

Retos en la reforma de la evaluación en educación superior: una mirada desde la lejanía

Challenges in reforming higher education assessment: a perspective from afar

Boud, D. 

Universidad de Deakin (Melbourne, Australia), Universidad de Tecnología de Sydney (Australia) & Universidad de Middlesex (Londres, Reino Unido)

Abstract

Can we be sure that assessment in higher education meets the need of developing and assuring high quality learning outcomes? Current assessment is typically a collection of conventional practices that have never been seriously questioned. Ten years ago, as part of a national project, representatives from Australian universities came together to identify an agenda for change in assessment. The resulting document—Assessment 2020: Seven propositions for assessment reform in higher education—focused on how assessment needed to change to support long term learning. That is, not how students can pass the next exam, but learning that is useful beyond the point of graduation.

From a learning-centred view this paper examine progress on assessment reform in universities internationally from the perspective of one of the players. It starts by considering Assessment 2020 to see where action is still needed. It reviews some of the major shifts in assessment in higher education and considers their implications. These include the move from comparing students (norm-referencing) to judging outcomes against standards (standards-based); and importantly, the conceptual shift from the single purpose of assessment as certifying students to multiple purposes including aiding learning and building the capacity of students to make their own judgements.

Keywords: Assessment; Assessment reform; Higher Education; Evaluative Judgement

Resumen

¿Podemos estar seguros de que la evaluación en la educación superior satisface la necesidad de desarrollar y garantizar resultados de aprendizaje de alta calidad? La evaluación actual es típicamente una colección de prácticas convencionales que nunca han sido seriamente cuestionadas. Hace diez años, como parte de un proyecto nacional, representantes de universidades australianas se reunieron para identificar una agenda para el cambio en la evaluación. El documento resultante *Evaluación 2020: siete propuestas para la reforma de la evaluación en la educación superior*, se centró en cómo debía cambiar la evaluación para apoyar el aprendizaje a lo largo de la vida. Es decir, no se trata de cómo los estudiantes pueden aprobar el próximo examen, sino de que aprendan lo que será útil más allá del momento de la graduación.

Desde una perspectiva centrada en el aprendizaje, este documento analiza a nivel internacional el progreso que se ha producido en las universidades en torno a la reforma de la evaluación desde la perspectiva de uno de sus actores. Se inicia considerando el documento Evaluación 2020 para ver dónde aún se necesitan acciones. Se revisa algunos de los principales cambios en la evaluación en la educación superior y sus implicaciones. Estos incluyen el tránsito de comparar estudiantes (evaluación referida a normas) a juzgar resultados contra estándares (evaluación basada en estándares) y, lo que es más importante, el cambio conceptual de pasar de un propósito simple de la evaluación, como es el de certificar a los estudiantes, a considerar múltiples propósitos, incluyendo ayudar al aprendizaje y al desarrollo de la capacidad de los estudiantes para emitir sus propios juicios.

Palabras clave: Evaluación; Reforma de la evaluación; Educación superior; Juicio evaluativo

Received/Recibido 2020 May 23 | Approved /Aprobado 2020 May 26 | Published/Publicado 2020 June 24

Todos tenemos la obligación, como parte del papel de la universidad en tanto institución pública, de declarar que nuestros graduados han cumplido con los requisitos académicos de sus titulaciones. Sin embargo, no podemos asumir que, por haberse hecho de una manera particular en el pasado, la evaluación sigue satisfaciendo las necesidades del presente. Debemos evaluar de manera que no socave el propio aprendizaje que estamos tratando de promover; debemos ser conscientes de que la expectativa sobre la evaluación tiene un profundo efecto sobre qué y cómo aprenden los estudiantes.

¿Qué trabajo específico debe hacer la evaluación? Es necesario considerar las diversas funciones importantes y legítimas que tiene. Está bien establecido que la evaluación y las actividades relacionadas con la evaluación tienen tres propósitos fundamentales (Boud, 2000; Boud y Soler, 2016):

- *Juzgar y certificar* lo que los estudiantes han aprendido. Es decir, validar que los estudiantes han cumplido con los requisitos del curso. A menudo, esto se denomina evaluación sumativa.
- *Ayudar a los estudiantes a aprender* para cumplir con los requisitos del curso. Esto ocurre a menudo mediante el uso de diversos procesos de retroalimentación y se denomina evaluación formativa
- *Desarrollar la capacidad de los estudiantes para valorar su propio trabajo* más allá del espacio temporal del curso. Esto implica el desarrollo progresivo de su propio criterio evaluativo y se denomina evaluación sostenible.

El primero de estos propósitos casi siempre se produce de alguna forma, pero es cuestionable su eficacia para lograrlo. Por ejemplo, ¿qué constituye el cumplimiento de los requisitos del curso? ¿Qué parámetros deben utilizarse para emitir un juicio? Estos propósitos se consideran cada vez más en términos de permitir y asegurar que los resultados de aprendizaje esperados de un curso se han alcanzado. Sin embargo, por lo general esto se hace de manera deficiente. La

forma en que se informa de la evaluación no permite normalmente establecer una conexión entre las calificaciones de los estudiantes y los resultados de aprendizaje declarados para un curso.

Desde el punto de vista individual del profesor dentro de una disciplina académica, los cambios en la evaluación a menudo parecen ser tan lentos que son casi imperceptibles. Sin embargo, si se consideran las universidades en su conjunto durante varios años, se han producido cambios importantes y se pueden anticipar muchos más. A medida que las universidades se ven más condicionadas por la influencia y la reglamentación externas, la presión de la globalización y las exigencias de los estudiantes, se observan nuevas tendencias en todo el mundo.

La evaluación tiene un legado de prácticas basadas en la disciplina para las que no existen bases bien articuladas y un conjunto de actividades basadas en reglas que suelen ser una reacción a problemas anteriores con los que se ha tropezado. En respuesta al aumento del número de estudiantes, se han introducido métodos de evaluación que, si bien supuestamente reducen el trabajo de calificación, como las pruebas de opción múltiple, también requieren una inversión considerablemente mayor en el diseño y el desarrollo, que no se tiene en cuenta. Se han encontrado cantidades excesivas de evaluación en algunas áreas y a menudo se tomaron decisiones administrativas arbitrarias para reducirlas sin pensar mucho en las consecuencias para el aprendizaje de los estudiantes. Si bien la evaluación formativa y la retroalimentación se han reconocido en principio a raíz del impulso de Black y Wiliam (1998), la mayoría de los cursos carecían de este aspecto y las presiones para reducir la evaluación tuvieron la consecuencia no deseada de reducir la práctica de los estudiantes y las oportunidades de retroalimentación. Grandes áreas del programa de estudios fueron infraevaluadas y otras más grandes fueron sobreevaluadas. Los estudiantes podían obtener sus titulaciones mediante la memorización estratégica, conservando al mismo tiempo las ideas

erróneas fundamentales de los conceptos clave. Estas prácticas de evaluación convencionales tenían características diferentes en los distintos países y en las distintas disciplinas, pero ninguna de ellas estaba exenta de censura. En los países en los que existían encuestas estándar para evaluar la experiencia de los estudiantes, las escalas de evaluación y retroalimentación eran típicamente las menos valoradas de cualquier aspecto de sus cursos.

