

¿EVALUAMOS LO QUE REALMENTE IMPORTA? EL DESAFÍO DE LA EVALUACIÓN AUTÉNTICA EN EDUCACIÓN SUPERIOR

Verónica Villarroel¹,
Daniela Bruna²

RESUMEN

El presente ensayo propone la evaluación auténtica como una herramienta necesaria de incorporar en educación superior. Los instrumentos de evaluación de los aprendizajes, que tradicionalmente se utilizan en la universidad, presentan grandes debilidades a la hora de medir el conocimiento profundo de los estudiantes. En cambio, la evaluación auténtica otorga pertinencia al vincular lo que ocurre en las aulas con la vida real y laboral, midiendo saberes en contexto. Se examinan las tres dimensiones que conforman la evaluación auténtica y que facilitan su implementación: 1) realismo, entendido como la inclusión de contextos del mundo del trabajo; 2) desafío cognitivo, que implica medir habilidades cognitivas de orden superior con las que el estudiante debe construir conocimiento; y 3) juicio evaluativo, referido a incluir procesos de retroalimentación que permitan a los estudiantes comprender e integrar en su comportamiento los criterios de buena calidad para transferirlos en otros contextos.

Conceptos clave: educación superior, evaluación auténtica, evaluación para el aprendizaje, pruebas escritas.

DO WE ASSESS WHAT REALLY MATTERS? THE CHALLENGE OF AUTHENTIC ASSESSMENT IN HIGHER EDUCATION

ABSTRACT

The present essay proposes authentic assessment as a tool that is necessary to incorporate into higher education. The instruments traditionally used in universities to assess student learning are poorly adept at measuring in-depth knowledge. Conversely, authentic assessments provide relevancy, by connecting what happens in the classroom with real life, measuring knowledge in context. The three dimensions which comprise an authentic assessment, and favour its implementation, are examined: 1) realism, understood as the inclusion of working world contexts; 2) cognitive challenge, which implies measuring higher order thinking skills that require students to construct knowledge; and 3) evaluative judgement, which refers to the inclusion of feedback processes, that allow students to understand and adopt good quality standards, in order to transfer them to other contexts.

Key concepts: assessment for learning, authentic assessment, higher education, written tests.

¹ Universidad del Desarrollo, Concepción, Chile. Contacto: vvillarroel@udd.cl

² Universidad del Desarrollo, Concepción, Chile. Contacto: dbrunaj@udd.cl

Introducción

La masiva expansión de la educación superior en el mundo ha hecho que la oferta de graduados sea más amplia y, por ende, exista mayor competencia a la hora de postular a un trabajo (Brooks, Waters & Pimlott-Wilson, 2012; Davies, Mangan, Hughes & Slack, 2013; Tholen, James, Warhurst & Commander, 2016; Yeom, 2016). Esto tiene impacto en la empleabilidad, evidenciándose cierto desajuste entre las habilidades y competencias requeridas por el mundo del trabajo y las formadas en la universidad (Figuereido, Biscaia, Rocha & Teixeira, 2017).

Para convertirse en un buen profesional no solo es necesario dominar el conocimiento o las habilidades técnicas de la disciplina, también se exige el desarrollo de habilidades de pensamiento complejo necesarias en el lugar de trabajo (Guzzomi, Male & Miller, 2015; Medland, 2016; Partnership for 21st Century Skills, 2010). Algunas de las competencias más requeridas por los empleadores son la capacidad de resolver problemas, pensar críticamente y comunicarse adecuadamente (Artes, Hooley & Mellors-Bourne, 2016; Kalyuga, Renkl & Pass, 2010; Oliveri & Markle, 2017). Si bien estas habilidades forman parte de los perfiles de egreso declarados en la mayoría de las carreras, no cuentan con un adecuado seguimiento de las trayectorias y logros (Morales y Zambrano, 2016).

En gran parte de las universidades, los programas de pregrado se han movido hacia un modelo basado en competencias (Knight, 2011; Manzanares y Santamaría, 2016). Aun cuando sus ventajas están relacionadas con procesos educativos más dinámicos y prácticos, que buscan la formación integral de los estudiantes y la evaluación de los aprendizajes al momento de su egreso (Pavié, 2011; Tobón, 2013), se ha encontrado que entre las principales dificultades de su implementación está la escasa modificación que han tenido los procesos evaluativos (Contreras, 2010; González y Hassall, 2014; Himmel, 2003). Por un lado, no hay coherencia entre las instituciones de educación superior en torno a lo que involucra la evaluación de competencias (Ampuero y Casas, 2013) y, por otro, se carece de instrumentos de evaluación coherentes y pertinentes para medir las

habilidades declaradas en los perfiles de egreso (Möller y Gómez, 2014; Morales y Zambrano, 2016).

