



www.universitaria.cl

Dieudo LECLERCQ



Álvaro CABRERA MARAY



UNIVERSIDAD
DE CHILE



Directores de la publicación:

Dieudonné Leclercq
Universidad de Liège (ULg)

Álvaro Cabrera Maray
Universidad de Chile (UCH)

IDEAS e INNOVACIONES
Innovaciones en Dispositivos de Evaluación
de los Aprendizajes en la enseñanza Superior
2014

Se pueden bajar gratuitamente
desde <http://orbi.uliege.be>, después Leclercq D., o
desde www.evaluaraprendizajes.cl

- Los **resúmenes** de los 23 capítulos
del libro IDEAS <http://hdl.handle.net/2268/173543>
- El **índice** de este libro para buscar entre
entradas de 1500 conceptos y
400 de autores <http://hdl.handle.net/2268/180060>

Dieudonné Leclercq

Dr. en Educación (1975) en « La Metacognición vía la autoevaluación con grados de certeza » y con postdoctorales en las universidades de Pittsburgh y UCLA. Fue profesor en las Universidades de Namur (1975-1980) y de Liège (1980-2010). Es emérito desde 2010. Enseña como invitado en las Ues. de Liège y Paris 13. Recibió el título de *Honorary Member of the World Cultural Council* (México). Ha colaborado, en Chile, con la U de Chile (UCH -Santiago), la UMCE, la UCT (Temuco), la UC del Maule, la UNAB y la UCSC (Concepción). En Perú con la PUCP y el SINEACE (Lima), la UNSAAC (Cusco) y la UNTRM (Chachapoyas). En México con la U A Chapingo. En España con la U de Sevilla y la U de Deusto (Bilbao). d.leclercq@uliege.be

Álvaro Cabrera Maray

Licenciado en Artes mención Teoría de la Música, y Master en Pedagogía en Educación Superior de la U. de Liège (Bélgica). Ha sido profesor en la Facultad de Artes y en Cursos de formación General, trabajando en el Depto. Estudios de Pregrado de la U. de Chile a cargo del Área de Formación. Integró la Red nacional de Centros de Enseñanza-Aprendizaje y la de expertos SCT-Chile sobre sistema de créditos transferibles. Trabajaba en el Ministerio de Educación de Chile, coordinando los programas de la reforma educacional en Educación Superior. alvarocabreramaray@gmail.com

Contenidos del libro IDEAS:

ES: Calificación ; Evaluación ; Productos ; Meta-cognición ; Resolución de problemas ; Proyectos ; Trabajo de grupo ; Portafolio ; Vigilancia cognitiva ; Pruebas de Progreso ; Taxonomía de Bloom ; Auto-evaluación ; Grados de certeza ; Test de Concordancia de Script ; Retroinformación ; calidades ; validez

EN : Assessment ; Evaluation ; Outcomes ; OSCE ; MCQ ; PARMs ; Metacognition ; Problem solving ; Projects ; Group produced work ; Portfolio ; Cognitive vigilance ; Progress Tests ; Bloom's Taxonomy ; Self-assessment ; Confidence Degrees ; Concordance Script Test ; Feedbacks ; Edometrics ; Metacognitive Spectral Test ; ETIC PRAD ; quality ; validity

FR : Notation ; Evaluation ; Résultats ; ECOS ; QCM ; PARMs ; Métacognition ; Résolution de problèmes ; Projets ; Travail de groupe ; Portfolio ; Vigilance cognitive ; Tests de progression ; Taxonomie de Bloom ; Auto-évaluation ; Degrés de certitude ; Test de Concordance de Script ; Rétro-information ; Edumétrie ; Test Spectral Métacognitif ; qualités d'une évaluation ; validité d'une mesure

IDEAS = Innovaciones en Dispositivos de Evaluación de los Aprendizajes en la educación Superior

La lista de los capítulos y el resumen de cada uno

aparece a continuación después de este capítulo.

P A R T E

4

Principios
estratégicos
en evaluación

CAPÍTULO XX

Concebir Dispositivos de Evaluación de los Aprendizajes (DEA) al nivel de un programa

ÁLVARO CABRERA, CEES VAN DER VLEUTEN Y DIEUDONNÉ LECLERCO

A. ¿Por qué planificar un Dispositivo de Evaluación de los Aprendizajes (DEA) al nivel de un programa?

Como parte del movimiento por implementar el aprendizaje basado en competencias, han surgido iniciativas que buscan declarar resultados de aprendizaje y competencias no solo a nivel de cada curso, sino que definidos pensando en un programa completo; en la totalidad del plan de formación (Adam, 2006; Bowe y Fitzmaurice, 2009; Prégent *et al.*, 2009). Este movimiento se basa en los principios de la Triple Concordancia¹⁸⁵ (entre objetivos, métodos y evaluaciones) y del LLL¹⁸⁶ (Lifelong Learning o Aprendizaje a lo largo de la vida).

P. Knight (2000) lo demostró¹⁸⁷: los docentes universitarios tienen más o menos los mismos objetivos en términos de las competencias genéricas que sus cursos se proponen desarrollar (espíritu crítico, búsqueda de información en la literatura científica, tolerancia a la incertidumbre, etc.). Pero en programas tradicionales, con frecuencia es hacia el final de la carrera que esos objetivos se desarrollan y se evalúan. Durante los primeros años de la carrera cada docente tiene la responsabilidad sobre sus propias evaluaciones (con función formativa o certificativa-sancionadora), sin tener en cuenta los métodos, momentos y contenidos de las evaluaciones de los otros docentes, las que a menudo incluso ignora.

