman, Strickland, Johnson, y Payne, 2011; Smith, Worsfold, Davies, Fisher, y McPhail, 2013; Welch y James, 2007), o la participación de los discentes en las sesiones de moderación (Bird y Yucel, 2013; Crotwell Timmerman et al., 2011).

El objeto de esta investigación es determinar la utilidad de las rúbricas en la incorporación de estudiantes de posgrado en equipos de evaluación para desarrollar su juicio profesional, y valorar hasta qué punto el uso de rúbricas debe venir acompañado de otros recursos de aprendizaje. Para ello, se diseñaron dos intervenciones. La primera intervención permitió examinar cómo los estudiantes utilizaron las rúbricas tras recibir un programa de formación. La segunda replicó el análisis pero después de que los estudiantes participaran en un proceso de rediseño de las rúbricas y asistieran a las sesiones de moderación junto con los profesores. Con el fin de comprobar si los estudiantes evaluadores habían desarrollado estrategias de evaluación consistentes, sus puntuaciones se compararon con el juicio experto de los profesores mediante un análisis de la varianza y del grado de acuerdo entre evaluadores.

Descripción del trabajo

Participantes

La muestra estuvo compuesta por un total de 184 estudiantes de reciente ingreso en las titulaciones de Bellas Artes, Diseño y Conservación-Restauración de la Universidad de Barcelona. En un 76,28% eran mujeres de 18 años de edad (según el valor de la moda) que mayoritariamente (en un 72,16%) se dedicaba solo a los estudios. Todos estaban matriculados en una asignatura de formación básica programada en el primer y segundo semestre de la titulación y todos llevaron a cabo proyectos de aprendizaje-servicio en grupos de unos cinco miembros cada uno.

En el primer semestre, los dos profesores responsables de la asignatura y dos estudiantes de posgrado evaluaron la defensa oral de los proyectos elaborados por 20 grupos de estudiantes en tres momentos diferentes del curso. Los estudiantes de posgrado ejercían como tutores en un programa de tutorías entre iguales diseñado para sus compañeros de reciente ingreso.



IMPACTOS DE LA INNOVACIÓN EN LA DOCENCIA Y EL APRENDIZAJE

Estos tutores eran una mujer de 25 años y un hombre de 26 que estaban cursando estudios de máster y doctorado respectivamente. La segunda intervención transcurrió durante el segundo semestre, y en ella los dos profesores y los dos estudiantes anteriores junto con una tercera evaluaron a 18 grupos de trabajo también en tres ocasiones. A diferencia de sus dos compañeros, esta tercera estudiante también de máster no había participado en el programa de tutoría entre iguales. De modo que su incorporación tuvo como finalidad examinar la posible influencia de dicho programa sobre la competencia evaluadora.

Diseño

De acuerdo con la clasificación de Montero y León (2007), se realizó un cuasi experimento pre-post para estimar el desarrollo de la competencia evaluadora en los estudiantes de posgrado. Los profesores diseñaron un sistema de tres rúbricas antes de empezar el curso que estaban pensadas para ayudar a los estudiantes a comprender los objetivos de aprendizaje y proporcionarles información útil para el desarrollo de sus proyectos; centrar las sesiones de trabajo entre estudiantes y profesores; y facilitar una evaluación continuada en la que todos los participantes (profesores, estudiantes evaluadores y estudiantes de reciente ingreso) puntuaran individualmente las presentaciones orales de sus proyectos. Las rúbricas mantuvieron una estructura idéntica de cuatro niveles de desempeño. Cada nivel incluía y mejoraba las descripciones cualitativas del nivel precedente e incorporaba algún otro aspecto que añadía complejidad al desarrollo del proyecto. Esta progresividad de los criterios según se ascendía en los niveles de consecución se trasladó también a las tres fases en que se distribuyó el desarrollo de los proyectos. A medida que avanzaban las fases, las rúbricas mantenían los criterios o dimensiones de la fase anterior -y así se hacía posible la solución de las insuficiencias detectadas- y se añadían criterios que recogían otros aspectos fundamentales del desarrollo de los proyectos.