Por su parte, las pruebas escritas son las más utilizadas en educación superior (Gitanjali, 2016; Pfund, Norcross, Hailstorks, Stamm & Christidis, 2018) y, entre ellas, las preguntas de opción múltiple son las de uso más frecuente (Ghosh, Bowles, Ranmuthugala & Brooks, 2017; Villarroel, Bruna, Bustos, Bruna y Márquez, 2018). Los defensores del uso de este tipo de pruebas afirman que reducen el plagio (Richardson, 2015), se asocian a una mayor fiabilidad y validez en la medición del aprendizaje (McColongue, 2012), además de permitir una corrección más eficiente y rápida, especialmente cuando se trabaja con grandes grupos de alumnos (McCabe & O'Connor, 2014). Sin embargo, también figuran otras razones más locales. En las universidades chilenas, existe una heterogeneidad de roles docentes (por ejemplo, investigador, instructor o roles de gestión), sin que necesariamente existan metas, derechos ni deberes claros y diferenciados asociados a ellos (Lolas, 2008). Si bien ha habido un avance en la profesionalización docente, esta es más visible en el área de investigación que en la docencia propiamente tal (Bernasconi, 2008). A pesar de la importancia que se otorga a la enseñanza en la formación de personas y profesionales, la docencia enfrenta cada vez más mayores desafíos que beneficios para quienes la ejercen. La matrícula se ha expandido significativamente, siendo más diversa la composición del estudiantado y con una gran cantidad de alumnos que conforman la primera generación de su familia en acceder a la educación superior (González, 2015). Por otra parte, existen muchos profesores con contrato de tiempo parcial realizando docencia en distintas instituciones a veces, incluso, en condiciones de precariedad laboral (Sisto, 2005). También es posible ver profesionales que, además de su trabajo tradicional, destinan algún tiempo a la docencia, sin que esta sea su rol o preocupación principal, pues constituye una segunda ocupación (Castillo, 2009). En estos casos, el tiempo disponible y destinado a pensar y a perfeccionar el proceso de enseñanza-aprendizaje es limitado.

Si bien el escenario que se describe es complejo, este artículo busca demostrar cómo el sistema de evaluación de respuesta cerrada

y descontextualizada se convierte en una inversión que no renta a futuro. Es decir, el beneficio de ahorrar tiempo en la construcción del instrumento de evaluación y en su corrección, se desvanece cuando los aprendizajes obtenidos por los estudiantes se olvidan rápidamente, lo que deriva en tener que reforzar los mismos contenidos en cursos superiores, gastando tiempo no contemplado para ello.

Se ha visto que este tipo de evaluación tradicional convierte a los estudiantes en aprendices pasivos (Altay, 2014; Struyven & Devesa, 2016) que buscan memorizar más que comprender (Flores, Veiga-Simao, Barros & Pereira, 2015), promoviendo un enfoque superficial del aprendizaje por sobre el profundo (Beyaztas & Senemoglu, 2015; Endedijk & Vermunt, 2013). Un aprendizaje profundo se relaciona con el número de conexiones de conceptos e ideas, el nivel de razonamiento y el uso de procesos de automonitoreo que los estudiantes necesitan desplegar y usar para dar una respuesta. Es decir, el nivel de profundidad del conocimiento está relacionado con la calidad del procesamiento cognitivo que un ítem requiere para ser resuelto (Webb, 2012).

¿Por qué es un problema construir pruebas escritas focalizadas en la memorización? Hay tres razones de peso. Primero, recordar representa el nivel más bajo en la evaluación del conocimiento (Gitajali, 2016) y, aunque es cierto que los estudiantes necesitan retener conceptos, hechos o cifras, aprender no se reduce a ello. Los estudiantes requieren desarrollar e integrar habilidades para participar constructivamente en el mundo actual, resolviendo problemas donde es necesario analizar información, sintetizar hallazgos, aplicar conceptos de manera contextualizada, reflexionar críticamente acerca de lo aprendido, tomar y comunicar decisiones, e innovar y crear nuevas soluciones (O'Sullivan & Dallas, 2017; Partnership for 21st Century Skills, 2010; Tucker, 2014). Segundo, cuando los estudiantes memorizan mecánicamente los contenidos para rendir una prueba, existe una mayor probabilidad de olvidarlos rápidamente (Rawson, Dunlosky & Sciarтели, 2013). Por el contrario, cuando deben usar habilidades cognitivas de orden superior para responder una prueba, los estudiantes obtienen una comprensión más profunda del conocimiento (Jensen, McDaniel, Woodard & Kummer, 2014),

mostrando un recuerdo más estable de lo aprendido a lo largo del tiempo (Rawson, et al., 2013).

Junto con estos dos argumentos, es importante destacar un tercero que es central: el impacto de la evaluación en el proceso de aprendizaje de los alumnos y de enseñanza de los profesores (Bosco & Ferns, 2014; Kearney, Perkins & Kennedy-Clark, 2015) debido al efecto colateral del proceso evaluativo (backwash effect, en inglés) (Boud, 2007; 2010; Watkins, Dahlin & Ekholm, 2005). En el caso de los estudiantes, la evaluación les informa lo que es importante aprender y cómo hacerlo, modelando y consolidando el uso de las habilidades cognitivas (Bloxham, den Outer, Hudson & Price, 2016; Vu & Dall'Alba, 2014). Si un examen solo mide habilidades memorísticas, el alumno concluirá que aprender es repetir información; en cambio, si la prueba solicita aplicar lo aprendido para resolver algún problema, entenderá que aprender es hacer uso del conocimiento para construir algo nuevo. Por estos motivos, mejorar la evaluación es una manera de impactar en la calidad y profundidad del aprendizaje de los alumnos (Ashford-Rowe, Herrington & Brown, 2014; Edström, 2008; Syafei, 2012).