Este documento adopta una visión centrada en el aprendizaje en el entorno actual. Aborda algunos de estos cambios y los desafíos que se han generado en el proceso. Su objetivo es proporcionar una visión general de los cambios que se han producido en la evaluación en la enseñanza superior y los diversos impulsores del cambio. Se identifican las esferas clave en las que se debe seguir centrando la atención en la reforma de la evaluación y se sugieren prioridades que beneficiarán a todos los propósitos de la evaluación. La primera parte del documento lo hace mediante una revisión de un documento de consenso creado hace 10 años como programa de reforma de la evaluación en las universidades que se centraba en lo que había que hacer para mejorar las prácticas de evaluación en las universidades a largo plazo. En él se ofrece la valoración del autor sobre lo que se ha logrado y lo que aún queda por hacer.

La segunda parte del documento se centra en lo que se necesita para los próximos diez años. Se basa en una mezcla de trabajo académico y observaciones políticas. Algunos de los desarrollos continúan la agenda existente, pero otros temas quedan por explorar. Es importante reconocer en este debate que los países, las instituciones y las disciplinas académicas no avanzan al mismo ritmo a lo largo de los años. En la mayoría de los países, es posible encontrar en cualquier institución o incluso en cualquier facultad, lo mejor y lo peor de la práctica de evaluación en estrecha proximidad. Por lo tanto, solamente podemos hablar en términos de tendencias generales, no de logros específicos y, de hecho, se necesita el plazo de un decenio para reconocer siquiera las tendencias. El documento es una reflexión

personal que no intenta repasar la considerable literatura, de figuras importantes como Royce Sadler y John Hattie, para influir. Se utiliza el estudio del autor para ilustrar el debate.

Parte 1. Revisión de lo que se ha logrado

Un punto de partida útil para el examen de los progresos realizados en la evaluación es una iniciativa de 2010 que dio lugar al documento: *Evaluación 2020: Siete propuestas para la reforma de la evaluación en la enseñanza superior*. Este fue el resultado de un proyecto nacional en el que participaron todas las universidades australianas. La iniciativa reconocía que era necesario cambiar la práctica de la evaluación en todo el sistema y que era preciso centrarse más en el aprendizaje a lo largo de la vida y en una evaluación que apoyara el plan de estudios. Se basaba en la idea de que:

Las universidades se enfrentan a cambios sustanciales en un contexto mundial que evoluciona rápidamente. Los retos que plantea el cumplimiento de las nuevas expectativas sobre los estándares académicos más allá del próximo decenio significan que será necesario repensar y renovar la evaluación (Boud & Associates, 2010, p. 1)

El proyecto trató de proporcionar un estímulo a los que participan en la reelaboración de la evaluación a todos los niveles. El proceso que se utilizó consistió en aprovechar los conocimientos especializados de un grupo de investigadores en evaluación muy experimentados, profesionales del desarrollo académico y administradores académicos para determinar las mejores ideas actuales sobre las formas en que la evaluación debe abordar las demandas inmediatas y futuras. Para obtener el compromiso de todos los interesados, el documento final fue redactado conjuntamente por representantes de casi todas las universidades australianas y un grupo de expertos en evaluación. Entre los representantes figuraban muchos de los niveles más altos de los responsables de las políticas de evaluación de cada institución (presidentes

de juntas académicas, rectores, presidentes de comités de evaluación de universidades).

El producto final adoptó la forma de un documento de consenso en el que se incluyó a los que participan en la elaboración y aplicación de políticas y en la investigación sobre la práctica de la evaluación en las universidades. Por consiguiente, las propuestas identificadas en él no eran radicales, sino que representaban compromisos que reflejaban tanto lo que se aceptaba como buena práctica como lo que podía adoptarse razonablemente en todo el sistema. Si bien el documento fue consultado con frecuencia y tuvo influencia en la formulación de la política de evaluación de algunas universidades, es difícil vincular directamente cualquier cambio concreto con propuestas específicas.

Los principios en los que se basó se enunciaron de la siguiente manera:

1. La evaluación enmarca poderosamente cómo los estudiantes aprenden y qué es lo que logran. Es una de las influencias más significativas en la experiencia de los estudiantes en la educación superior y todo lo que ganan con ella. La razón para mejorar la práctica de la evaluación es el enorme impacto que tiene en la calidad del aprendizaje.

2. La evaluación es la realización de juicios sobre cómo el trabajo de los estudiantes cumple con los estándares apropiados. Tradicionalmente, los maestros, los calificadores y los examinadores han sido los encargados de esa responsabilidad. Sin embargo, los propios estudiantes deben desarrollar la capacidad de juzgar tanto su propio trabajo como el de los demás con el fin de convertirse en estudiantes y practicantes eficaces.

3. La evaluación desempeña un papel fundamental tanto en el fomento del aprendizaje como en la certificación de los estudiantes. Sin embargo, a menos que primero satisfaga el propósito educativo de asegurar que los estudiantes puedan identificar trabajos de alta calidad y puedan relacionar estos conocimientos

con su propio trabajo, la probabilidad de que ellos mismos alcancen altos estándares se reduce. (p.1)

Las Propuestas

Las siete propuestas que se elaboraron tenían por objeto articular un programa en el que se esbozaban importantes orientaciones para la política y la práctica de la evaluación. Se denominaron *Evaluación 2020*, en el supuesto de que se necesitarían diez años para que se produjeran cambios significativos en los sistemas de evaluación. En la Tabla 1 figura el conjunto completo de propuestas, sin incluir su justificación.

Cabe señalar los siguientes puntos al respecto. En primer lugar, incorporan un enfoque en todos los propósitos de la evaluación: sumativo, formativo y sostenible. Sugieren que no es significativo centrarse exclusivamente en uno u otro de estos propósitos de la evaluación, todos deben ser considerados conjuntamente. Un estudio de las políticas de evaluación universitarias, publicado antes de las propuestas (Boud, 2007), había identificado que la gran mayoría de las políticas de evaluación universitarias se centran en lo sumativo, solo un poco en lo formativo y casi nada en lo sostenible. Esta era una preocupación del grupo que formuló las propuestas. En segundo lugar, sin que se afirmara de manera tan explícita, se suponía que la evaluación se basa ahora en estándares y actúa para valorar el rendimiento en función de los resultados de aprendizaje. Se excluye la evaluación normativa ya que para entonces se había desalentado o prohibido en la mayoría de las universidades australianas y esto se ha establecido en los estándares nacionales de calidad desde entonces. En tercer lugar, la atención se centra en todo momento en el aprendizaje de los estudiantes. Las pruebas que no tienen una influencia positiva en el aprendizaje de los estudiantes, es decir, la medición de resultados sin validez consecuente, no tienen cabida en la evaluación universitaria. Por último, la evaluación no es una práctica unilateral impuesta a los estudiantes. Para que la evaluación influya en el aprendizaje, los estudiantes deben participar

activamente en los procesos de evaluación y no limitarse a responder a las indicaciones no negociables de los demás.