De esta situación resultan varias desventajas:

- (1) Nadie se preocupa de asegurar la validez Teórica
 - “de concepción”, por ejemplo, la coherencia de la evaluación con los resultados de aprendizaje y los métodos.
 - “de cobertura”: que todos los aprendizajes esperados sean evaluados, evitando que algunos lo sean evaluados muchas veces y otros nunca.

¹⁸⁵ Porque la visión panorámica del plan de formación permite detectar vacíos y repeticiones poco productivas en la relación entre objetivos/aprendizajes esperados, métodos y evaluaciones. La gestión coordinada del conjunto de actividades de aprendizaje y evaluación debiese tender a asegurar que se alcance la Triple Coherencia.

¹⁸⁶ Porque al proveer al estudiante de una estructura (orden y progresión) de los logros de aprendizaje esperados se facilita el aprendizaje autoadministrado y la autorregulación; porque esta estructura también permite al estudiante entender cómo el programa de formación le ayudará a aprender; porque el desarrollo y evaluación de todas las competencias de un plan (asegurado por la Triple Coherencia) demanda variedad de estrategias, métodos e instrumentos, lo que provoca variedad de experiencias y de canales de aproximación al estudiante, lo cual favorece su aprendizaje.

¹⁸⁷ Ver Capítulo 7, sección A.3.

- (2) Estas pruebas aisladas no presentan situaciones complejas, aunque es la definición misma de las competencias: actuar en situaciones complejas. En consecuencia, estos exámenes separados no preparan a los estudiantes en la integración de los saberes aprendidos en varias asignaturas, ni en su movilización en una actuación en situación.
- (3) Cuando los/las docentes formulan por separado sus exámenes con la intención de evaluar desempeños complejos (competencias), resulta en un trabajo excesivo de corrección para cada uno de ellos, midiendo muchas veces las mismas cosas que sus colegas.
- (4) Se producen acumulaciones de trabajo en algunos periodos del año para los estudiantes.
- (5) Frente a estas acumulaciones de trabajo el estudiante está con frecuencia sometido a un dilema que llamamos “de las dos trampas”. Ilustramos a continuación el dilema con un ejemplo ficticio.

EL DILEMA DE LAS DOS TRAMPAS

“Ud. debe producir un trabajo escrito que trate de la resolución de un problema y que sea (1) pertinente, (2) original, y (3) detallado” es la consigna que ha recibido un estudiante de tres profesores diferentes, a cada uno de los cuales debe entregar un trabajo distinto.

El escenario de la trampa 1 ocurre cuando los profesores X, Y y Z se dan cuenta de que en los tres informes —que en principio debían ser diferentes según el criterio de originalidad— hay 3 secciones (D, E y F) que son comunes: crimen contra la originalidad. La acusación es que el estudiante ha intentado engañar a cada uno de los tres profesores entregándoles (parcialmente) el mismo trabajo. ¡Culpable!

El escenario de la trampa 2 ocurre cuando el estudiante ha gestionado de modo que no existe ninguna intersección entre los tres informes, lo que resulta en que los tres profesores consideran que el informe que han recibido no es bastante pertinente o detallado (otras condiciones impuestas por la consigna). El profesor X se queja de que el estudiante no ha evocado o detallado el contenido D (que el estudiante ha elegido reservar para el informe al docente Y). El docente Z se queja de que el estudiante no ha tratado del contenido o asunto F (que el estudiante reservó para el profesor X). Crímenes contra la pertinencia y el detalle. ¡Culpable!

Un escenario sin estas trampas ocurriría si se hubiera decidido que todos los profesores (X, Y y Z) leerán los tres informes o, más racionalmente, que el estudiante escribirá un solo informe que será leído y calificado por los tres docentes. Por supuesto, un informe así debe presentar el nivel de calidad de los anteriores tres trabajos, pero sin ser tres veces más extenso (ganancia en términos de cantidad-síntesis)... y desarrollar ideas y soluciones que integren todo lo que ha aprendido. Pero, ¿es realista este último desafío?

B. ¿Por qué es difícil hacerlo? (las desventajas de un DEA concebido en común)

En carreras tradicionales rara vez se implementa un Dispositivo de Evaluación de los Aprendizajes concebido en común, con excepción de la tesis al final del plan de formación. Esto ocurre por varias razones:

- (1) Cada docente desea que su contenido sea evaluado en los niveles de la taxonomía de los procesos cognitivos¹⁸⁸ que él o ella prefiere y que sabe cómo evaluar.
- (2) A cada profesor le gusta leer (y evaluar) únicamente los contenidos de su especialidad.
- (3) Coordinarse entre profesores requiere tiempo. Pertenecer a un comité de elaboración de pruebas comunes demanda tiempo (ver Capítulo 18).
- (4) Las decisiones comunes pueden resultar en una limitación del poder y de la libertad de cada docente. En particular, cada docente debe respetar las fechas clave (*deadlines*) de producción del material de evaluación.
- (5) Cada docente pierde el poder de decidir, por sí solo, el éxito o fracaso anual o semestral de un estudiante sobre la base de los resultados en su propia asignatura.
- (6) Cada docente pierde el poder de innovar en términos de principios o técnicas de evaluación. Tiene que convencer a sus colegas para que sus ideas sean aceptadas al nivel del colectivo. Sin embargo, si logra convencerlos, su idea será implementada a un nivel más amplio.
- (7) En algunos programas (en la Facultad de Medicina de la Universidad de Maastricht, por ejemplo), los profesores (o mentores/tutores) son docentes distintos a aquellos que crean las pruebas, para evitar la confusión de roles, lo que exige fuertemente al sistema de gestión de la docencia de la institución.