En el caso de los profesores, las investigaciones han mostrado que, al introducir cambios en la forma de evaluar, implícita e indirectamente se impacta en la enseñanza (Boud, 2007; Boud & Molloy, 2013). Cuando el docente implementa evaluaciones para el aprendizaje, motivando la construcción de conocimiento, comienza también a cambiar su enfoque hacia la enseñanza del aprendizaje profundo, usando para ello prácticas problematizadoras donde los estudiantes tienen un rol más activo (Biggs & Tang, 2011).

Por lo tanto, uno de los mayores desafíos de la educación superior es aplicar métodos de evaluación pertinentes que permitan medir lo que los estudiantes pueden hacer con el conocimiento de lo que han aprendido de manera contextualizada, respondiendo al perfil de egreso comprometido y a las necesidades del mundo del trabajo (Nicholson Perry, Donovan, Knight & Shires, 2017; Oliveri & Markle, 2017). La evaluación auténtica permite hacerlo.

¿Por qué evaluación auténtica?

La evaluación auténtica permite relacionar lo que ocurre en la sala de clases con situaciones de la vida diaria y profesional, creando un vínculo entre lo que se aprende y su uso para la resolución de problemas cotidianos (Gulikers, Bastiaens & Kirschner, 2004; Neely & Tucker, 2012). También involucra a los estudiantes con problemas o preguntas importantes que resultan valiosos más allá del aula. En este sentido, las tareas de evaluación son réplicas o analogías de problemas del mundo externo y pretenden que los estudiantes usen el conocimiento para mostrar desempeños efectivos y creativos través de su aplicación en problemas auténticos (Herrington & Kervin, 2007; Kearney, 2013; Saye, 2013).

Son tres las dimensiones que constituyen el núcleo de la evaluación auténtica: realismo, desafío cognitivo y juicio evaluativo. El realismo se refiere a contextualizar y ubicar la evaluación en función de problemas profesionales o de la vida diaria, de modo que los estudiantes utilicen lo que han aprendido para ofrecer una respuesta o solución que les permita valorar el conocimiento como un medio para comprender y transformar el mundo en el que viven. El desafío cognitivo se refiere a la necesidad de que la evaluación mida habilidades cognitivas de orden superior, movilizándolo al estudiante para utilizar y transferir el conocimiento. Finalmente, como la evaluación auténtica está vinculada con el mundo del trabajo, busca que los estudiantes desarrollen e incorporen criterios de calidad que les permitan juzgar, objetivamente, su propio trabajo y el de sus pares, reflexionando acerca de las mejoras que puedan tener. De esta forma, logran la autonomía para monitorear y regular su proceso de aprendizaje (Villaruel, Bloxham, Bruna, Bruna & Herrera-Seda, 2018).

En relación con sus beneficios, los estudios señalan que la evaluación auténtica impacta en la calidad y en la profundidad del aprendizaje logrado por el estudiante (Dochy & McDowell, 1997) y en el desarrollo de habilidades cognitivas de orden superior (Ashford-Rowe, et al., 2014; Koçyiğit & Zembat, 2013). Asimismo, apoya el desarrollo en ellos de una confianza personal

(Martínez, O'Brien, Roberts & Whyte, 2018) y los prepara para la autonomía (Carter, Sidebotham, Crreedy, Fenwick & Gamble, 2015; Raymond, Homer, Smith & Gray, 2012). Además, mejora el compromiso académico (Kearney & Perkins, 2014), la motivación para el proceso de aprendizaje (Nicol, Thompson & Breslin, 2014), la capacidad de autorregulación (Ling Lau, 2013), y la metacognición y la autorreflexión del estudiante (Vanaki & Memarian, 2009). A continuación, se describirán los tres ámbitos característicos de la evaluación auténtica: realismo, desafío cognitivo y juicio evaluativo.

La evaluación auténtica simula contextos realistas y problematizadores

En la evaluación auténtica, la validez y pertinencia de lo que se mide es fundamental. Para facilitar la selección de lo que se va a evaluar, se propone revisar tres fuentes: el perfil de egreso, los resultados de aprendizaje del curso y los requerimientos del mundo del trabajo para esa disciplina. Estos tres elementos permiten aprehender lo nuclear de la asignatura, aportando pertinencia a la evaluación a partir de la reflexión en torno a tres preguntas:

- ¿cómo se conecta y contribuye esta asignatura con las competencias del perfil de egreso en esta carrera?
- ¿qué serán capaces de hacer los estudiantes al final del curso?, y
- ¿las capacidades adquiridas en el curso permiten a los graduados responder a las tareas o funciones que necesita la profesión? (Villarroel, Bloxham, et al., 2018).

Entonces, una de las cosas que distingue a la evaluación auténtica es su realismo (Bosco & Ferns 2014; Saye 2013). Esto se refiere a la simulación de situaciones reales o del mundo real que funciona como un proxy (mejor modelo) para el desempeño profesional. Al crear una situación problemática, ubicamos al estudiante en un contexto real que lo impulsa a tomar decisiones respecto de qué debe hacer. De esta manera, el alumno no solo debe reproducir el contenido del curso, sino que tiene que discriminar qué áreas de su aprendizaje son necesarias para responder a la pregunta.