Si en la primera intervención, se programaron unas sesiones de formación sobre la utilización de las rúbricas, en la segunda, los tres estudiantes de posgrado participaron en el rediseño de las rúbricas y en las sesiones de moderación que siguieron a cada una de las fases de evaluación de los proyectos. El programa de formación constó de un total de 6 horas durante las cuales se trató con los estudiantes qué es un proyecto de aprendizaje-servicio, qué estructura fundamental debe organizar un proyecto de este tipo, qué problemas acostumbran a presentar los estudiantes en su diseño y planificación, cuáles son los aspectos principales de una exposición oral y cómo todo ello venía recogido en las rúbricas. La segunda intervención estuvo precedida por la fase de rediseño de las rúbricas. En dicha fase, los dos profesores participantes y los dos estudiantes tutores analizaron por separado los puntos fuertes y débiles del sistema de rúbricas inicial a partir de su experiencia en el primer semestre; elaboraron



IMPACTOS DE LA INNOVACIÓN EN LA DOCENCIA Y EL APRENDIZAJE

propuestas de cambio y las discutieron con la tercera estudiante que todavía no había tenido ocasión de utilizar las rúbricas.

Tanto los estudiantes de posgrado como los profesores asistieron a las presentaciones con la plantilla de evaluación y, con el fin de evitar discrepancias atribuibles a las condiciones ambientales, se sentaron juntos y en el mismo lugar durante todas las sesiones de exposición oral; tomaron nota sobre los aspectos que consideraban relevantes de cada criterio de evaluación recogido en las rúbricas y, en la segunda intervención, describieron cómo habían visto los proyectos en las reuniones celebradas tras cada sesión de exposición. A diferencia de otras experiencias, en esta segunda intervención, los evaluadores calificaron individualmente cada uno de los criterios después de las sesiones de moderación.

Análisis

Teniendo en cuenta los objetivos del estudio y la naturaleza de los datos recopilados, en primer lugar se realizó un análisis de la varianza (ANOVA) con las puntuaciones finales de los evaluadores en cada una de las tres fases de evaluación. Acto seguido, se compararon las puntuaciones de los estudiantes de posgrado con el juicio experto de los profesores mediante el cálculo del Kappa de Cohen.

Resultados

Las Tablas 1 y 2 resumen los resultados de la primera y segunda intervención analizadas respectivamente.

Según prueban las Tabla 1 y 2, en la primera intervención se detectaron diferencias estadísticamente significativas entre los evaluadores; diferencias que desaparecieron en la segunda intervención. De ello se puede inferir que las discrepancias en la utilización de las rúbricas de la primera intervención se resolvieron en la segunda, cuando los evaluadores puntuaron de un modo similar y, por lo tanto, utilizaron de una forma compartida el sistema de evaluación. Hasta qué punto podía defenderse esta semejanza en la utilización de las rúbricas lo estableció el grado de acuerdo logrado entre pares de evaluadores en ambas intervenciones. La Tabla 3 resume los resultados.

Los coeficientes Kappa mostraron una distribución errática en las tres fases de la primera intervención. Incluso podría afirmarse que los coeficientes empeoraron a medida que se sucedieron las fases de evaluación y que afectaron por igual a estudiantes y profesores. Durante la segunda intervención se alcanzaron valores de 1 en 6 de los 10 coeficientes



IMPACTOS DE LA INNOVACIÓN EN LA DOCENCIA Y EL APRENDIZAJE

calculados y tampoco se detectaron diferencias rotundas en función de quién utilizó las rúbricas: estudiantes y profesores experimentaron un proceso similar de familiarización con el uso de las rúbricas que puede atribuirse al papel desempeñado por las reuniones en las que todos los evaluadores discutieron y negociaron el significado y la aplicación de dichas rúbricas.

Discusión

En las dos intervenciones analizadas, las rúbricas fueron el único recurso común y ayudaron a que los evaluadores utilizaran los mismos criterios, se centraran en los mismos aspectos a evaluar y previnieran la intrusión de prejuicios (Reynolds, 2009; Taub et al., 2011; Venning y Buisman-Pijlman, 2013; Wolf y Stevens, 2007). Ello fue posible por tres razones. En primer lugar, porque se intentó utilizar una terminología adaptada a los estudiantes que tenían que utilizarlas siguiendo las recomendaciones de la literatura especializada (Lapsley y Moody, 2007; O'Donovan, Price, y Rust, 2004; Price y Rust, 1999). En segundo, porque se trabajó con la especificidad de las rúbricas. Si la claridad del recurso era la clave para asegurar la comprensión de los evaluadores, su especificidad garantizaba una aplicación similar por parte de los estudiantes. El diseño de las rúbricas fue específico para la disciplina (Dunbar, Brooks, y Kubicka-Miller, 2006; Torrance, 2007) y los objetivos de aprendizaje, pero también su uso se planteó de manera progresiva por fases y dimensiones. En tercer y último lugar, porque las rúbricas eran analíticas y, en este sentido, transmitieron a los estudiantes información específica (Moskal, 2001; Wolf y Goodwin, 2007) sin requerir de ellos un criterio técnico basado en una impresión general (Dunbar, Brooks, y Kubicka-Miller, 2006).

