

III ENCUENTRO INTERNACIONAL UNIVERSITARIO



LA EVALUACIÓN DE COMPETENCIAS EN LA **EDUCACIÓN SUPERIOR**

Ponencias y conversatorios

Leandro García-Calderón Palomino, editor



LA EVALUACIÓN DE
COMPETENCIAS EN LA
EDUCACIÓN SUPERIOR

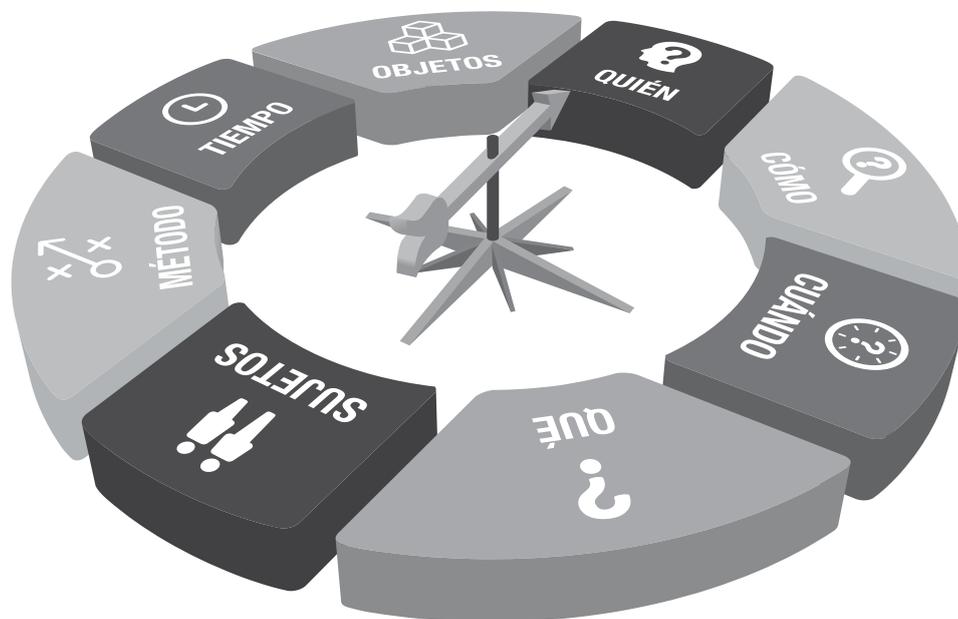
Ponencias y conversatorios

III Encuentro Internacional Universitario

LA EVALUACIÓN DE COMPETENCIAS EN LA EDUCACIÓN SUPERIOR

Ponencias y conversatorios

Leandro García-Calderón Palomino, editor



La evaluación de competencias en la educación superior

Editor: Leandro García-Calderón Palomino

Revisión general: María Teresa Moreno Alcázar

© Pontificia Universidad Católica del Perú
Dirección de Asuntos Académicos, 2017
Av. Universitaria 1801, Lima 32, Perú
Teléfono: (51 1) 626-2000
www.pucp.edu.pe

Diseño de carátula:

Interactiva Studio S.A.C., basado en la rosa de los vientos de D. Leclercq

Diseño y diagramación:

Carmen Inga Colonia

Corrección de estilo y cuidado de la edición:

Daniel Enrique Amayo Magallanes

Primera edición digital: agosto de 2017

ISBN: 978-612-47238-2-7

Ruta de acceso:

http://cdno2.pucp.education/academico/2017/08/16123430/III_EncuentroInt_evaluacion_competencias_en_edusup.pdf