Adopción de las propuestas

En los diez años desde que se formularon, ¿cuánto se ha avanzado en la evaluación universitaria? La respuesta a esta pregunta varía enormemente y solo puedo recurrir a mi

experiencia directa con las universidades australianas y con algunas en otros lugares. Unas pocas universidades han adoptado las propuestas de manera sustancial y han revisado sus políticas de evaluación a la luz de las mismas. Para otras es difícil determinar el alcance del cambio ya que también ha habido otros impulsores del cambio durante este período.

Tabla 1. Evaluación 2020: Siete propuestas para la reforma de la evaluación en la educación superior

LA EVALUACIÓN TIENE MÁS EFECTO CUANDO:	
1. ... la evaluación se utiliza para involucrar a los estudiantes en un aprendizaje que sea productivo.	<ul style="list-style-type: none"> i. ... la evaluación está diseñada para centrar a los estudiantes en el aprendizaje. ii. ...la evaluación se reconoce como una actividad de aprendizaje que requiere un compromiso con las tareas apropiadas.
2. ... la retroalimentación se usa para mejorar activamente el aprendizaje de los estudiantes.	<ul style="list-style-type: none"> i. ...la retroalimentación es informativa y de apoyo, y facilita una actitud positiva para el aprendizaje futuro. ii. ...los estudiantes buscan y usan la retroalimentación oportuna para mejorar la calidad de su aprendizaje y trabajo. iii. ...los estudiantes reciben regularmente información específica, no solo notas y calificaciones, sobre cómo mejorar la calidad de su trabajo.
3. ... los estudiantes y los profesores se convierten en socios responsables en el aprendizaje y la evaluación.	<ul style="list-style-type: none"> i. ... los estudiantes progresivamente asumen la responsabilidad de los procesos de evaluación y retroalimentación. ii. ... los estudiantes desarrollan y demuestran la capacidad de valorar la calidad de su propio trabajo y el de los demás según los estándares acordados. iii. ... el diálogo y la interacción sobre los procesos y estándares de evaluación son habituales entre el profesorado y los estudiantes.
4. ... los estudiantes son iniciados en las prácticas de evaluación y en la cultura de la educación superior.	<ul style="list-style-type: none"> i. ... las prácticas de evaluación están cuidadosamente estructuradas en las primeras etapas de los cursos para asegurar que los estudiantes hagan una transición exitosa a los estudios universitarios en su campo elegido. ii. ... las prácticas de evaluación responden a la diversidad de expectativas y experiencias de los estudiantes que ingresan.
5. ... la evaluación para el aprendizaje se coloca en el centro del diseño de la asignatura y del programa.	<ul style="list-style-type: none"> i. ... el diseño de la evaluación se reconoce como una parte integral de la planificación del plan de estudios desde las primeras etapas del desarrollo del curso. ii. ... la evaluación se organiza holísticamente a través de temas y programas con tareas integradas complementarias.
6. ... la evaluación para el aprendizaje es un foco para el desarrollo del personal y de las instituciones.	<ul style="list-style-type: none"> i. ... los enfoques profesionales y académicos de evaluación por parte del personal académico son desarrollados, desplegados, reconocidos y recompensados por las instituciones. ii. ... las prácticas de evaluación y el plan de estudios deberían revisarse a la luz de las percepciones de los graduados y los empleadores sobre la preparación de los graduados. iii. ... la evaluación de los logros de los estudiantes se juzga en función de estándares nacionales e internacionales coherentes que están sujetas a un diálogo, revisión y justificación continuos dentro de las comunidades disciplinarias y profesionales.
7. ... la evaluación proporciona una representación inclusiva y confiable de los logros de los estudiantes.	<ul style="list-style-type: none"> i. ... los resultados de las evaluaciones intermedias utilizadas para la retroalimentación sobre el aprendizaje y el progreso no juegan un papel significativo en la determinación de las calificaciones finales de los estudiantes. ii. ...la evidencia del logro alcanzado para determinar las calificaciones finales se basa en la evaluación del aprendizaje integrado. iii. ...la certificación retrata con precisión y riqueza los logros de los graduados y estudiantes para informar las futuras carreras y aprendizajes.

Las propuestas fueron más influyentes a nivel de diseño de la evaluación, tanto para los programas como para las unidades de curso¹ específicas. Se han producido cambios con respecto a las tres primeras propuestas, con el reconocimiento de que la evaluación puede influir positivamente en el aprendizaje y que los estudiantes deben participar activamente en él. Las actividades de evaluación entre iguales han proliferado particularmente en el área formativa. Las pruebas entre iguales (el uso de pares o compañeros para generar calificaciones formales) se critican cada vez más por no ser creíbles, pero los procesos de retroalimentación entre iguales están resultando valiosos y se están adoptando ampliamente.

La evaluación no se ve como una medida o juicio que deja al "sujeto" (el estudiante) sin afectar, sino como un acto que influye profundamente, de muchas maneras, en el aprendizaje de los estudiantes y sus prioridades. La mala elección de las actividades de evaluación conduce a un aprendizaje deficiente y distorsiona lo que los estudiantes terminan siendo capaces de hacer. Todavía hay un desafortunado exceso de énfasis en la evaluación de los conocimientos de bajo nivel y una falta de énfasis en la evaluación de la aplicación de esos conocimientos. Los exámenes con demasiada frecuencia incitan a los estudiantes a memorizar en lugar de comprender. Hay una conciencia mucho mayor de que esto es un problema ahora, pero se siguen dando excusas de que las preguntas de opción múltiple o de respuesta corta deben dominar la evaluación debido al gran tamaño de las clases. Lo que se olvida en este proceso es que, si bien son fáciles de calificar, es extremadamente difícil diseñar pruebas de opción múltiple de alta calidad. Por lo tanto, vemos una proliferación de pruebas mal construidas que comunican el mensaje a los estudiantes de que lo importante es aprender los conocimientos de bajo nivel necesarios para aprobarlas. Esto fomenta los malos hábitos de estudio entre los estudiantes.

Las ilustraciones de los sistemas de evaluación en cursos con un gran número de estudiantes (más de 1000) muestran que se puede hacer mucho para mejorar la calidad del aprendizaje mediante el diseño de evaluaciones dentro de los niveles de recursos existentes: véanse los estudios de casos presentados en www.feedbackforlearning.org.

En relación con la cuarta propuesta, uno de los mayores éxitos ha sido el aumento del interés en lo que se ha denominado *pedagogía de la transición* para ayudar a los estudiantes a introducirse en la cultura y las prácticas del aprendizaje de la educación superior (Kift et al., 2010; www.transitionpedagogy.com). Se considera que las clases de primer año no solo introducen a los estudiantes en las disciplinas que estudiarán, sino también en los procesos de aprendizaje y de aprender a aprender que necesitarán a lo largo de sus cursos. La inversión en la calidad de las asignaturas de primer año y en la evaluación del primer año no solo tiene un beneficio a largo plazo para el resto del curso, sino que permite retener a los estudiantes que de otro modo fracasarían.