Sin embargo, las pruebas estadísticas constataron una diferencia importante entre la primera y la segunda intervención, y evidenciaron la constitución de un equipo de evaluación consistente solo cuando se produjeron procesos de discusión y negociación entre los participantes. Nótese, en este sentido, que las parejas de profesores experimentaron una mejora similar a la de los estudiantes en sus grados de acuerdo, y que estas no difirieron mucho en sus niveles de coincidencia con las parejas estudiante-profesor y estudiante-estudiante. Dado que la intervención de la tercera estudiante tampoco alteró significativamente los grados de acuerdo, el factor que más influyó en la construcción de esta comunidad de prácticas fue la participación de los estudiantes de posgrado en el rediseño de las rúbricas y en las sesiones de moderación.

Sin negar los resultados de investigaciones que demuestran la utilidad de los programas de formación (Linn, 1994; Croswell et al., 2011; Welch y James, 2007), este estudio no pudo constatar que este tipo de programas fuera un elemento decisivo o mejorase significativamente el juicio de los participantes (Smith et al., 2013). En la segunda intervención se exploraron formas de participación más profundas de los estudiantes y se lograron mayores



IMPACTOS DE LA INNOVACIÓN EN LA DOCENCIA Y EL APRENDIZAJE

beneficios en lo que al desarrollo del juicio profesional se refiere (Oakleaf, 2008; Sadler, 2009). Coincidimos con Diller y Phelps (2008) en que la colaboración de todos los evaluadores en la creación de las rúbricas ayuda a evitar los problemas que tienen su origen en la interpretación de los criterios y procedimientos. El diseño es al fin y al cabo una construcción del consenso sobre los atributos del objeto a evaluar y sobre su valor relativo en función del contexto y la finalidad educativos.

Pero no basta con comprender el significado de los criterios de evaluación ni con tener una visión global sobre cómo utilizarlos. Aquello que define el juicio experto es su aplicación específica a cada objeto evaluado o la capacidad de interpretar dicho objeto en toda su complejidad. Por lo general, las sesiones de moderación parten de la identificación de las discrepancias y persiguen alcanzar puntuaciones compartidas y refinar el sistema de evaluación mediante la revisión consensuada de los criterios y procedimientos de aplicación. Algunos autores han cuestionado la influencia positiva de este tipo de sesiones de moderación en la fiabilidad y el acuerdo (e.g. Bloxham, 2009; Vardi, 2013). A nuestro juicio, su debilidad radica en que la búsqueda del consenso se realiza cuando los evaluadores ya han emitido una valoración, de manera que la obtención de las mismas puntuaciones es más compleja: primero, porque cada evaluador tiene una idea muy arraigada del valor del objeto evaluado y, segundo, porque la negociación involucra aspectos espurios al proceso evaluador como, por ejemplo, cuestiones de rango y prestigio profesional de los participantes. Por esta razón, preferimos implantar una estrategia de sesión similar a los estudios de Bird y Yucel (2013) y Hughes y Cappa (2007). Nuestra propuesta difirió en que las sesiones de moderación se programaron inmediatamente después de cada sesión de evaluación y antecedieron a la fase final en la que cada evaluador puntuó en solitario. Su objeto no fue alcanzar una comprensión compartida en lo relativo a la definición de los criterios - puesto que esto ya había sido el objeto de la fase de diseño - sino comparar y debatir las diferentes percepciones que cada evaluador tuvo de cada uno de los proyectos expuestos. Así, se logró que los estudiantes de posgrado y los profesores construyeran colectivamente una misma visión de los objetos evaluados.