Al diseñar una evaluación, el realismo se puede conseguir presentando un contexto real que describa y entregue un marco en el que se resuelva un problema, tanto en pruebas escritas como en tareas basadas en el desempeño. La inclusión del contexto brinda autenticidad en una prueba escrita, en preguntas de resolución de problemas, preguntas de desarrollo breves y extensas, análisis de casos e incluso preguntas de respuestas múltiples. La información presentada en un contexto realista y problematizado puede mostrar más acerca de la perspectiva de un fenómeno o presentar límites o restricciones que son necesarios de considerar y analizar para responder al problema (Villarroel, Bloxham, et al., 2018).

La evaluación auténtica mide habilidades cognitivas de orden superior

La evaluación auténtica posee un nivel de complejidad cognitiva que demanda al estudiante construir conocimiento, haciendo uso de habilidades cognitivas de orden superior (Thornburn, 2008). A través del estímulo de la evaluación se generan procesos de resolución de problemas, aplicación del conocimiento y toma de decisiones que corresponden al desarrollo de habilidades cognitivas y metacognitivas, y no solo a la repetición del conocimiento declarativo o conceptual (Newmann, King & Carmichael, 2007).

Según la taxonomía de Bloom, Masia y Krathwohl (1984) y otras actualizaciones como las de Anderson y Krathwohl (2001), las habilidades cognitivas de orden superior estarían relacionadas con el análisis, la evaluación y la creación. Inicialmente, se planteó que esta secuencia de dominios cognitivos era jerárquica y acumulativa (por ejemplo, para lograr habilidades analíticas era necesario dominar habilidades previas como aplicar, comprender o recordar). Sin embargo, hoy en día se sabe que para desarrollar habilidades más desafiantes y de mayor jerarquía, los estudiantes deben utilizar e integrar aquellas de niveles inferiores (Anderson & Krathwohl 2001; Jensen, McDaniel, Woodard & Kummer, 2014).

En su estudio, Villarroel, Bruna, Bruna, Herrera y Márquez (2017) analizaron de 4.401 ítems de 244 pruebas de tercer y cuarto

año de seis programas de pregrado: Enfermería, Ingeniería Comercial, Psicología, Tecnología Médica, Fonoaudiología y Bioingeniería de distintas casas de estudio de Chile. Del total de ítems analizados, 64,2% evaluaban habilidades cognitivas de bajo nivel, es decir, vinculadas con el reconocimiento de información memorizada; 23,6% medían habilidades cognitivas analíticas de complejidad moderada, como comparar, argumentar y concluir acerca de conocimiento; y 12,2% valoraban habilidades de orden superior referidas a evaluar, criticar o crear conocimiento. Los ítems de respuesta cerrada tienen mayores debilidades en la medición de habilidades complejas que son cognitivamente desafiantes para los estudiantes, lo cual evidencia que las pruebas escritas no están midiendo pensamiento complejo ni aprendizaje profundo, como se esperaría en la educación superior, más aún en los cursos más avanzados de una carrera.

La evaluación auténtica involucra retroalimentación para desarrollar juicio evaluativo

La retroalimentación es importante en cualquier tipo de evaluación. Puede ser una de las influencias más poderosas en el aprendizaje y en el logro de los estudiantes (Hattie & Timperley, 2007; Jonsson, 2012). Para que un proceso de retroalimentación sea auténtico es necesario incluir actividades que preparen a los estudiantes para lo que tendrán que hacer en el mundo trabajo (Kearney et al., 2015). Involucrarlos en procesos de retroalimentación auténticos les ayuda a lograr conocimientos, habilidades y actitudes que apoyan sus procesos aprendizaje continuo (Boud, 2010), promueven el desarrollo de la autonomía (Carter, et al., 2015) y la práctica reflexiva (Tait-McCutcheon, Drake & Sherley, 2011).

La retroalimentación, por su parte, tiene un carácter dialógico (Carless, Salter, Yang & Lam, 2011), ya que los estudiantes tienen un papel activo en la construcción del significado de lo que aprenden (López-Pastor y Sicilia-Camacho, 2017) a través de una relación intersubjetiva con el docente, intercambiando y negociando puntos de vista (Mercer & Howe, 2012). Esto los involucra activamente en el análisis, defensa y recepción de los comentarios del docente (Boud & Molloy, 2013), y los ayuda a comprender y a buscar información

para mejorar su desempeño en actividades futuras (Lipnevich, Berg & Smith, 2016).

Una forma de hacerlo es involucrar a los estudiantes en la construcción de los criterios de evaluación de una prueba o tarea de desempeño. El acto de crear conjuntamente los criterios de corrección involucra a los estudiantes en una comprensión profunda del conocimiento (O'Donovan, Price & Rust, 2008), ya que deben volver a estudiar, revisar y buscar información para crear la pauta o rúbrica. Otra estrategia se relaciona con involucrarlos en procesos de auto y coevaluación que promuevan la capacidad del alumno para juzgar su propio trabajo y el de otros, como ocurre en el mundo del trabajo. Dichas prácticas de evaluación promueven el desarrollo de un “juicio evaluativo” (Boud, Ajjawi, Dawson & Tai, 2018), que contribuye a que los estudiantes se conviertan en profesionales autónomos en el futuro.