C. Los ritmos de implementación de las innovaciones

C.1. ¿Nacer o renacer?

Cada institución, en función de su historia y su nivel de desarrollo, implementa las innovaciones a su propio ritmo, en particular cuando se trata de luchar contra la inercia¹⁸⁹ o de cambiar el rumbo. Se debe notar que la Universidad de Maastricht, que evocamos tan frecuentemente, no tuvo que afrontar estas dificultades de inercia y cambio de rumbo. La Facultad de Medicina de la Universidad de Maastricht nació, en 1974, con el proyecto pedagógico del ABP —Aprendizaje Basado en Problemas— como

¹⁸⁸ Como los descritos por Bloom *et al.* (1956): conocimiento de memoria, comprensión, aplicación, análisis, síntesis-producción-expresión, juicio-evaluación.

¹⁸⁹ Esta “inercia” se debe entender como en física: “tendencia a conservar el movimiento adquirido”.

un elemento central de su plan de desarrollo. El edificio de la facultad, las salas y espacios de indagación y estudio grupal, fueron concebidos arquitecturalmente integrando esta perspectiva. Los docentes fueron contratados sobre la base de su compromiso con los principios del ABP. Incluso los estudiantes, cuando eligen esta facultad (entre todas las que existen en el país), se comprometen con este método, del cual están perfectamente conscientes. Habitualmente, una facultad que se propone realizar innovaciones curriculares no está en condiciones tan favorables. *Es más fácil nacer que renacer.*

C.2. ¿Evaluar performances aisladas (recursos/saberes) o las competencias?

En la docencia tradicional los recursos (aislados) son los únicos componentes que son evaluados, medidos y calificados. Las competencias no lo son, lo que dificulta la integración y promueve que el estudiante estudie “para la prueba”. Sin embargo, consideramos que implementar dispositivos de evaluación que intenten evaluar únicamente resultados de aprendizaje y competencias sería reemplazar un extremo por otro, con el peligro de no poder asegurar que cada estudiante domina todos los recursos que debe dominar.

Es la posición de los autores que, en la actual fase de transición desde la educación con enfoque centrado en la enseñanza/el profesor hacia la formación con el centro en el aprendizaje/el estudiante, los educadores debemos tender a implementar sistemas de evaluación variados y combinados, midiendo los recursos por separado cuando sea pertinente y TAMBIÉN evaluándolos movilizados en actuaciones.

C.3. Componentes del DEA al nivel de un programa

La sección D, a continuación, presenta cuatro ejemplos de combinaciones de métodos y tiempos (¿en cuáles años introducir cuáles evaluaciones?). Tres de estos ejemplos han incorporado, en varios niveles de profundidad, un mismo principio metodológico: el Aprendizaje Basado en Competencias (ABC) en forma de la metodología de Aprendizaje Basado en Problemas (ver Capítulo 18).

El ejemplo más avanzado es la Facultad de Medicina de la Universidad de Maastricht, lo que sin duda está relacionado con haber comenzado este camino en 1974. En Maastricht se combinan los métodos de evaluación descritos en los puntos 1 a 5, de acuerdo con el listado que sigue. En la sección D veremos con cuál ritmo se desarrollaron históricamente y en cuáles años de la carrera:

1. *Las Pruebas por Unidades de Contenido (PUC)*. Es el modo de evaluación más clásico. La originalidad de la Facultad de Medicina de la Universidad de Maastricht es que una prueba de este tipo tiene lugar cada seis semanas. Se llaman “Bloc tests”, y tratan de los contenidos que han sido aprendidos durante estas semanas en los grupos tutoriales, porque los casos de los pacientes simulados conciernen contenidos específicos, que pueden y deben ser evaluados. Estos contenidos siguen una lógica de complejidad creciente a medida que avanzan los años de la carrera.

2. *Las Pruebas de Progreso o PdP* (ver Capítulo 18): desde el primer año hasta el último, cada 3 meses (4 veces por año), todos los estudiantes contestan a una prueba de 200 preguntas sobre toda la medicina (todas las especialidades médicas y todos los sistemas del cuerpo humano). Las PdP aseguran que cada una de las especialidades esté representada en cada prueba, con las preguntas de los profesores especialistas. Por ejemplo, el profesor de cardiología puede estar seguro de que en cada PdP aparecerán 13 preguntas de cardiología, lo que resultará en que en seis años de duración de la carrera cada estudiante habrá contestado a 312 preguntas en este tema.
3. *Los Exámenes Clínicos Objetivos y Estructurados o ECOEs* (ver Capítulo 8) en Maastricht están reservados a los estudiantes del último año. Se basan en el principio de simulación de encuentros entre un médico (en formación) y un paciente (simulado).
4. *El portafolio* (ver Capítulo 12), contiene varios documentos que constituyen evidencias de aprendizajes. Incluye observaciones de terreno, reflexiones, e incluso un artículo, si se quiere incluir una investigación personal. Este portafolio se desarrolla continuamente durante los 6 años de formación, de modo que provee muchas oportunidades de reflexibilidad (desafío clave de un portafolio) y de retroalimentaciones (de los pares, de los tutores). En Maastricht, el portafolio permite *integrar* producciones de naturalezas variadas, y será leído por un jurado de varios docentes. Por ejemplo, en el portafolio se integran reflexiones sobre los resultados en las 24 PdP sucesivas, y las regulaciones de aprendizaje que el estudiante ha llevado a cabo. Integra también evaluaciones de aprendizaje que el estudiante ha llevado a cabo. Integra también evaluaciones en 360° (ver Capítulo 3E). Este portafolio, que se desarrolla durante años y combina informaciones muy diversas (incluido lo que en Maastricht llaman “Aprendizaje en terreno, en el lugar de trabajo”) no puede ser reemplazado por un “súper examen” al final de la carrera.
5. *Las Pruebas Integradoras (PI) de varios contenidos*. En la Facultad de Medicina de la Universidad de Liège, desde la introducción de los ARPCs (Aprendizaje en la Resolución de Problemas Complejos; ver sección D2), los estudiantes de los Masters 3 y 4 rinden una prueba integradora, que equivale a una evaluación común a todos los profesores.