Conclusión

Nuestro estudio demuestra que un breve programa de formación sobre la naturaleza de los objetos evaluados y el significado y el uso de las rúbricas no es suficiente; hay que incorporar a los estudiantes de posgrado en el equipo de evaluación. En este sentido, se puede afirmar que la mejor contribución de las rúbricas radica en que concretan y realimentan los procesos de negociación y acuerdo que determinan los criterios y procedimientos de evaluación y las percepciones de los propios fenómenos evaluados. Las rúbricas así concebidas fortalecen los



IMPACTOS DE LA INNOVACIÓN EN LA DOCENCIA Y EL APRENDIZAJE

equipos de evaluadores y promueven el surgimiento de comunidades de práctica en donde se puede desarrollar el juicio profesional en contextos reales.

4.1. FIGURA O IMAGEN 1

Tabla 1. ANOVA conjunto por fase en la primera intervención

Fase	Dispersión	Suma de cuadrados	gl.	Media cuadrática	F ^a	Sig. (p<0,05)
1	Inter sujetos	26,37	3	8,79	3,40	0,01
	Intra sujetos	1024,98	396	2,59		
	Total	1051,35	399			
2	Inter sujetos	15,68	3	5,23	2,94	0,03
	Intra sujetos	702,24	396	1,77		
	Total	717,92	399			
3	Between	48,40	3	16,13	6,28	0,00
	Within	1016,40	396	2,57		
	Total	1064,80	399			

^aMáximo valor explicado por el azar = 2,6

4.2. FIGURA O IMAGEN 2

Tabla 2. ANOVA conjunto por fase en la segunda intervención

Fase	Dispersión	Suma de cuadrados	gl.	Media cuadrática	F ^a	Sig. (p<0,05)
1	Inter sujetos	10,37	4	2,58	1,52	0,19
	Intra sujetos	730,46	430	1,69		
	Total	740,79	434			
2	Inter sujetos	8,84	4	2,21	1,07	0,36
	Intra sujetos	882,79	430	2,05		
	Total	891,64	434			
3	Between	7,36	4	1,84	1,60	0,17
	Within	492,49	430	1,14		
	Total	499,85	434			

^aMáximo valor explicado por el azar = 2,4

4.3. FIGURA O IMAGEN 3

Tabla 3. Coeficientes Kappa por parejas de evaluadores y fases en las dos intervenciones

Evaluadores	Primera intervención			Segunda intervención		
	Fase 1	Fase 2	Fase 3	Fase 1	Fase 2	Fase 3
P1-P2	0,44	0,55	0,15	0,57	0,85	1
P1-E1	0,46	0,36	0,75	0,40	1	1
P1-E2	0,52	0,49	0,75	0,46	0,52	1
P1-E3				0,75	0,70	0,82
P2-E1	0,44	0,25	0,01	0,40	0,85	1
P2-E2	0,61	0,29	0,03	0,46	0,64	1
P2-E3				0,51	0,83	0,82
E1-E2	0,51	0,26	0,52	0,21	0,52	1
E1-E3				0,39	0,70	0,82
E2-E3				0,44	0,79	0,82

Nota: P = Profesor. E = Estudiante



IMPACTOS DE LA INNOVACIÓN EN LA DOCENCIA Y EL APRENDIZAJE

5. REFERENCIAS BIBLIOGRÁFICAS

- Beck, R. J., Skinner, W. F., y Schwabrow, L. A. (2013). A study of sustainable assessment theory in higher education tutorials. *Assessment & Evaluation in Higher Education*, 38(3), 326-348.
- Bird, F. L., y Yucel, R. (2013). Improving marking reliability of scientific writing with the Developing Understanding of Assessment for Learning programme. *Assessment & Evaluation in Higher Education*, 38(5), 536-553.
- Bloxham, S. (2009). Marking and moderation in the UK: False assumptions and wasted resources. *Assessment & Evaluation in Higher Education*, 34(2), 209-220.
- Brooks, V. (2012). Marking as judgment. *Research Papers in Education*, 27(1), 63-80.
- Carroll, T. (2005). Induction of teachers into 21st century learning communities: Creating the next generation of educational practice. *The New Educator*, 1(3), 199-204.
- Crotwell Timmerman, B. E., Strickland, D. C., Johnson, R. L., y Payne, J. R. (2011). Development of a 'universal' rubric for assessing undergraduates' scientific reasoning skills using scientific writing. *Assessment & Evaluation in Higher Education*, 36(5), 509-547.
- Diller, K. R., y Phelps, S. F. (2008). Learning Outcomes, Portfolios, and Rubrics, Oh My! Authentic Assessment of an Information Literacy Program. *Portal: Libraries and the Academy*, 8(1), 75-89.
- Duguid, P. (2005). 'The art of knowing': Social and tacit dimensions of knowledge and the limits of the community of practice. *The Information Society*, 21(2), 109-118.
- 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) 115-128.
- Hughes, C., y Cappa, C. (2007). Developing generic criteria and standards for assessment in law: processes and (by)products. *Assessment & Evaluation in Higher Education*, 32(4), 417-432.
- Klein, J. H. (2008). Some directions for research in knowledge sharing. *Knowledge Management Research & Practice*, 6, 41-46.
- Knight, P. T. (2006). The local practices of assessment, *Assessment & Evaluation in Higher Education*, 31(4), 435-452.
- Lapsley, R., y Moody, R. (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), 167-172.