Conclusiones

La evaluación del aprendizaje es una de las prácticas docentes que mayor impacto ha mostrado tener sobre el aprendizaje de los estudiantes. La forma en que este es evaluado, configura la calidad de los resultados, habilidades y competencias que alcanzarán los aprendices. Así mismo, aprender tiene relación con hacer uso del conocimiento para algo. Ese algo puede ser la comprensión de un fenómeno social, la resolución de un problema disciplinar, o bien la evaluación de la calidad o efectividad de un producto. Lo importante es aplicar y construir conocimiento y no solo reproducirlo literalmente, ya que la memorización es un medio para aprender, pero no es el fin en sí mismo. En ese sentido, la utilización del modelo de evaluación auténtica aporta mayores beneficios en términos de la calidad del aprendizaje, por la vinculación que esta tiene con la vida real y el mundo laboral.

Se espera que este trabajo sea de utilidad en los procesos de diseño de evaluaciones en las instituciones de educación universitaria, a nivel nacional e internacional, promoviendo la reflexión en torno a estas temáticas.

Referencias

- Altay, B. (2014). User-centered design through learner-centered instruction. *Teaching in Higher Education*, 19(2), 138-155. <https://doi.org/10.1080/13562517.2013.827646>
- Ampuero, N. y Casas, M. (2013). *Evaluación de aprendizajes en el contexto de innovaciones curriculares en el pregrado en universidades chilenas*. Santiago de Chile: Centro Interuniversitario de Desarrollo, Cinda.
- Anderson, L. & Krathwohl, D. A. (2001). *Taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Addison Wesley Longman.
- Artes, J., Hooley, T., & Mellors-Bourne, R. (2016). *Employability: A review of literature 2012-2016*. New York: HEA.
- Ashford-Rowe, K., Herrington, J., & Brown, C. (2014). Establishing the critical elements that determine authentic assessment. *Assessment & Evaluation in Higher Education*, 39(2), 205-222. <https://doi.org/10.1080/02602938.2013.819566>
- Bernasconi, A. (2008). La profesionalización de la academia en Chile. *Calidad en la Educación*, 28, 15-27. <https://doi.org/10.31619/caledu.n28.199>
- Beyaztas, D. & Senemoglu, N (2015). Learning approaches of successful students and factors affecting their learning approaches. *Eğitim ve Bilim*, 40(179), 193-216. <https://doi.org/10.15390/EB.2015.4214>
- Biggs, J. & Tang, C. (2011). *Teaching for quality learning at university: What the student does*. Maidenhead, Berkshire: Open University Press.
- Bloom, B., Masia, B., & Krathwohl, D. (1964). *Taxonomy of educational objectives*. New York: McKay.
- Bloxham, S., den-Outer, B., Hudson, J., & Price, M. (2016). Let's stop the pretence of consistent marking: Exploring the multiple limitations of assessment criteria. *Assessment & Evaluation in Higher Education*, 41(3), 466-481. <https://doi.org/10.1080/02602938.2015.1024607>
- Bosco, A.M. & Ferns, S. (2014). Embedding of authentic assessment in work-integrated learning curriculum. *Asia-Pacific Journal of Cooperative Education*, 15(4), 281-290.
- Boud, D. (2007). Reframing assessment as if learning were important. En D. Boud & N. Falchikov (Eds.), *Rethinking assessment in higher education; learning for the longer term* (pp. 14-25). London: Routledge.
- Boud, D. (2010). Sustainable assessment: Rethinking assessment for the learning society. *Studies in Continuing Education*, 22(2), 151-167. <https://doi.org/10.1080/713695728>

- Boud, D., Ajjawi, R., Dawson, P., & Tai, J. (2018). *Developing evaluative judgement in higher education. Assessment for knowing and producing quality work*. New York: Routledge.
- Boud, D. & Molloy, E. (2013). Rethinking models of feedback for learning: The challenge of design. *Assessment & Evaluation in Higher Education*, 38(6), 698-712. <https://doi.org/10.1080/02602938.2012.691462>
- Brooks, R., Waters, J., & Pimlott-Wilson, H. (2012). International education and the employability of UK students. *British Educational Research Journal*, 38(2), 281-298.
- Carless, D., Salter, D., Yang, M., & Lam, J. (2011). Developing sustainable feedback practices. *Studies in Higher Education*, 36(4), 395-407. <https://doi.org/10.1080/03075071003642449>
- Carter, A., Sidebotham, M., Creedy, D., Fenwick, J., & Gamble, J. (2015). Strengthening partnership: The involvement of health care providers in the evaluation of authentic assessment within midwifery undergraduate education. *Nurse Education in Practice*, 15(4), 327-32. <https://doi.org/10.1016/j.nepr.2015.01.013>
- Castillo, M. (2009). La profesión docente. *Revista Médica de Chile*, 138(7), 902-907. <https://doi.org/10.4067/S0034-98872010000700017>
- Contreras, G. (2010). Diagnóstico de dificultades de la evaluación de aprendizaje en la universidad: un caso particular de Chile. *Educación y Educadores*, 13(2), 2019-238.
- Davies, P., Mangan, J., Hughes, A., & Slack, K. (2013). Labour market motivation and undergraduates' choice of degree subjects. *British Educational Research Journal*, 39(2), 361-382. <https://doi.org/10.1080/01411926.2011.646239>
- Dochy, F. & McDowell, L. (1997). Assessment as a tool for learning. *Studies in Educational Evaluation*, 23(4), 279-298.
- Edström, K. (2008). Doing course evaluation as if learning matters most. *Higher Education Research & Development*, 27(2), 95-106. <https://doi.org/10.1080/07294360701805234>
- Endedijk, M. D. & Vermunt, J. D. (2013). Relations between student teachers' learning patterns and their concrete learning activities. *Studies in Educational Evaluation*, 39(1), 56-65. <https://doi.org/10.1016/j.stueduc.2012.10.001>
- Figueredo, H., Biscaia, R., Rocha, V., & Teixeira, P. (2017). Should we start worrying? Mass higher education, skill demand and increasingly complex landscape of young graduates' employment. *Studies in Higher*