D. Ejemplos de proyectos de DEA al nivel de un programa

A continuación presentamos cuatro ejemplos de carreras, de entre 5 y 7 años de duración, que combinan (en distintos porcentajes) evaluación de recursos aislados y evaluación de desempeños complejos en sus DEAs, considerados *a nivel del programa completo*. Llamamos la atención del lector sobre las fechas en que las innovaciones fueron introducidas.

D.1. La carrera de Medicina de la Universidad de Maastricht (Holanda)¹⁹⁰

A) CONTEXTO

Esta carrera es conocida mundialmente por su opción por el Aprendizaje Basado en Problemas (ABP) como el principal método de formación, lo que en su caso implica ABP desde el primer día y ausencia casi total de clases expositivas.

B) EL DEA

El Dispositivo de Evaluación de los Aprendizajes (DEA) de esta carrera consiste en 5 tipos de instrumentos, que la Tabla 1 muestra de acuerdo al año de formación en que son utilizados (con función sancionadora), junto a la fecha en que fueron introducidos.

- PUC (Pruebas de Unidades de Contenidos) fueron utilizadas desde el comienzo (1974).
- PP o Pruebas Prácticas fueron rápidamente añadidas (1979).
- ECOEs o Exámenes Clínicos Objetivos Estructurados (ver Capítulo 8) fueron introducidos en 2000, para evaluar desempeños en situaciones complejas simuladas. Esto vino a fortalecer la coherencia con el enfoque metodológico del ABP.
- PdP o Pruebas de Progreso (ver Capítulo 18), introducidas en 1978, constituyeron una innovación mayor: evalúan recursos aislados, y se repiten en forma equivalente cuatro veces cada año durante toda la carrera, permitiendo una visión longitudinal de la evolución del dominio de recursos.
- Portafolios son la más reciente innovación (2006), para integrar estos variados componentes, sobre los cuales se superpone la reflexividad.

Tabla 1: Fechas de introducción de innovaciones en evaluación al nivel de la carrera de Medicina en la Universidad de Maastricht

U. Maastricht Medicina		1	2	3	4	5	6
MÉTODOS	Aprendizaje Basado en Problemas (ABP)	1974					
	Pruebas por unidades de contenido	1974	x	x	x	x	x
Evaluación (DEA)	Pruebas prácticas	1979	x	x	x	x	x
	Pruebas de Progreso (PdP)	1978	x	x	x	x	x
	Examen Clínico Objetivo Estruct. (ECOEs)						2000-
	Portafolio						2006-

¹⁹⁰ Van der Vleuten, 1998.

D.2. La carrera de Medicina de la Universidad de Liège (Bélgica)¹⁹¹

A) CONTEXTO

Esta facultad combina una enseñanza y evaluación tradicional (centrada en los recursos aislados), con métodos integradores e instrumentos para la evaluación de desempeños complejos, introducidos en la carrera en los primeros años del siglo XXI. Por ejemplo,

- en 2000 se introdujeron los seminarios de ARC (Aprendizaje del Razonamiento Clínico) al nivel del cuarto año de carrera y en los años que siguen.
- en 2002 se introdujo el método de ABP en el tercer año (solamente), y dos años después (2004) este método fue extendido a los estudiantes de un semestre más abajo (4º semestre) y uno más arriba (7º semestre), de los 14 semestres que componen el plan de formación.
- en 2003 se introdujeron los seminarios de ARPC (Aprendizaje en la Resolución de Problemas Complejos), donde grupos de estudiantes de los últimos dos años (6º y 7º años) tratan casos multidimensionales (que a partir de un problema médico, involucran aspectos sociales, legales, policiales, etc.) con la presencia de expertos de estos diversos ámbitos.

B) EL DEA

- PUCs (Pruebas de Unidades de Contenido) eran, desde hace muchos años, el único método de evaluación. Algunos departamentos introdujeron innovaciones en las PSMS, pero estuvieron limitadas a esos departamentos.
- ECOEs (Exámenes Clínicos Objetivos Estructurados) fueron introducidos en 2005 en pediatría y en 2007 en medicina general.
- EMS (Entrevista Médica Simulada) son estaciones de ECOE del doble de duración –15 minutos– donde el estudiante debe conducir una entrevista médica con todos sus pasos.
- PIs (Pruebas Integradoras de varios contenidos) fueron introducidas con los ARPCS en 2003.
- TCS (Tests de Concordancia de Scripts (ver Capítulo 19) fueron introducidos en 2009, desde el año 3 al año 6, como instrumento para evaluar el razonamiento clínico, que se ubica en un punto intermedio entre la evaluación de recursos aislados y la evaluación de desempeños complejos.

¹⁹¹ Giet et al., 2005.