El diseño de la evaluación para el aprendizaje, el tema de la quinta propuesta, se ha tomado más en serio (Bearman et al, 2016). Esto ha sido impulsado por dos cambios. En primer lugar, la atención prestada a los estándares y los resultados de aprendizaje, y la cartografía de los planes de estudio que se asocia a ella, ha hecho que sea más fácil advertir que las evaluaciones no abordan los resultados de aprendizaje según lo especificado (véase www.assuringlearning.com). Las instituciones se han avergonzado de que esta transparencia haya revelado que no están haciendo lo que afirman hacer en cuanto a los resultados obtenidos por sus graduados reales en el momento de la graduación. En segundo lugar, el deseo de que los programas sean acreditados por organismos internacionales externos (e.g. la AACSB) ha centrado la atención de las escuelas de negocios en particular en el diseño de su evaluación y en la necesidad de alinearla

¹ El término unidad de curso o unidad se utiliza en todo momento para referirse a las asignaturas o módulos individuales de un programa de grado.

con resultados de aprendizaje explícitos. El diseño de la evaluación era un término utilizado solo por los profesionales de la educación hace una década, pero está adquiriendo un uso más general. El tema recibió un impulso gracias al Proyecto de Decisiones sobre el Diseño de Evaluaciones, que desarrolló un sitio web con recursos para ayudar a los académicos en el diseño de las evaluaciones: www.assessmentdecisions.org.

La evaluación como foco para el desarrollo institucional y del personal (cuerpo docente), la sexta propuesta, se ha convertido en un diferenciador de las instituciones. Si bien se ha hecho más hincapié en el desarrollo en general, ha aumentado la brecha entre las instituciones que se toman en serio el desarrollo educativo y que invierten en él, a nivel central y local y las que han tomado medidas simbólicas. Ha aumentado el número de cursos para académicos que incluyen un componente sustancial sobre evaluación y se han multiplicado las iniciativas institucionales para reformar los cursos a fin de utilizar la evaluación de manera más racional. En algunas disciplinas, en particular la medicina, se está incorporando la noción de evaluación programática. Este enfoque adopta una visión de la evaluación a nivel de todo el programa, separa las actividades de evaluación de las unidades individuales del curso y garantiza que se han alcanzado todos los resultados necesarios para el curso en su conjunto (Van der Vleuten et al, 2012).

Probablemente el menor movimiento se ha producido en el área de la propuesta final. Esto no es sorprendente ya que se centró en el área más intratable para el cambio: las representaciones sumativas de los resultados de la evaluación. Aunque algunas universidades de algunos países (p. ej., el Reino Unido) hacen caso omiso de las calificaciones otorgadas en el primer año a los efectos de determinar los promedios de las calificaciones o el estado de las distinciones, sigue habiendo un enfoque poco saludable e inválido en las calificaciones. Las calificaciones se otorgan sin asegurar los resultados de aprendizaje, el desempeño inadecuado en un área se compensa con un buen desempeño en otra y las bajas

calificaciones en las primeras partes de un curso siguen siendo una desventaja para los estudiantes, incluso cuando sobresalen exactamente en la misma área más adelante. Lo más importante es que el expediente académico que se proporciona a los graduados al final de un programa no comunica a las partes externas casi ninguna información útil sobre lo que un estudiante sabe y puede hacer.

No debemos sobreestimar la magnitud de los efectos en ninguna de las áreas de las propuestas. Hay un camino considerable por recorrer, como también lo ejemplifican los estudios en el contexto europeo (Panadero et al, 2019; Pereira et al., 2017; Rodríguez-Gómez et al., 2016; Quesada-Serra et al., 2016), particularmente en aquellas instituciones que no se han dado cuenta suficientemente de que su futuro depende en gran medida de la calidad de sus cursos y no solo de su investigación.

Otros cambios importantes en la evaluación

Los cambios en las áreas identificadas en las propuestas deben considerarse en el contexto de los grandes cambios que tienen lugar en el mismo período de tiempo. Dos son dignos de mención. Probablemente el mayor cambio en las políticas de evaluación de la enseñanza superior haya sido hacer que se basen menos en normas y más en principios. Las antiguas políticas eran a menudo acumulaciones de normas concebidas para abordar problemas que se habían producido a lo largo de los años. A menudo se centraban en la realización de exámenes, no en la evaluación en general. Un enfoque basado en principios significa que hay más flexibilidad en los procedimientos de evaluación específicos (la evaluación no tiene por qué ser la misma en toda la institución o incluso en toda una facultad) pero todos los procesos de evaluación deben ajustarse a los principios articulados en la política. Esto, por ejemplo, puede significar que se eliminen las normas que exigen que ciertas proporciones de las notas se basen en los exámenes y que se abandonen las normas sobre el número de tareas de evaluación por unidad del curso. Irónicamente, incluso cuando esto ha ocurrido, muchos académicos todavía creen que las reglas siguen existiendo. La evolución hacia un

enfoque de la política de evaluación basado en principios no fue impulsada por las propuestas, sino por los cambios en la gestión de la universidad, a menudo para proteger a la institución de litigios no deseados por parte de los estudiantes.

El segundo cambio importante en la evaluación durante el último decenio ha sido el aumento de la evaluación auténtica. La evaluación auténtica implica el uso de tareas y procesos de evaluación que representan los tipos de tareas y procesos que se encuentran en el mundo de la práctica. Muchas disciplinas profesionales ya habían dado este paso durante el siglo pasado, pero, incluso en ese caso, las tareas en materias básicas o asignaturas académicas, como parte de un título profesional, a menudo no eran auténticas. Esto ha cambiado mucho y hay indicios de que se ha adoptado la evaluación auténtica en cursos que no conducen directamente a una profesión. Aunque es mucho más probable que las tareas sean auténticas, las pruebas y exámenes han sido más resistentes a esta tendencia. Además, el cambio a la evaluación auténtica varía sustancialmente entre los países. Los países en los que tradicionalmente se ha hecho gran hincapié en las pruebas y los exámenes muestran menos signos de pasar a la evaluación auténtica, pero hay indicios, incluso en países como Chile, que hacen (demasiado) hincapié en las pruebas de opción múltiple, de que se está considerando la evaluación auténtica (Villarreal et al, 2019). En países como Australia, la primera oleada de evaluación auténtica en todos los cursos ha pasado y las instituciones se enfrentan al reto de tratar con académicos que se resisten a ella.

Por supuesto, ha habido cambios en el campo más amplio de la evaluación, pero las iniciativas en torno a la evaluación del sector educativo han ejercido muy poca influencia en la enseñanza superior de la mayoría de los países. En el contexto australiano, los cambios en la evaluación de las universidades se vieron poco influidos por éstas y, solo de manera limitada, por los especialistas en evaluación de renombre internacional que trabajan en el país. Esto es común en otros lugares.

Parte 2. ¿Adónde vamos desde aquí?