- Education*, 42(8), 1401-1420. <https://doi.org/10.1080/03075079.2015.1101754>
- Flores, M. A., Veiga-Simão, A. M. Barros, A., & Pereira, D. (2015). Perceptions of effectiveness, fairness and feedback of assessment methods: A study in higher education. *Studies in Higher Education*, 40(9), 1523-1534. <https://doi.org/10.1080/03075079.20140881348>
- Ghosh, S., Bowles, M., Ranmuthugala, D., & Brooks, B. (2015). Authentic assessment in seafarer education: Using literature review to investigate its' validity and reliability through rubrics. *WMU J Marit Aff*, 15(2), 317-336. <https://doi.org/10.1007/s13437-015-0094-0>
- Gitanjali, M. (2016). The three Rs of written assessment: The JIPMER experience. *Journal of Pharmacology and Pharmacotherapeutics*, 7(3), 115-119. <https://doi.org/10.4103/0976-500X.189650>
- González, C. (2015). Perspectivas y desafíos de la docencia en la educación superior chilena. En A. Bernasconi (Ed.), *Educación superior en Chile. Transformación, desarrollo y crisis* (pp. 371-408). Santiago de Chile: Ediciones Universidad Católica de Chile.
- González, A. y Hassall, J. L. (2014). Consolidación de la formación por competencias en la universidad española: estudio de caso. *Educación XXI*, 2(17), 145-168. <https://doi.org/10.5944/educxx1.17.2.11483>
- Gulikers, J., Bastiaens, T., & Kirschner, P. (2004). A five-dimensional framework for authentic assessment. *Educational Technology Research and Development*, 52(3), 67-86. <https://doi.org/10.1007/BF02504676>
- Guzzomi, A., Male, S., & Miller, K. (2015). Students' responses to authentic assessment designed to develop commitment to performing at their best. *European Journal of Engineering Education*, 42(3), 1-22. <https://doi.org/10.1080/03043797.2015.1121465>
- Hattie J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112. <https://doi.org/10.3102/003465430298487>
- Herrington, J. & Kervin, L. (2007). Authentic learning supported by technology: Ten suggestions and cases of integration in classrooms. *Educational Media International*, 44(3), 219-236. <https://doi.org/10.1080/09523980701491666>
- Himmel, E. (2003). Evaluación de aprendizajes en educación superior: una reflexión necesaria. *Revista de Pensamiento Educativo*, 33, 199-211.
- Jensen, J.L., McDaniel, M.A., Woodard, S.M. (2014). Teaching to the test... or testing to teach: Exams requiring higher order thinking skills encourage greater conceptual understanding. *Educational Psychology Review*, 26(2), 307- 329. <https://doi.org/10.1007/s10648-013-9248-9>

- Jonsson, A. (2012). Facilitating productive use of feedback in higher education. *Active Learning in Higher Education*, 14(1), 63-76. <https://doi.org/10.1177/1469787412467125>
- Kalyuga, S., Renkl, A., & Paas, F. (2010). Facilitating flexible problem solving: A cognitive load perspective. *Educational Psychology Review*, 22, 175-186. <https://doi.org/10.1007/s10648-010-9132-9>
- Kearney, S. (2013). Improving engagement: The use of authentic self-and peer-assessment for learning' to enhance the student learning experience. *Assessment & Evaluation in Higher Education*, 38(7), 875-891. <https://doi.org/10.1080/02602938.2012.751963>
- Kearney, S. & Perkins, T. (2014). Engaging students through assessment: The success and limitations of the ASPAL (Authentic Self- and Peer-Assessment for Learning) model. *Journal of University Teaching and Learning Practice*, 11(3), 1-13.
- Kearney, S., Perkins, T., & Kennedy-Clark, S. (2015). Using self-and peer-assessments for summative purposes: Analysing the relative validity of the AASL (Authentic Assessment for Sustainable Learning) model. *Assessment & Evaluation in Higher Education*, 41(6), 1-14. <https://doi.org/10.1080/02602938.2015.1039484>
- Knight, B. (2011). Training in professional psychology in the US: An increased focus on competency attainment. *Australian Psychologist*, 46, 140-141. <https://doi.org/10.1111/j1742-9544.2011.00026.x>
- Koçyiğit, S. & Zembat, R. (2013). The effects of authentic tasks on preservice teachers' attitudes towards classes and problem-solving skills. *Educational Sciences: Theory and Practice*, 13(2), 1045-1051.
- Ling Lau, K. (2013). Chinese language teachers' perception and implementation of self-regulated learning-based instruction. *Teacher and Teaching Education*, 31, 56-66. <https://doi.org/10.1016/j.tate.2012.12.001>
- Lipnevich, A. A., Berg, D. A., & Smith, J. K. (2016). Toward a model of student response to feedback. G. T.L Brown & L.R. Harris (Eds.), *Handbook of Human and Social Conditions in Assessment* (pp. 169-185). New York: Routledge.
- Lolas, F. (2008). Sobre académicos, academia y universidad. Necesidad de una clarificación. *Calidad en la Educación*, 28, 30-37. <https://doi.org/10.31619/caledu.n28.200>
- López-Pastor, V. & Sicilia-Camacho, A. (2017). Formative and shared assessment in higher education. Lessons learned and challenges for the future. *Assessment & Evaluation in Higher Education*, 42(2), 77-97. <https://doi.org/10.1080/02602938.2015.1083535>