Tabla 2: Fechas de introducción de innovaciones en evaluación al nivel de la carrera de Medicina en la Universidad de Liège

U. Liège Medicina		1	2	3	4	5	6	7	
MÉTODOS	Aprendizaje en el Razonam. Clínico (ARC)				2000-				
	Aprendizaje por Resolución de Problemas Complejos (ARPC)						2003-		
	Aprendizaje Basado en Problemas (ABP)			2002-					
	Aprendizaje Basado en Problemas (ABP)		2004-						
EVALUACIÓN (DEA)	Pruebas por unidades de contenido	? x x x x x x x x x x x x x x x hasta hoy							
	Pruebas prácticas	? x x x x x x x x x x x x x x x hasta hoy							
	Pruebas integrando varios contenidos						2003-		
	Pruebas de Progreso (PdP)								
	TCS (Test de Concordancia de Scripts)				2009				
	Examen Clínico Objetivo Estruct. (ECOE)						2007-		
Portafolio							2010-		

D.3. La carrera de Farmacia en la Universidad de Liège (Bélgica)¹⁹²

En esta carrera la enseñanza se mantiene en su mayoría dentro de los cánones tradicionales, y es durante el sexto y último año de carrera que se introduce el ABP y el Portafolio como instrumento de evaluación de desempeños complejos (documenta 6 meses de práctica profesional en terreno). Al final del sexto año además se realiza un ECOE, descrito en el capítulo 8.

Tabla 3: Fechas de introducción de innovaciones en evaluación al nivel de la carrera de Farmacia en la Universidad de Liège

U. LIÈGE FARMACIA		1	2	3	4	5	6
MÉTODOS	Aprendizaje Basado en Problemas (ABP)						
	Pruebas por unidades de contenido	? x x x x x x x x x x x x x hasta hoy					
Evaluación (DEA)	Pruebas prácticas	? x x x x x x x x x x x x x hasta hoy					
	Pruebas integrando varios contenidos						2000-
	Pruebas de Progreso (PdP)						
	Examen Clínico Objetivo Estruct. (ECOE)						2008-
	Portafolio						2009-

¹⁹² Philippe et al. (2009).

E. Tendencias generales

No intentamos generalizar lo que se observa en los tres casos que preceden. Son tan solo ejemplos prototípicos que pueden servir para ilustrar tendencias.

E.1. El orden de las innovaciones varía de una institución a otra

En algunas situaciones las innovaciones comienzan con los métodos de formación (por ejemplo en Maastricht con el ABP). En otras, las innovaciones de métodos y de evaluación son simultáneas (por ejemplo: en la Facultad de Medicina de la ULG, con los seminarios de ARPCs y las PI). En otras, es solo en la evaluación donde hay innovaciones (por ejemplo, en el Instituto de Farmacia de la ULG con el uso de ECOES y de portafolios).

E.2. Los avances son por saltos

Innovar necesita tiempo, en parte porque los actores deben esperar que pasen las olas de desequilibrio introducidas por la innovación anterior, antes de iniciar otra.

E.3. La relación entre el momento de la formación y el tipo de evaluación que predomina

- a) Los primeros años tienden a concentrarse en la evaluación de recursos en forma aislada (procesos mentales simples, como memorización, comprensión, aplicación).
- b) Más adelante pueden introducirse evaluaciones de recursos integrados en torno a problemas básicos, y que suman el análisis a los procesos mentales simples (por ejemplo, describir la mitosis desde la química, la física y la biología celular, y comentar las relaciones).
- c) Luego aparecen instrumentos que además demandan el proceso mental de evaluación (juicio, valoración) y evalúan los recursos integrados, como aquellos que evalúan aspectos del razonamiento profesional en contexto de test (Tests de Concordancia de Scripts; Análisis Fraccionado de Casos – ver capítulos 19 y 7).
- d) Hacia el final de los programas aparecen las evaluaciones de desempeños complejos en situaciones controladas pero intencionadamente “reales” en sus características esenciales (simulaciones), como los ECOE, las EMS, algunos tipos de evaluaciones orales (como un “alegato simulado” en Derecho), algunos tipos de reportes escritos (como un ensayo que se escribe “como si fuera a ser publicado” pero se mantiene en el contexto escolar). Demandan la creación además de los demás procesos mentales.
- e) Mención especial merecen los procesos metacognitivos, que deben ser fomentados y evaluados durante todo el plan de formación. Esto es posible hacerlo desde el inicio, incluso en la evaluación de recursos aislados (ver Capítulo 16). Evidencias de

reflexividad y metacognición pueden ser intencionadas usando instrumentos de evaluación como el Test Espectral Metacognitivo (ver Capítulo 9), y las secciones metacognitivas que exija el formato del portafolio integrador.

E.4. El rol integrador del portafolio

Los desempeños en situaciones profesionales o académicas cercanas a lo que espera al estudiante luego de su titulación o egreso son evaluados hacia el final de los procesos formativos, mediante

- (1) los proyectos estudiantiles, sobre todo cuando tienen un socio comunitario o profesional externo a la Universidad (como en el Aprendizaje + Servicio), o con una audiencia en vivo (como en el caso de conciertos, obras de teatro / danza, y/o exposiciones de estudiantes de Artes),
- (2) los internados, terrenos, prácticas profesionales o estadías en equipos de investigación.

En ambos casos el *portafolio* es una buena opción de instrumento para evaluar los aprendizajes producidos en estas situaciones cercanas a la realidad profesional.

F. Conclusiones

Los métodos e instrumentos para evaluar *recursos* siguen siendo de relevancia para la evaluación de los aprendizajes en la educación superior, y en particular es urgente su revisión y ajuste para hacerlos más pertinentes y de mejor calidad (por ejemplo, evaluar procesos mentales superiores a la memorización, o sumar la seguridad a la exactitud de la respuesta –para estimar el realismo y el conocimiento parcial). Junto a ellos se deben implementar métodos e instrumentos para evaluar la movilización de esos recursos en *actuaciones complejas*, crecientemente cercanas a la realidad profesional.