ISBN: 978-612-47238-2-7



Índice

Presentación	9
Introducción	11
PONENCIAS	13
Diez preguntas para autoevaluar el perfil de egreso, el programa y las asignaturas. Desafíos ejemplificados / Dieudonné Leclercq	15
Evaluación de competencias a nivel microcurricular: aplicación en un curso real / Álvaro Cabrera Maray	55
La evaluación de competencias en la educación superior / José Rafael López Islas	91
Seguimiento de egresados y evaluación de la formación por competencias / Juan Fernando Vega Ganoza	117
El reto de la evaluación por competencias en las asignaturas de Humanidades / Gemma Puigvert Planagumà, Josep Antoni Iglesias-Fonseca	163
CONVERSATORIOS	187
Primer conversatorio: Marco general de la evaluación de competencias y la evaluación en el aula	189
Segundo conversatorio: Sistemas de evaluación en la educación superior: evaluación del plan de estudios y los estudios de egresados	215
Sobre los ponentes	235

Presentación

Desde hace tres años, la Dirección de Asuntos Académicos de la PUCP organiza encuentros internacionales universitarios que tienen como propósito primordial esclarecer el camino que implica adoptar un currículo por competencias en el ámbito de la educación superior. En el año 2014, en el primer encuentro, se mostraron las experiencias en la adopción, implementación y evaluación de propuestas formativas por competencias en universidades de Chile, Colombia y México. En el año 2015, en el segundo encuentro, fueron las “competencias genéricas” las que articularon los ejes que fueron trabajados por especialistas de España, Canadá y México.

Para el año 2016, nos comprometimos a organizar un encuentro relacionado con la evaluación de competencias en los diferentes niveles del currículo. Con el objeto de lograr este fin, nos pusimos en contacto con especialistas amigos, no solo porque conocíamos de sus calidades como personas, sino porque tenían la comprensión y la experiencia para ayudarnos a recorrer el camino que emprendimos hace más de una década.

Para este III Encuentro Internacional Universitario: La evaluación de competencias en la educación superior, convocamos a Dieudonné Leclercq de la Universidad de Lieja, Álvaro Cabrera Maray del Teachers College, Columbia University, José Rafael López Islas del Tecnológico de Monterrey y Juan Fernando Vega Ganoza de la PUCP. Para ellos solo tengo palabras de agradecimiento por sus deseos de enseñar y compartir su inapreciable conocimiento.

Además, quiero dar las gracias a los profesores de la PUCP, Patricia Escobar, del Departamento de Educación, y Sandro Paz, del Departamento de Ingeniería, quienes no dudaron un instante cuando les pedimos participar en los conversatorios del segundo día del encuentro.

Asimismo, agradezco a la Dra. Gemma Puigvert, profesora titular de Filología Latina de la Universidad Abierta de Cataluña, quien ha contribuido con nosotros con un artículo sobre su experiencia en la implementación del enfoque por competencias en un curso de humanidades.

Finalmente, quiero reconocer los esfuerzos y dedicación de los integrantes del equipo de Gestión Curricular para organizar este encuentro. Hago extensivo este reconocimiento a los otros integrantes de la Dirección de Asuntos Académicos que han participado en este evento contribuyendo con su creatividad, experiencia y conocimiento.

En este III Encuentro Internacional Universitario: La evaluación de competencias en la educación superior, cuyos temas, ponencias, conversatorios y conclusiones son el motivo de la presente publicación, ha quedado claro que el proceso evaluativo es clave en la consecución del fin primordial del currículo: formar ciudadanos competentes que, enfrentados a un problema, movilicen los recursos necesarios, internos y externos, para hallar una solución que evidencie su formación profesional y su valía como personas.

JORGE ZEGARRA PELLANNE
Director de Asuntos Académicos de la PUCP

Introducción

Los días 11 y 12 de octubre de 2016 se llevó a cabo, en el campus de la PUCP, el III Encuentro Internacional Universitario: La evaluación de competencias en la educación superior. Tal como indica el nombre del encuentro, este devino naturalmente de los dos anteriores para ocuparse de un tema crucial en el reto de implementar currículos por competencias en las universidades: la evaluación del perfil de egreso como impacto de la formación del profesional en el mundo laboral, la evaluación del plan de estudios como camino hacia el logro del perfil y la evaluación a nivel del aula como aseguramiento del desarrollo de las competencias en la estructura fina de la propuesta formativa.