- Manzanares, A. y Santamaría, J. S. (2016). La dimensión pedagógica de la evaluación por competencias y la promoción del desarrollo profesional en el estudiante universitario. *Revista Iberoamericana de Evaluación Educativa*, 5(1), 187-202.
- Martínez, M., O'Brien, M., Roberts, K., & Whyte, D. (2018). Critical pedagogy an assessment in higher education: The ideal of "authenticity" in learning. *Active Learning in Higher Education*, 19(1), 9-21. <https://doi.org/10.1177/1469787417723244>
- McCabe, A. & O'Connor, U. (2014). Student-centred learning: the role and responsibility of the lecturer. *Teaching in Higher Education*, 19(4), 350-359. <https://doi.org/10.1080/13562517.2013.860111>
- McColongue, T. (2012). But is it fair? Developing students' understanding of grading complex written work through peer assessment. *Assessment & Evaluation in Higher Education*, 37(1), 113-123. <https://doi.org/10.1080/02602938.2010.515010>
- Medland, E. (2016). Assessment in higher education: Drivers, barriers and directions for change in the UK. *Assessment & Evaluation in Higher Education*, 41(1), 81-96. <https://doi.org/10.1080/02602938.2014.982072>
- Mercer, N. & Howe, C. (2012). Explaining the dialogic processes of teaching and learning: The value of sociocultural theory. *Learning, Culture and Social Interaction*, 1, 12-21. <https://doi.org/10.1016/j.lcsi.2012.03.001>
- Möller, I. y Gómez, H. (2014). Coherencia entre perfiles de egreso e instrumentos de evaluación en carreras de educación básica en Chile. *Calidad en la Educación*, 41, 17-49. <https://doi.org/10.4067/s0718-45652014000200002>
- Morales, S. y Zambrano, H. (2016). Coherencia evaluativa en formación universitaria por competencias: estudio en futuros educadores de Chile. *Infancia Imágenes*, 15(1), 9-26. <https://doi.org/10.14483/udistrital.jour.infimg.2016.1.a01>
- Neely, P. & Tucker, J. (2012). Using business simulations as authentic assessment tools. *American Journal of Business Education*, 5(4), 449-456.
- Newmann, F., King, B., & Carmichael, D. (2007). *Authentic instruction & assessment*. Iowa: Iowa Department of Education.
- Nicholson Perry, K., Donovan, M., Knight, R., & Shires, A. (2017). Addressing professional competency problems in clinical psychology trainees. *Australian Psychologist*, 52(2), 121-129. <https://doi.org/10.1111/ap.12268>

- Nicol, D., Thomson, A., & Breslin, C. (2014). Rethinking feedback practices in higher education: A peer review perspective. *Assessment & Evaluation in Higher Education*, 39(1), 102-122. <https://doi.org/10.1080/02602938.2013.795518>
- O'Donovan, B., Price, M., & Rust, C. (2008). Developing student understanding of assessment standards: A nested hierarchy of approaches. *Teaching in Higher Education*, 13(2), 205-217.
- Oliveri, M. E. & Markle, R. (2017). Continuing a culture of evidence: Expanding skills in higher education. *ETS Research Report*, 1, 1-8. <https://doi.org/10.1002/ets2.12137>
- O'Sullivan, M. K. & Dallas, K. B. (2017). A collaborative approach to implementing 21st century skills in a high school senior research class. *Education Libraries*, 33(1), 3-9. <https://doi.org/10.26443/el.v33i1.284>
- Partnership for 21st Century Skills. (2010). *American Management Association Critical Skills Survey*. Tucson: Autor.
- Pavié, A. (2011). Enfoque basado en competencias: orientaciones sobre sus procedimientos de evaluación. *Revista Electrónica de Desarrollo de Competencias (REDEC)*, 1(7), 114-147.
- Pfund, R., Norcross, J., Hailstorks, R., Stamm, K., & Christidis, P. (2018). Introduction to Psychology: Course, purposes, learning outcomes, and assessment practices. *Teaching of Psychology*, 45(3), 213-219. <https://doi.org/10.1177/0098628318779257>
- Richardson, J. T. (2015). Coursework versus examinations in end-of-module assessment: a literature review. *Assessment & Evaluation in Higher Education*, 40(3), 439-455.
- Rawson, K., Dunlosky, J., & Sciarrelli, S. (2013). The power of successive relearning: Improving performance on course exams and long-term retention. *Educational Psychology Review*, 25, 523-548. <https://doi.org/10.1007/s10648-013-9240-4>
- Raymond, J., Homer, C., Smith, R., & Gray, J. (2012). Learning through authentic assessment. An evaluation of a new development in the undergraduate midwifery curriculum. *Nurse Educational and Practice*, 13(5), 471-476. <https://doi.org/10.1016/j.nepr.2012.10.006>
- Saye, J. (2013). Authentic pedagogy: Its presence in social studies classrooms and relationship to student performance on state-mandated tests. *Theory & Research in Social Education*, 41, 89-132. <https://doi.org/10.1080/00933104.2013.756785>
- Sisto, V. (2005). Flexibilización laboral de la docencia universitaria y la gestación de la universidad sin órganos Un análisis desde la