Por esta razón, este libro aborda ambos aspectos, dedicando los Capítulos 6 al 11 a la *evaluación de desempeños complejos*, los capítulos 13 al 19 a la *evaluación de recursos*, y el Capítulo 12 *al portafolio integrador*.

Referencias

- ADAM, S. (2006). An introduction to learning outcomes. In EUA Bologna Handbook, European university association, article B 2.3-1. Berlin: Raabe.
- BLOOM, B.S., ENGELHART, M.D., FORST, E.J., HILL, W.H. Y KRATHWOHL, D.R. (1956). Taxonomy of educational objectives. Handbook 1: Cognitive domain, New York, McKay.
- BOWE, B. Y FITZMAURICE, M. (2009). Guide to Writing Learning Outcomes. Dublin Institute of Technology (www.bcit.ca/files/ltc/pdf/htoutcomes.pdf).

- GIET, D., MASSART, V., STIR, A., FREYENS, A., FIRKET, P. Y BONIVER, J. (2005). Approche de la complexité contextuelle et de la pluridisciplinarité de l'action médicale: mise en place de séances d'apprentissage à la résolution de problèmes complexes (ARPC) en fin de 2^e cycle d'études médicales. *Pédagogie médicale*, vol. 6, no2, pp. 88-97.
- KNIGHT, P. (2000). Skills Plus: Employability in higher education. <http://www.lancs.ac.uk/users/edres/research/skillsplus/index.htm>. (consultado el 7 de junio de 2007).
- PHILIPPE, G., ANGENOT, L., GIET, D., LECLERCQ, D., MASSART, V. Y BOURGUIGNON, J.-P. (2009). Comparaison de séminaires d'apprentissage à la résolution de problèmes complexes et de conférences magistrales sur le développement de compétences pharmaceutiques. *Pédagogie Médicale*, 10(S1), S97.
- PRÉSENT R., BERNARD H. Y KOZANITIS A. (2009). Enseigner à l'université dans une approche programme, Presses internationales Polytechnique.
- VAN DER VLEUTEN, C. (1998). The assessment of professional competence: Developments, research and practical implications. *Advances in Health Sciences Education*. 1, p. 46-67.

IDEAS E INNOVACIONES Dispositivos de Evaluación de los Aprendizajes en la educación

Dieudonné LECLERCQ y Álvaro CABRERA MARAY 2014

Resumen de cada capítulo

Los editores y autores principales del libro

p. 11-13

Prologo

Álvaro Cabrera &
Dieudonné
Leclercq

Parte 1: Conceptos clave en educación

p. 15-20

1	ATOME (Alineamiento en un Tablero de Objetivos, Métodos y Evaluaciones. Da una visión panorámica de los tres pilares de un programa de formación: los objetivos (y sus 4 niveles de alcance), los Métodos (y sus 8 Eventos de Enseñanza-Aprendizaje), las evaluaciones (y sus 4 niveles de profundidad), insistiendo sobre la Triple Concordancia (u alineamiento) O-M-E y dando ejemplos de inconsistencia.	D.Leclercq & Álvaro Cabrera p. 23-34
2	Los componentes de un dispositivo de evaluación de los aprendizajes (DEA) Da una visión de los vínculos entre las finalidades (formativas o sancionantes) de la evaluación, las competencias que desarrollar y los recursos que dominar, las condiciones de un dispositivo, las herramientas y los criterios de calidad de cada componente de un DEA.	D. Leclercq p. 35-50
3	El prisma de las características de un Dispositivo de Evaluación de los Aprendizajes (DEA) Presenta las características y las condiciones de un DEA como las facetas de un prisma: Quien (los agentes) evalúa, cuando (de manera definitiva o mejorable), quienes (individuo o grupo), para quienes (pública o confidencial), como (objetivamente o subjetivamente; estandarizada o adaptativa), que modifican la medición o su interpretación.	D. Leclercq p. 51-82
4	ETIC PRAD: Ocho criterios de validez de un Dispositivo de Evaluación de los Aprendizajes (DEA) Presenta 8 tipos de validez de un componente de un DEA: Ecológica (cerca de la situación real), Teórica (razonamiento o teoría que lo funda), Informativa (o diagnóstica), Consecuencial (lo que resulta del componente), Predictiva (correlada con otras mediciones), Replicabilidad (o fiabilidad), Aceptabilidad (para los profesores, los estudiantes, el público), Deontológica (equitativo).	D. Leclercq p. 83-92
5	Autodescribir y evaluar el Dispositivo de Evaluación de los Aprendizajes (DEA) de un curso Propone una secuencia que puede seguir un profesor para definir un DEA para su curso, es decir sus objetivos, sus métodos y sus evaluaciones, presentándoles en una tabla de modo que aparecen los vínculos y las ausencias de vínculos.	D. Leclercq & Álvaro Cabrera p. 93-102