A pesar de algunos cambios importantes, en la mayoría de las instituciones todavía hay cuestiones considerables que deben abordarse en la evaluación. Algunos de ellos requieren una mayor investigación antes de que podamos ofrecer recomendaciones seguras para el cambio, pero muchos otros no. Un extraño se sorprendería al ver cuánta práctica que no puede defenderse sobre la base de ningún conocimiento evaluativo sigue existiendo. Si bien algunos de los cambios que deben hacerse pueden ser discutidos, muchas otras prácticas se han fijado simplemente por costumbre e inercia. Las instituciones deben estar dispuestas a aceptar la necesidad de abordar algunas de estas cuestiones pendientes y se requiere cierta sensibilidad política interna para lograr que se apliquen (p. ej., Deneen y Boud, 2014). La segunda parte se centra, en primer lugar, en las cuestiones no resueltas que probablemente no requieran más investigación académica y, en segundo lugar, considera las que requieren más investigación y desarrollo antes de que puedan aplicarse plenamente.

Completar el movimiento hacia un marco de evaluación basado en estándares

La mayoría de los sistemas de enseñanza superior de los países de la OCDE han adoptado un marco de cualificaciones basado en estándares, al menos al nivel más alto. Esto implica la especificación de los resultados de aprendizaje para todos los programas y el uso de la evaluación para asegurar que estos resultados se han cumplido. Este marco a menudo no se traduce plenamente en las especificaciones de los cursos ofrecidos. Sin embargo, en algunos países, el enfoque basado en estándares y orientado a los resultados se ha incorporado en la legislación utilizada para controlar las universidades (p. ej., en Australia), y en otros se ha aceptado sin ninguna imposición. Es común, incluso en países como Australia, donde la expectativa es muy fuerte y las instituciones aceptan el cambio, que las políticas todavía incorporen supuestos sobre las calificaciones y la evaluación referida a norma que son incompatibles con un enfoque basado en

estándares. Aún más común es que las facultades y los cursos queden muy a la zaga de los cambios de política institucional que se han producido, no principalmente porque los cambios sean discutidos, sino porque no hay una razón directa para cambiar.

¿Qué se necesita para que la evaluación funcione en un marco basado en estándares? Todo gira en torno a tener declaraciones significativas de los resultados de aprendizaje a un nivel de abstracción adecuado. Si son muchos, tienden a convertirse en excesivamente conductuales; si son pocos, no abarcan la gama de resultados necesarios. Una de las principales trampas es hacer demasiado hincapié en las características técnicas y de conocimiento del aprendizaje a expensas de la utilización del conocimiento en términos de lo que los estudiantes pueden hacer. Un resultado efectivo de aprendizaje debería representar claramente lo que un estudiante puede hacer como resultado de un determinado curso o programa de estudio. Se requiere de un estándar umbral para definir lo que todos los estudiantes deben ser capaces de demostrar y niveles adicionales pueden indicar logros más allá de esto. No se debe permitir que los estudiantes que no cumplan con el estándar de umbral pasen, aunque les vaya bien en otras partes del curso. En el resultado de aprendizaje hay, por supuesto, un cuerpo considerable de conocimientos, pero no es necesario que el conocimiento detallado esté en primer plano, sino solo lo que le permitirá hacer al estudiante. Si los resultados de aprendizaje están bien redactados, y a menudo requerirán al menos tres iteraciones importantes antes de estar suficientemente bien elaborados, entonces las actividades de evaluación pueden ser impulsadas por ellos.

Hay dos niveles clave de resultados de aprendizaje: para el programa en su conjunto y para las unidades, módulos o materias individuales dentro de él. Tradicionalmente, la atención se ha centrado en la unidad o materia individual, pero cada vez más debe centrarse en los resultados generales del programa. Independientemente de lo que se evalúe, son estos resultados generales de aprendizaje los que deben ser prioritarios ya que son los que representan la cualificación en la que está

inscrito un estudiante; son éstos los que deben garantizarse mediante la evaluación sumativa. El dilema que hay que afrontar es si las evaluaciones dentro de las unidades de un curso específico permitirán asegurar estos resultados. Esto solo puede ocurrir si existe una correspondencia clara entre los resultados de aprendizaje de las unidades y los resultados de aprendizaje de los cursos y un acuerdo de que los resultados de aprendizaje de los cursos/programas siempre tienen prioridad.

Una dificultad importante para la evaluación son los sistemas de registro de evaluación heredados que utiliza cada universidad, que no se han actualizado desde una época anterior. Normalmente, permiten registrar las notas (letras o números) de cada estudiante en cada tarea. Para cada unidad se determina un promedio ponderado y los promedios sobre las unidades se calculan como un promedio de calificaciones. Estas diversas formas de promedios sobre diferentes resultados de aprendizaje son incompatibles con un enfoque basado en estándares. En ese enfoque, la evaluación debe garantizar que cada resultado de aprendizaje se ha alcanzado en el nivel de umbral. El exceso de rendimiento en uno de ellos no puede compensar el déficit de rendimiento en otro. El promedio de los diferentes resultados del aprendizaje carece de sentido, ya que representan, o deberían representar, logros fundamentalmente diferentes (Boud, 2017). Por lo tanto, lo que tenemos es un sistema disfuncional que socava los principios en los que se basa actualmente la evaluación. El cambio solo se producirá cuando sea posible que los sistemas de registro de la evaluación consignen los niveles de logro por resultados de aprendizaje, sin acumulación entre los diferentes resultados. Esto significa que, si una unidad de curso o una asignatura tiene tres resultados de aprendizaje, entonces es necesario registrar tres calificaciones o notas, puesto que sería inapropiado hacer un promedio entre ellos.

Incluso una vez que se adecuen los sistemas de registro, sigue existiendo el problema de asegurar los resultados de aprendizaje. El hecho de tener varias calificaciones para un solo resultado no resuelve este problema. Lo que hay que identificar es qué cuenta como

haber alcanzado el resultado. Además, ¿en qué etapa se puede decir que se ha alcanzado un resultado de aprendizaje de un curso (como distinto de una unidad)? Si queremos mantener la práctica de que la evaluación se realice dentro de las unidades del curso (en lugar de a nivel de curso o programa), entonces hay que tomar decisiones sobre cuál es el evento de evaluación que asegura cada una de ellas. Es común ayudar a este proceso definiendo cada evento de evaluación para cada unidad del curso en términos de si representa un nivel de logro introductorio, un nivel intermedio o un nivel final asegurado. Estos pueden ser mapeados a través de las unidades para asegurar que todos los resultados han sido abordados y que algunos no son evaluados en exceso. (Véase, www.assuringlearning.com). El problema del aseguramiento o garantía plantea la cuestión de si las evaluaciones introductorias o intermedias deben "contar" a efectos de la graduación. Es necesario realizarlas como parte del proceso de aprendizaje y puede ser necesario registrarlas, pero ¿cuál es la justificación para que aparezcan en un documento público? Seguramente, todo lo que se necesita es un registro validado de lo que los estudiantes son capaces de hacer al final.