- subjetividad laboral del docente en condiciones de precariedad. En P. Gentili y B. Levy (Eds.), *Espacio público y privatización del conocimiento: estudios sobre políticas universitarias en América Latina* (pp. 523-574). Buenos Aires: Consejo Latinoamericano de Ciencias Sociales, CLACSO.
- Struyven, K. & Devesa, J. (2016). Students' perceptions of novel forms of assessment. En G. T. L. Brown & L. R. Harris (Eds.), *Handbook of human and social conditions in assessment* (pp. 129-144). New York: Routledge.
- Syafei, M. (2012). Backwash effects of portfolio assessment in academic writing classes. *TEFLIN Journal*, 23(2), 206-221.
- Tait-McCutcheon, S., Drake, M., & Sherley, B. (2011). From direct instruction to active construction: Teaching and learning basic facts. *Mathematics Education Research Journal*, 23(3), 321. <https://doi.org/10.1007/s13394-011-0018-z>
- Tholen, G., James, S., Warhurst, C., & Commander, L. (2016). Higher education, graduate skills and the skills of graduates: The case of graduates as residential sales estate agents. *British Educational Research Journal*, 42(3), 508-523. <https://doi.org/10.1002/berj.3222>
- Thornburn, M. (2008). Articulating a Merleau-Pontain phenomenology of physical education: The quest for active student engagement and authentic assessment in high-stakes examination awards. *European Physical Education Review*, 4(2), 263-280. <https://doi.org/10.1177/1356336X08090709>
- Tobón, S. (2013). *Formación integral y competencias, pensamiento complejo, currículo, didáctica y evaluación*. Bogotá: ECOE Ediciones.
- Tucker, S. Y. (2014). Transforming pedagogies: Integrating 21st century skills and Web 2.0 technology. *Turkish Online Journal of Distance Education*, 15(1), 166-173. <https://doi.org/10.17718/tojde.32300>
- Vanaki, Z. & Memarian, R. (2009). Professional ethics: Beyond the clinical competency. *Journal Professional Nursing*, 25, 285-291. <https://doi.org/10.1016/j.profnurs.2009.01.009>
- Villarroel, V., Bloxham, S., Bruna, D., Bruna, C., & Herrera-Seda, C. (2018). Authentic assessment: Creating a blueprint for course design. *Assessment & Evaluation in Higher Education*, 83(5), 840-854. <https://doi.org/10.1080/02602938.2017.14123969>
- Villarroel, D. Bruna, D., Bruna, C., Herrera, C., y Márquez, C. (2017). Implementación de la evaluación auténtica en educación superior. Un análisis del cambio en la construcción de pruebas escrita. En O. Jerez y C. Silva (Eds.), *Innovando en educación superior: experiencias*

clave en Latinoamérica y el Caribe 2016-2017. Volumen 2: Metodologías activas de enseñanza y aprendizaje (pp. 249-256). Santiago de Chile: Facultad de Economía y Negocios, Universidad de Chile.

- Villarroel, V., Bruna, D., Bustos, C., Bruna, C., y Márquez, C. (2018). Análisis de pruebas escritas bajo los principios de la evaluación auténtica. Estudio comparativo entre carreras de la salud y otras carreras de dos universidades de la Región del Biobío. *Revista Médica de Chile*, 146(1), 46-52. <https://doi.org/10.4067/s0034-98872018000100046>
- Vu, T. & Dall'Alba, G. (2014). Authentic assessment for student learning: An ontological conceptualization. *Educational Philosophy and Theory*, 46(7), 778-791. <https://doi.org/10.1080/00131857.2013.795110>
- Watkins, D., Dahlin, B., & Ekholm, M. (2005). Awareness of backwash effect of assessment: A phenomenographic study of the views of Hong Kong and Swedish lecturers. *Instructional Science*, 33, 283-309. <https://doi.org/10.1007/s11251-005-3002-8>
- Webb, N. (2002). *Alignment analysis of State F Language Arts Standards and Assessments Grades 5, 8, and 11*. Washington, DC: Council of Chief State School Officers.
- Yeom, M. (2016). Critical reflection on the massification of higher education in Korea: Consequences for graduate employment and policy issues. *Journal of Education and Work*, 29(1), 48-63. <https://doi.org/10.1080/13639080.2015.1049026>

Recibido: 18/07/2018

Aceptado: 22/05/2019