6	<p>La calificación subjetiva de los desempeños complejos: Criterios y rubricas Presenta la docimología y sus evidencias de los efectos de notación o de calificación subjetiva (ley de Posthumus, ausencia de concordancia intra y inter-jueces, efectos de halo, de secuencia, de estereotipo, de confirmación (o de inercia). Además de esta docimología “negativa”, presenta principios de una docimología positiva y varios tipos de escalas (ej: la de Mercali) y rubricas.</p>	<p>D. Leclercq & Álvaro Cabrera p. 103-128</p>
7	<p>Evaluar la capacidad de resolver problemas Explica la diferencia entre una pregunta y un problema, el cono de la experiencia (Dale), y las heurísticas de Polya para resolver problemas. Da varios ejemplos de evaluaciones apropiadas a medir la capacidad y detectar los procesos utilizados en la resolución de problemas: las cascadas convergentes y divergentes, las análisis fraccionadas de casos (AFC), la facilitación progresiva, la medición de la búsqueda de información (Shannon, Rimoldi). Da ejemplos de medición de la creatividad, de la capacidad de aproximación y una teoría de la auto-fijación de la dificultad, como de la perseverancia.</p>	<p>D. Leclercq, S. Delcomminette (HERS) & A. Cabrera p. 129-152</p>
8	<p>ECOE: Exámenes Clínicos Objetivos y Estructurados Esta técnica consiste en una sucesión de estaciones en cada de cuales se juegan roles (simulaciones) donde el profesor juega el paciente (el estudiante jugando el del medico o de la enfermera) u el cliente (el estudiante jugando el del farmacéutico), o... para medir competencias, es decir capacidad de actuar en situación compleja. El sistema de notación incluye las actitudes, las destrezas, y la cognición. Las reacciones de los participantes como la predictividad de estas mediciones son presentadas.</p>	<p>G. Philippe (ULg), D. Leclercq & J-P. Bourguignon (ULg) p. 153-170</p>
9	<p>Meta cognición y Tests Espectrales Metacognitivos (TEMs) Para los docentes que quieren desarrollar y medir capacidades como la vigilancia cognitiva, el espíritu crítico, la auto-evaluación (y la meta cognición) y el desarrollo epistemológico es presentada el método “Test Espectrales Meta cognitivos” que combina PSM con SGI (cap. 13, 14 y 15), grados de certeza (cap. 15 y 16), debate y reflexión meta cognitiva. Presenta los aspectos técnicos como los resultados obtenidos en varios ámbitos (cognitivo, epistemológico, meta cognitivo).</p>	<p>D. Leclercq & Álvaro Cabrera p. 171-196</p>
10	<p>Evaluar los Aprendizajes en la Pedagogía Por Proyectos (PPP) La PPP permite de desarrollar y medir competencias complejas (incluido trabajar en equipo), con un enfoque sobre rubricas, tan como sus componentes (recursos) en términos de cognición, actitudes, destrezas. Se puede aplicar los principios de evaluación a 360° (por los pares, por su mismo, por los docentes, por el público). El capítulo plantea (y ilustra sobre un caso) el problema de la convergencia (o ausencia de congruencia) entre estas varias fuentes de evaluación, y el problema de la ponderación de los criterios.</p>	<p>Álvaro Cabrera p. 197-220</p>
11	<p>Evaluar la contribución de cada participante a un trabajo grupal Distingue colaboración y cooperación, presenta los elementos que deben ser parte de un contrato al inicio, y después presenta 6 métodos para evaluar el valor añadido de cada participante al trabajo de grupo. Ilustra el método 4 (declaraciones de participación) con un ejemplo, el de PARMs (Proyectos de Animación Reciproca Multimedia) y sus criterios DECLAR, el método 5 (observación continua con la simulación de actividad parlamentaria y el método 6 (observar la colaboración) con la pauta de Bales. .</p>	<p>D. Leclercq, P. Gillet (ULg), M. Erpicum (ULg) & A. Cabrera p. 221-242</p>
12	<p>Los Portfolios: Hacia una evaluación más integrada y coherente con el concepto de desempeño complejo Este principio (y método) de evaluación sirve no solo a evaluar desempeños complejos como estancias en terreno, sino de constituir una integración de varias evaluaciones. Es ilustrado en dos carreras de la universidad de Liège: Formasup o Master en Pedagogía Universitaria (con sus instrucciones o consignas de redacción del portfolio) y el Master en Logopedia (que permite de discutir de 4 niveles de calidad de evidencias).</p>	<p>M. Poumay (ULg) & Chr. Maillard (ULg) p. 243-260</p>