IMPACTOS DE LA INNOVACIÓN EN LA DOCENCIA Y EL APRENDIZAJE

Linn, R. L. (1994). Performance assessment: Policy promises and technical measurement standards. *Educational Researcher*, 23(9), 4-14.

Montero, I., y León, O. (2007). Guía para nombrar los estudios de investigación en Psicología, *International Journal of Clinical and Health Psychology*, 7(3), 847-862.

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

Moskal, B. M., y Leydens, J. A. (2001). Scoring Rubric Development: Validity and Reliability. En L. M. Rudner y W. D. Schafer (Eds.), *Practical Assessment, Research & Evaluation* 2000-2001 (pp. 71-81). Recuperado de <http://www.eric.ed.gov/PDFS/ED458254.pdf>

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), 325-335.

O'Malley, G. (2010). Designing induction as professional learning community. *The Educational Forum*, 74(4), 318-327.

Oakleaf, M. (2008). Dangers and Opportunities: A Conceptual Map of Information Literacy Assessment Approaches. *Portal: Libraries and the Academy*, 8(3), 233-253.

Polanyi, M. (1958). *Personal knowledge: Towards a post-critical philosophy*. London: Routledge & Kegan Paul.

Polanyi, M. (1962). Tacit knowing: Its bearing on some problems in philosophy. *Reviews of Modern Physics*, 34(4), 601-616.

Price, M. (2005). Assessment standards: The role of communities of practice and the scholarship of assessment. *Assessment & Evaluation in Higher Education*, 30(3), 215-230.

Price, M., y Rust, C. (1999). The experience of introducing a common criteria assessment grid across an academic department. *Quality in Higher Education*, 5(2), 133-144.

Reynolds, J., Smith, R., Moskovitz, C., y Sayle, A. (2009). BioTAP: A Systematic Approach to Teaching Scientific Writing and Evaluating Undergraduate Theses. *BioScience*, 59(10), 896-903.

Sadler, D. R. (2009). Indeterminacy in the use of preset criteria for assessment and grading. *Assessment & Evaluation in Higher Education*, 34(2), 159-179.

Smith, C. D., Worsfold, K., Davies, L., Fisher, R., y McPhail, R. (2013). Assessment literacy and student learning: the case for explicitly developing students 'assessment literacy'. *Assessment & Evaluation in Higher Education*, 38(1), 44-60.



IMPACTOS DE LA INNOVACIÓN EN LA DOCENCIA Y EL APRENDIZAJE

Taub, D. J., Servaty-Seib, H. L., Wachter Morris, C.A., Prieto-Welch, S. L., y Werden, D. (2011). Developing Skills in Providing Outreach Programs : Construction and Use of the POSE (Performance of Outreach Skills Evaluation) Rubric. Counseling Outcome Research and Evaluation, 2(1), 59-72.

Torrance, H. (2007). Assessment as Learning? How the Use of Explicit Learning Objectives, Assessment Criteria and Feedback in Post-Secondary Education and Training can come to Dominate Learning. Assessment in Education, 14(3), 281-294.

Tsoukas, H. (2002). Do we really understand tacit knowledge? Paper presented at the Knowledge, Economy and Society Seminar, LSE Department of Information Systems, London, June 14.

Vardi, I. (2013). Effectively feeding forward from one written assessment task to the next. Assessment & Evaluation in Higher Education, 38(5), 599-610.

Venning, J., y Buisman-Pijlman, F. (2013). Integrating assessment matrices in feedback loops to promote research skill development in postgraduate research projects. Assessment & Evaluation in Higher Education, 38(5), 567-579.

Ward, H. C., y Selvester, P. M. (2012). Faculty learning communities: improving teaching in higher education. Educational Studies, 38(1), 111-121.

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.

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.