En el próximo decenio de desarrollo será necesario hacer mucho más hincapié en la noción de evaluación como aseguramiento o garantía. Hay una serie de iniciativas que contribuyen a ello, además de la cartografía de los programas de estudio a que se hace referencia en el último párrafo. La primera es la separación de la enseñanza y el aprendizaje de la evaluación. Una iniciativa de la Universidad de Brunel en Londres, por ejemplo, permite que la evaluación solo se produzca en "unidades de evaluación" en las que no hay enseñanza (<https://bilt.online/wp-content/uploads/2018/11/IPA-Practical-Guide-0918.pdf>). La enseñanza se produce en "unidades de enseñanza" en las que solo ocurre la evaluación formativa. Las "unidades de evaluación" suelen abarcar múltiples "unidades de enseñanza", de modo que se reduce la carga global de la evaluación y esta se centra en los resultados de aprendizaje generales del curso. Una segunda iniciativa,

mencionada anteriormente, es la noción de evaluación programática. Toma la idea de separar la evaluación sumativa de la enseñanza y considera toda la evaluación en función de la medida en que ésta valora el rendimiento de los estudiantes en relación con los resultados de aprendizaje del curso (Van der Vleuten et al, 2012; www.bradford.ac.uk/pass/).

No debemos aceptar sin críticas el enfoque basado en estándares y criterios y la centralidad de los resultados de aprendizaje. Estos han sido un paso importante en la reforma de la evaluación, pero no son una panacea. Mientras se producen las críticas (p. ej., Ajjawi et al., 2019), tenemos que ser conscientes de no adoptar una postura reaccionaria que nos lleve a un pasado idealizado que no existía. El paso a los estándares y los resultados del aprendizaje fue necesario para hacer frente a la enorme expansión de la educación superior y la rendición de cuentas del dinero público. La disciplina de centrarse en los resultados de aprendizaje es importante, pero debemos ser conscientes de que algunas de las rigideces administrativas que se han producido junto con esta medida pueden ser contraproducentes.

Reposicionando a los estudiantes como aprendices y productores

Probablemente el mayor desafío *educativo* de la educación superior es el de cambiar la identidad de los estudiantes de la de "estudiante" a la de "aprendiz". Ser un estudiante a menudo significa hacer lo que se requiere (asistir a conferencias, hacer exámenes, etc.) en otras palabras, seguir las instrucciones de los demás para graduarse. El papel pasivo se refuerza con las prácticas de la universidad. Los estudiantes son tratados como una masa, las decisiones se toman unilateralmente sobre lo que deben aprender, lo que deben hacer y cómo son evaluados. Esta es una posición profundamente inútil que los estudiantes no han de adoptar si quieren convertirse en graduados responsables en el mundo, con control sobre su propia dirección y su propio aprendizaje. La identidad de "estudiante" no los equipa para ser un practicante efectivo de cualquier tipo en la sociedad actual. Ser un "aprendiz" es ser

proactivo, tomar la iniciativa para decidir en qué concentrarse y juzgar el propio desempeño en el proceso. Para los aprendices, otras personas pueden ser recursos que se pueden utilizar y de los que se puede aprender, no directores de sus acciones.

Aunque se necesitan muchas cosas en el plan de estudios y la pedagogía para abordar el posicionamiento pasivo de los estudiantes, nos centramos aquí en la evaluación. Hay tres direcciones para explorar e investigar más y luego implementar el cambio. La primera es con respecto al desarrollo del juicio evaluativo de los estudiantes, la segunda y relacionada con ella, es la forma en que se conceptualiza y se lleva a cabo la retroalimentación. En tercer lugar, está la de las nuevas formas de representación de los logros.

1. Desarrollando el juicio evaluativo de los estudiantes

No basta con que un graduado tenga ciertos conocimientos y pueda hacer ciertas cosas. Por supuesto, cuando se relacionan con los resultados de aprendizaje del curso, deben estar consolidados. Sin embargo, se necesita más para que un graduado sea una persona efectiva en el mundo. Necesitan saber lo que saben y lo que pueden hacer y saber cuáles son los límites de esto. Todas las formas de trabajo requieren que los profesionales aprovechen sus conocimientos y capacidades. El graduado necesita determinar cómo lo que sabe y puede hacer se relaciona con el problema en cuestión y qué más necesita saber y ser capaz de hacer para abordarlo. Cuando los empleadores critican a las universidades por no proporcionar graduados útiles, tienden a no preocuparse por sus conocimientos y capacidades, sino por su capacidad para utilizarlos eficazmente en situaciones nuevas. Por lo tanto, una necesidad fundamental de la educación superior (y, de hecho, de otros sectores también) es que los estudiantes desarrollen la capacidad de analizar su propio aprendizaje y el de los demás, por ejemplo, "¿cómo sabré si esto es lo suficientemente bueno?", "¿estoy cumpliendo con el estándar para esta tarea en particular?" Esta capacidad se ha denominado juicio evaluativo, que se ha definido como "la capacidad de tomar

decisiones sobre la calidad del trabajo propio y ajeno" (Tai et al, 2018). Podría decirse que es más importante que cualquier resultado de aprendizaje concreto que anuncie un curso. Es necesario tanto para el aprendizaje dentro del curso como a lo largo de la vida.

Es probable que la elaboración de juicios evaluativos sea específica para cada ámbito y deberá promoverse en todo el programa de estudios. Requiere una práctica considerable e implica dos elementos clave: la apreciación de lo que constituye un buen trabajo y la formulación de juicios sobre el trabajo propio y el de los demás (Boud et al., 2018; Joughin et al., 2019). La primera precede a la segunda: sin saber cómo es una buena obra, ¿cómo se puede producir? Si bien el desarrollo del juicio evaluativo se ha basado en la literatura de la autoevaluación de los estudiantes y la evaluación entre iguales, no se trata simplemente de un replanteamiento de eso: es un objetivo del plan de estudios. No importa cuán pobre sea un estudiante a la hora de juzgar su propio trabajo o el de los demás, el objetivo es mejorar para que los estudiantes dependan cada vez menos del juicio de los profesores a medida que avanza el curso.

Si bien actualmente podemos establecer el desarrollo del juicio evaluativo como parte del programa de estudios, hay muchas preguntas que deben responderse antes de que pueda aplicarse plenamente. Entre ellas figuran: ¿cuáles son las formas eficaces de incorporarlo en los cursos que mejoran los resultados de aprendizaje? ¿Cuánta práctica sobre cuántas repeticiones, sobre cuántas unidades de curso, sobre cuántos años se necesita para su desarrollo? ¿Cuál es el equilibrio adecuado entre los juicios de los profesores, los juicios propios y los juicios de los compañeros? ¿Requiere estrategias de aplicación bastante diferentes para diferentes resultados y diferentes disciplinas? Parte de esta investigación ha comenzado en el contexto español con el estudio de Ibarra-Sáiz et al. (2020).