Cuando se tiene que evaluar competencias, dada la imbricación de los recursos que los estudiantes deben movilizar, se hace necesario poner en juego una inusual batería de técnicas de enseñanza-aprendizaje y, por consiguiente, métodos de evaluación que aseguren la adquisición de aquellas habilidades que los expertos de la carrera han considerado necesarias e imprescindibles para ejercer su profesión.

La definición de competencia ha evolucionado de acuerdo a la teorización de académicos de alta valía, así como de la experiencia acumulada en la implementación de modelos educativos por competencias, especialmente en el ámbito del espacio europeo de educación superior dentro del proceso de Bolonia. Actualmente, se entiende que un individuo es competente si sabe actuar en una situación, movilizando recursos internos y externos. Detrás de esta definición, subyace la convicción de que todos los seres humanos, independientemente de la carrera que hayamos emprendido y aprendido, estamos constantemente enfrentándonos a problemas de distinta índole, donde la emergencia y la incertidumbre no son la excepción, sino la regla.

En la definición anterior, está el mandato implícito de que la actuación tiene que ser evaluada en tanto es observable y, dado el despliegue de recursos de diverso jaez que hace un estudiante para abordar un problema, la evaluación tiene que ser multidimensional, caleidoscópica y planteada en situaciones muy cercanas a su desempeño futuro como profesional de una carrera.

A diferencia de los métodos de evaluación estándar, en un enfoque pedagógico que pretende desarrollar competencias, el error debe ser tomado en cuenta como oportunidad de aprendizaje y no necesariamente como indicio para sancionar. En este sentido, Leclercq reivindica las pruebas de opciones múltiples sobre la base de especificar los grados de certeza en las respuestas dadas: en ciertas circunstancias, un error cometido con mucha convicción puede ser extremadamente contraproducente.

En las propuestas formativas por competencias, el trabajo menudo en clase está orientado a contribuir al desarrollo y mantenimiento de las habilidades que deben tener los egresados de una carrera universitaria. Por tanto, se hace perentorio verificar que las diversas unidades —sean cursos, módulos, talleres o cualquier otra actividad— efectivamente conduzcan al logro del perfil; de allí la necesidad de establecer un mecanismo de aseguramiento de la calidad que permita comprobar la efectividad de los planes de estudio y eventualmente cambiarlos de manera parcial o total.

Por último, todo el trabajo de estudiantes, profesores y autoridades de una carrera en una universidad se pone a prueba en el espacio y el tiempo posteriores al egreso. En otras palabras, si un perfil de salida no contempla las necesidades actuales y proyectadas de la sociedad en que vivirá el futuro profesional, los años de formación universitaria podrían convertirse en una completa pérdida de tiempo. Por esa razón, un estudio de egresados debe indicarnos, en la medida de lo posible, el impacto de los perfiles de egreso de las carreras de una institución de educación superior en la vida de sus egresados y en la sociedad que constituyen.

Estos temas fueron tratados en el encuentro por cuatro especialistas en la enseñanza y evaluación de competencias: Dieudonné Leclercq de la Universidad de Lieja, Álvaro Cabrera del Teachers College de la Columbia University, José Rafael López del Tecnológico de Monterrey y Juan Fernando Vega de la PUCP. A todos ellos mi más profundo agradecimiento.

En este año 2017, en el que la Pontificia Universidad Católica del Perú cumple 100 años de fundación, la Dirección de Asuntos Académicos y los miembros del equipo de Gestión Curricular esperamos que esta publicación sea del interés y utilidad para todos aquellos que creen en la mejora continua de la formación universitaria como sustento de la misión que nuestra institución tiene: formar personas felices y altamente competentes en todos los aspectos de la vida.