13	<p>Las Preguntas de Selección Múltiples (PSM): del currículo escondido a la vigilancia cognitiva Presenta los retos del currículo oculto y de la espontaneidad vs la limitación a respuestas sobre sollicitación. Explica como la vigilancia cognitiva se puede entrenar y medir con una consigna valida por las PRB (Preguntas a respuesta Breve) y las PSM (Preguntas a Selección Múltiple): las Soluciones Generales Implícitas (SGI) como “Ninguna, Todas, falta datos, Absurdo”. Da una definición muy precisa de PSM, sus formas de presentación, sus ventajas y desventajas y presenta los modelos mentales que cada de 8 consignas (instrucciones) favorece. Presenta la fórmula que vincula la fiabilidad de la nota final en la prueba, el número de PSM y el número de soluciones en ella.</p>	<p>D. Leclercq & Álvaro Cabrera p. 261-286</p>
14	<p>Reglas de redacción de las Preguntas de Selección Múltiples y la habilidad para responder pruebas Presenta 24 reglas (repartidas en 5 categorías) y los dispositivos experimentales (preguntas sobre contenidos ficticios) que permiten verificarlas, tan como los resultados de estas verificaciones en caso de transgresión de las reglas.</p>	<p>D. Leclercq p. 287-300</p>
15	<p>Evaluar procesos cognitivos según la Taxonomía de Bloom Presenta modalidades de evaluación apropiadas a cada de los 6 niveles de los procesos mentales descritos en la taxonomía de Bloom: la memoria (de re-cognición y de evocación), la comprensión (con la definición de Smedslund), la aplicación, el análisis (y las Preguntas PRIM-BIS para diferenciar entre análisis y comprensión, la síntesis y la creación (y los criterios de Torrance), el juicio(incluido la capacidad de aproximar).</p>	<p>D. Leclercq p. 301-328</p>
16	<p>Auto-evaluación con grados de certeza: un microscopio para la evaluación de los aprendizajes Presenta los retos del uso de grados de certeza: epistemológico (de definición de “dominio”), de medición en investigación (la necesidad de un microscopio del pensamiento), de caracterización practica (utilizable – inutilizable) de niveles de conocimiento) y de fijación de umbrales de éxito os resultados y de excelencia. Presenta las condiciones metodológicas de uso (3 principios), las distribuciones espectrales de calidad de les respuestas, las nociones de meta memoria y de meta comprensión (el JOC o juicio de comprensión).</p>	<p>D. Leclercq p. 329-356</p>
17	<p>Grados de certeza y docimología: como calificar Denuncia varios sistemas de cotejo inapropiados y la importancia (impredecible) de tener en cuanta el realismo de las respuestas acertadas por un estudiante en una prueba. Explica como verificar (con la ley binomial) la presunción de realismo, cálculo de un índice de calibración. Trata de la sobrestimación y de resolución (Discriminación y lucidez), tan como de una pauta innovadora de cotejo basada en ;los grados de certeza.</p>	<p>D. Leclercq p. 357-386</p>
18	<p>PdP: Pruebas de Progreso Presenta una modalidad de evaluación en cual la universidad de Maastricht se ha ilustrada como pionera: la Pruebas de Progreso que consisten en presentar el mismo día a todos los estudiantes de una carrera (que sean de primer o de ultimo año) una prueba sobre todos los contenidos de la carrera (centenas de preguntas), cuatro veces por año (con pruebas “paralelas”). Las ventajas y desventajas son revisitadas, como el modo de comunicar los resultados, original también. Estos principios son ilustrados por su aplicación en Maastricht desde cuarenta años.</p>	<p>D. Leclercq, A. Cabrera & C. Van der Vleuten (U. Maastricht) p. 387-408</p>
19	<p>TCS : El Test de concordancia de Script Esta técnica ha sido concebida para medir la capacidad clínica de tratar la información. Ha sido utilizada principalmente en medicina (revisión de opinión desde una información adicional). Es ilustrada con un ejemplo y resultados de su aplicación en la univ. de Liège.</p>	<p>V. Massart (ULg), A. Collard (ULg) D. Giet (ULg) p. 409-418</p>

20	<p>Concebir Dispositivos de Evaluación de los Aprendizajes (DEA) al nivel de un programa Presenta tres experiencias de desarrollo de un DEA al nivel de una facultad: la de Farmacia en Liège y las de medicina en Liège y en Maastricht.</p>	<p>D. Leclercq, C. Van der Vleuten & A. Cabrera p. 419-430</p>
21	<p>Retroinformaciones (Feedbacks) Empieza con el problema de la profundidad de penetración de una retroinformación, desde sobre los detalles de ejecución de la tarea hasta el <i>Self</i> (es porque son presentadas las teorías de William James sobre la auto-estima y la <i>FIT</i> o <i>Feedback Intervention Theory</i>). Un modelo integrador (llamado CAIRO) es presentado. Varios modos de presentación de las retroinformaciones después de una prueba son presentados. Una modalidad, utilizada en la UCH (Universidad de Chile) que se focaliza al esencial, es presentada con un ejemplo.</p>	<p>D. Leclercq, M. de la Fuente (UCH) & A. Cabrera p. 431-454</p>
22	<p>Los roles de un SMART: Servicio Metodológico de Apoyo a la Realización de Tests Un (SMART) ayuda docentes en la concepción y la realización de pruebas estandarizadas y en el procedimiento de las respuestas de los estudiantes (calculo de varios índices relativos a cada pregunta y cada solución de las PSM), como en las retroinformaciones automatizadas a los estudiantes. Un enfoque especial es dedicado al uso de cajas de voto a distancia (<i>clickers</i>).</p>	<p>D. Leclercq & P. Detroz (ULg) p. 455-476</p>
23	<p>Índices cuantitativos en Docimología Consiste en un catálogo de conceptos útiles para tratar cuantitativamente los datos resultando de evaluaciones estandarizadas como</p> <ul style="list-style-type: none"> -los tipos de categorías (nominales, ordinales, métricas). -los índices relativos a una distribución : índices de centración (Modo, Mediana, Media), de dispersión (rango, cuartiles, desviación estándar), de posiciones relativas o normativas (la nota z, los percentiles) de la forma de la distribución (asimetría o <i>skewness</i>). -las presentaciones gráficas de distribuciones. -índices de comparación o de progreso: la amplitud del efecto (AE), la ganancia relativa (GR). -la fiabilidad de la nota (<i>reliability</i>) al total de la prueba y el alfa de Cronbach. -el umbral de éxito, fijado a priori o a posteriori. -el índice de discriminación (correlación punto <i>biserial</i> o <i>rpbis</i>) de un modo de respuesta aplicado a cada de las soluciones de cada PSM -el análisis automática de una prueba -el valor heurístico de los nubes de puntos. 	<p>D. Leclercq, R. Roco (Chile) & A. Cabrera p. 477-543</p>
24	<p>Index de los autores 426 autores citados.</p>	<p>D. Leclercq & A. Cabrera p. 545-549</p>
25	<p>Index de los conceptos Se puede bajar gratuitamente via http://hdl.handle.net/2268/180060</p>	<p>D. Leclercq & A. Cabrera</p>