2. Reconceptualizando y reencuadrando la retroalimentación

La retroalimentación es una característica importante del aprendizaje en general y del desarrollo del juicio evaluativo en particular. En los últimos años se ha producido un enorme cambio en la forma de conceptualizar la retroalimentación en la enseñanza superior, que ha pasado de la noción centrada en la retroalimentación del profesor, como un aporte a los estudiantes (comentarios sobre su trabajo), a la retroalimentación como un proceso en el que los estudiantes deben participar necesariamente (Boud y Molly, 2013, Carless, 2015). Este cambio ha sido tan grande que ha sido calificado como un cambio de paradigma por Winstone y Carless (2019). Las implicaciones de este cambio son profundas ya que influyen en cada parte de un curso en el que se despliega la retroalimentación. También socava el uso del término "*feed-forward*" utilizado en algunas publicaciones populares, ya que es superfluo y revela un malentendido sobre el significado mismo de la retroalimentación.

La retroalimentación está siendo reformulada como una idea centrada en el aprendizaje. Carless y Boud (2018) lo definen como "*un proceso a través del cual los alumnos dan sentido a la información de diversas fuentes y la utilizan para mejorar su trabajo o sus estrategias de aprendizaje*" (p. 1316). La implicación de esto es que la retroalimentación por defecto no es el suministro de información no solicitada de los profesores a los estudiantes sobre su trabajo evaluado, como complemento de las calificaciones. Se trata más bien de un proceso que pueden iniciar los estudiantes para obtener información de otros, no solo de los profesores, que luego pueden procesar para mejorar su trabajo. Los indicadores típicos de una visión centrada en el aprendizaje son: los estudiantes proponen a otros qué tipo de comentarios desean sobre su trabajo, se espera que los estudiantes respondan a los comentarios con un plan de lo que harán, las pruebas de una retroalimentación efectiva se encuentran en el trabajo posterior de los estudiantes. La retroalimentación desde esta perspectiva es un tipo de diálogo en el que los estudiantes necesariamente desempeñan un

papel activo. Ejemplos de este nuevo enfoque de la retroalimentación pueden encontrarse en la nueva colección de Henderson et al. (2019) y las sugerencias sobre la forma en que la instrucción por retroalimentación de los estudiantes puede incorporarse al plan de estudios se presentan en Malecka et al. (2020).

3. Nueva representación de los logros

El área en la que los estudiantes se han posicionado de forma más pasiva es la de la evaluación final o sumativa. La educación superior no solo tiene que ayudar a los estudiantes a aprender más eficazmente, sino que también tiene la responsabilidad, no solo de certificar lo que los estudiantes han logrado, sino de permitirles representar bien sus logros. Sin embargo, tiene que hacerlo sin inhibir los propios procesos de acción independiente que trata de fomentar. Las credenciales digitales requieren que se lleven registros de lo que puede hacer un estudiante exactamente, en qué condiciones, cómo se ha juzgado ese logro, según qué estándares y criterios, y por quién. Esta credencial puede insertarse, por ejemplo, en un perfil de LinkedIn y ser validada oficialmente por cualquier persona de la universidad que acceda al perfil con solo pulsar un botón. Esas credenciales pueden extenderse no solo para representar todos los resultados de aprendizaje alcanzados por los estudiantes, sino que pueden ser utilizadas por los estudiantes para diferenciarse de los demás mediante la exhibición de sus logros singulares (Jorre St Jorre et al., 2019).

Una vez establecido un depósito digital con información sobre lo que ha hecho un estudiante, ésta puede ser organizada por los estudiantes para permitir diversas formas de representación con diferentes propósitos: destacar logros particulares a un posible empleador o representarse a sí mismo en su campo profesional. La organización consiste en combinar los logros validados oficialmente con otros mecanismos, incluidos los logros extra curriculares, con un comentario del graduado que los sitúe en el contexto (Clarke y Boud, 2018). Actualmente, existen todos los elementos básicos para esas representaciones de la evaluación, pero se necesita mucha investigación y desarrollo para producir

tecnologías suficientemente facilitadoras y, lo que es más importante, para facilitar a los estudiantes la búsqueda de formas de registro y presentación a la luz de los avances producidos en el contexto de la comunicación y las redes sociales (Ajjawi et al., 2020).

Las tres cuestiones examinadas anteriormente han demostrado ser muy prometedoras y bien podrían transformar la evaluación en el futuro. Cada una de ellas requiere también una considerable investigación adicional ya que hay muchas características de ellas que están subdesarrolladas o inexploradas.

Conclusión

Si bien el cambio en la evaluación parece difícil de alcanzar año tras año, una perspectiva más amplia muestra que son muchas las transformaciones que se están produciendo. El reto es fomentar la adopción de los enfoques identificados anteriormente. El principal requisito para que esto ocurra es el compromiso, especialmente el compromiso de las instituciones cuyos dirigentes ya son conscientes de que su sostenibilidad implica atraer y satisfacer las cambiantes necesidades de los estudiantes. Si esto no se produce con la suficiente rapidez, algunas de esas instituciones estarán en peligro.

Referencias

- Ajjawi, R., Bearman, M., & Boud, D. (2019). Performing standards: A critical perspective on the contemporary use of standards in assessment, *Teaching in Higher Education*. <https://doi.org/10.1080/13562517.2019.1678579>
- Ajjawi, R., Boud, D., & Marshall, D. (2020). Repositioning assessment-as-portrayal: what can we learn from celebrity and persona studies? In Bearman, M., Dawson, P., Ajjawi, R., Tai, J. & Boud, D. (Eds.) *Re-imagining university assessment in a digital world*. Springer.
- Bearman, M., Dawson, P., Boud, D., Bennett, S., Hall, M., & Molloy, E. (2016) Support for assessment practice: developing the Assessment Design Decisions Framework. *Teaching in Higher Education*, 21(5), 545-556. <https://doi.org/10.1080/13562517.2016.1160217>
- Bearman, M., Dawson, P., Bennett, S., Hall, M., Molloy, E. Boud, D., & Joughin, G. (2017). How university teachers design assessments: a cross disciplinary study. *Higher Education*, 74 (1), 49-64. <https://doi.org/10.1007/s10734-016-0027-7>
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74. <https://doi.org/10.1080/0969595980050102>
- Boud, D. (2000). Sustainable assessment: rethinking assessment for the learning society. *Studies in Continuing Education*, 22(2), 151-167. <https://doi.org/10.1080/713695728>
- Boud, D. (2007). Reframing assessment as if learning was important. In Boud, D. & Falchikov, N. (Eds.) *Rethinking Assessment for Higher Education: Learning for the Longer Term*. Routledge, 14-25. <https://doi.org/10.4324/9780203964309>
- Boud, D. (2014). Shifting views of assessment: from secret teacher's business to sustaining learning. In Kreber, C., Anderson, C., Entwistle, N. and McArthur, J. (Eds) *Advances and Innovations in University Assessment and Feedback* (pp. 13-31). Edinburgh University Press. <https://doi.org/10.3366/edinburgh/9780748694549.003.0002>
- Boud, D. (2017). Standards-based assessment for an era of increasing transparency. In Carless, D., Bridges, S., Chan, C. and Glofcheski, R. (Eds). *Scaling up assessment for learning in higher education* (pp. 19-31). Springer. https://doi.org/10.1007/978-981-10-3045-1_2
- Boud, D., & Associates (2010). *Assessment 2020: Seven propositions for assessment reform in higher education*. Australian Learning and Teaching Council.
- Boud D., Ajjawi R., Dawson P., & Tai J. (Eds.). (2018). *Developing Evaluative Judgement in Higher Education: Assessment for Knowing and Producing Quality Work*. Routledge. <https://doi.org/10.4324/9781315109251>

- Boud, D., & Molloy, E. (2013). Rethinking models of feedback for learning: the challenge of design. *Assessment and Evaluation in Higher Education*, 38(6), 698-712. <https://doi.org/10.1080/02602938.2012.691462>
- Boud, D., & Soler, R. (2016). Sustainable assessment revisited. *Assessment and Evaluation in Higher Education*, 41(3), 400-413. <https://doi.org/10.1080/02602938.2015.1018133>
- Carless, D. (2015). *Excellence in university assessment: Learning from award winning practice*. Routledge. <https://doi.org/10.4324/9781315740621>
- Carless, D., & Boud, D. (2018). The development of student feedback literacy: enabling uptake of feedback. *Assessment and Evaluation in Higher Education*, 43(8), 1315-1325. <https://doi.org/10.1080/02602938.2018.1463354>
- Clarke, J. L., & Boud, D. (2018). Refocusing portfolio assessment: curating for feedback and portrayal. *Innovations in Education and Teaching International*, 55(4), 479-486. <https://doi.org/10.1080/14703297.2016.1250664>
- Deneen, C., & Boud, D. (2014). Patterns of resistance in managing assessment change. *Assessment and Evaluation in Higher Education*, 39(5), 577-591. <https://doi.org/10.1080/02602938.2013.859654>
- Henderson, M., Ajjawi, R., Boud, D., & Molloy, E. (Eds.) (2019). *The Impact of Feedback in Higher Education: Improving assessment outcomes for learners*. Palgrave Macmillan <https://doi.org/10.1007/978-3-030-25112-3>
- Ibarra-Sáiz, M. S., Rodríguez-Gómez, G., & Boud, D. (2020). Developing student competence through peer assessment: the role of feedback, self-regulation and evaluative judgement. *Higher Education*. <https://doi.org/10.1007/s10734-019-00469-2>
- Jorre St Jorre, T., Boud, D., & Johnson, E. (2019) Assessment for distinctiveness: recognising diversity of accomplishments, *Studies in Higher Education*, <https://doi.org/10.1080/03075079.2019.1689385>
- Joughin, G., Boud, D., & Dawson, P. (2019). Threats to student evaluative judgement and their management, *Higher Education Research and Development*, 38(3), 537-549. <https://doi.org/10.1080/07294360.2018.1544227>
- Kift, S., Nelson, K., & Clarke, J. (2010). Transition pedagogy: A third generation approach to FYE - A case study of policy and practice for the higher education sector. *The International Journal of the First Year in Higher Education*, 1(1), 1-20. <https://doi.org/10.5204/intjfyhe.v1i1.13>
- Malecka, B., Boud, D., & Carless, D. (2020). Eliciting, processing and enacting feedback: Mechanisms for embedding student feedback literacy within the curriculum. *Teaching in Higher Education*. <https://doi.org/10.1080/13562517.2020.1754784>
- Panadero, E., Fraile, J., Fernández Ruiz, J., Castilla-Estévez, D., & Ruiz, M. A. (2019). Spanish university assessment practices: examination tradition with diversity by faculty. *Assessment & Evaluation in Higher Education*, 44 (3), 379-397. <https://doi.org/10.1080/02602938.2018.1512553>
- Pereira, D., Niklasson, L., & Flores, M. A. (2017). Students' perceptions of assessment: a comparative analysis between Portugal and Sweden. *Higher Education*, 73(1), 153-173. <https://doi.org/10.1007/s10734-016-0005-0>
- Rodríguez-Gómez, G., Quesada-Serra, V., & Ibarra-Sáiz, M. S. (2016). Learning-oriented e-assessment: the effects of a training and guidance programme on lecturers' perceptions. *Assessment & Evaluation in Higher Education*, 41(1), 35-52. <https://doi.org/10.1080/02602938.2014.979132>
- Quesada-Serra, V., Rodríguez-Gómez, G., & Ibarra-Sáiz, M. S. (2016). What are we missing? Spanish lecturers' perceptions of their assessment practices. *Innovations in*

Education and Teaching International, 53(1), 48-59.

<https://doi.org/10.1080/14703297.2014.930353>

Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgement: enabling students to make decisions about the quality of work. *Higher Education*, 76, 467-481. <https://doi.org/10.1007/s10734-017-0220-3>

Van der Vleuten, C. P. M., Schuwirth, L. W. T., Driessen, E. W., Dijkstra, J., Tigelaar, D., Baartman, L. K. J., & van Tartwijk, J. (2012). A model for programmatic assessment fit for purpose. *Medical Teacher*, 34, 205-214.

<https://doi.org/10.3109/0142159X.2012.652239>

Villarroel, V., Boud, D., Bloxham, S., Bruna, D., & Bruna, C. (2020). Using principles of authentic assessment to redesign written examinations and tests. *Innovations in Education and Teaching International*, 34, 535-557.

<https://doi.org/10.1080/14703297.2018.1564882>

Winstone, N., & Carless, D. (2019). *Designing Effective Feedback Processes in Higher Education: A Learning-Focused Approach*. Routledge.

<https://doi.org/10.4324/9781351115940>

Author / Autor

Boud, D. (david.boud@deakin.edu.au)  0000-0002-6883-2722

Profesor y director del Centro para la Investigación en Evaluación y Aprendizaje Digital en la Universidad de Deakin (Melbourne) y Profesor emérito en la Universidad de Tecnología de Sydney. También es Profesor de Trabajo y Aprendizaje en la Universidad de Middlesex (Londres). Ha escrito numerosas publicaciones sobre enseñanza, aprendizaje y evaluación en educación superior y educación profesional. Actualmente su trabajo se centra en las áreas de evaluación para el aprendizaje en educación superior, formación académica y aprendizaje en el lugar de trabajo. Es uno de los autores más citados en el ámbito de la evaluación superior. Ha sido pionero en el desarrollo de enfoques de la evaluación centrados en el aprendizaje, particularmente en la construcción de competencias evaluativas para el aprendizaje a largo plazo (*Developing Evaluative Judgement in Higher Education*, Routledge, 2018) y diseñando nuevos enfoques de la retroalimentación (*Feedback in Higher and Professional Education*, Routledge, 2013). A lo largo de 2020 se publicará el texto editado bajo el título "Reimaginando la evaluación en la universidad en un mundo digital" (*Re-imagining University Assessment in a Digital World*, Springer, 2020)



Revista Electrónica de **I**nvestigación y **E**valuación **E**ducativa
E-Journal of Educational Research, Assessment and Evaluation

[ISSN: 1134-4032]



Esta obra tiene [licencia de Creative Commons Reconocimiento-NoComercial 4.0 Internacional](https://creativecommons.org/licenses/by-nc/4.0/).
This work is under a [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by-nc/4.0/).