LEANDRO GARCÍA-CALDERÓN PALOMINO
Equipo de Diseño Curricular
Dirección de Asuntos Académicos de la PUCP

Ponencias

Diez preguntas para autoevaluar el perfil de egreso, el programa y las asignaturas. Desafíos ejemplificados

■ Dieudonné Leclercq

A lo largo de mi carrera, en diversas publicaciones, he descrito conceptos e instrumentos haciendo uso de muchas ilustraciones en cuanto a lo que se “puede” hacer y lo que a veces se hace en lo que respecta a la evaluación de los aprendizajes, incluso en sitios remotos del mundo. El presente artículo trata de lo que, desde mi punto de vista, se “debe” hacer.

A menudo recomendaré al lector que profundice en los temas, consultando uno de los siguientes documentos en castellano, empezando con el más reciente y completo (las referencias completas figuran en la bibliografía al final del artículo)¹: *Ideas e innovaciones* (Leclercq & Cabrera, 2014). La expresión “Ideas” significa “Innovaciones en Dispositivos de Evaluación de los Aprendizajes en la enseñanza Superior”. Este fue el título inicial del libro que coordinamos Álvaro Cabrera y yo, y que la editorial decidió llamar *Ideas e innovaciones*. Este libro (de 552 páginas) contiene veintitrés capítulos escritos por diecisiete autores.

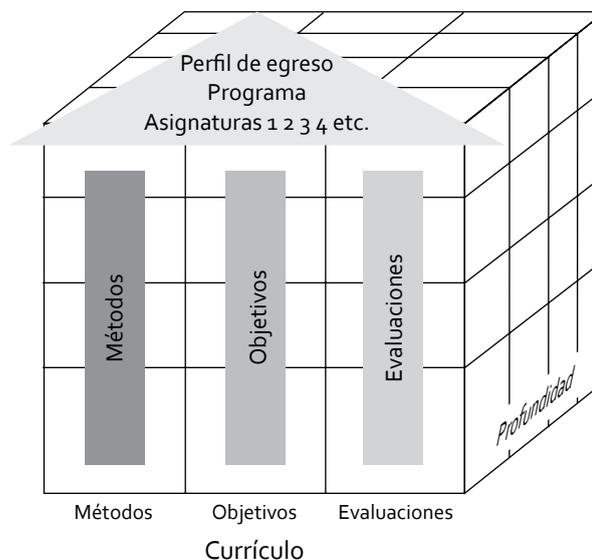
LA ARQUITECTURA DE UN CURRÍCULO

Comenzaré haciendo referencia a los dispositivos de evaluación de los aprendizajes (DEA) que constituyen uno de los tres pilares que sostienen el perfil del egresado, el programa (o plan de estudios) y las asignaturas o cursos (ver figura 1). Un DEA incluye no solo el pilar “Evaluaciones”

¹ Del libro *Ideas e innovaciones*, del cual se puede bajar gratuitamente, desde la web, los resúmenes de cada uno de los veintitrés capítulos (<http://hdl.handle.net/2268/173543>) y el índice analítico (<http://hdl.handle.net/2268/180060>). También es posible bajar de la Web, en forma gratuita, los otros tres documentos (véase Leclercq & Cabrera, 2014).

(a la derecha de la arquitectura), sino también el pilar central “Objetivos”, porque cada elemento de evaluación debe referirse a objetivos precisos.

FIGURA 1
Los tres pilares de un currículo



En el presente artículo, se explica los cuatro pisos de esta arquitectura; en la pregunta 4, el eje “Profundidad”; y en la pregunta 6, la necesaria “triple concordancia” entre los tres pilares².

LA ROSA DE LOS VIENTOS O EL PRISMA DE LAS FACETAS DE LAS EVALUACIONES

Tal como se expresa en el título de este artículo, los diez desafíos prioritarios de la autoevaluación de los DEA son expresados a manera de preguntas. Espero que los argumentos y las ilustraciones faciliten las tres operaciones de la metacognición profesional de un docente: autojuzgar, autodiagnosticar y autoregular sus acciones.

² Los eventos de aprendizaje-enseñanza, que constituyen los elementos clave que componen los métodos, son descritos por Leclercq y Cabrera (2011, pp. 23-26).

FIGURA 2

Afiche del III Encuentro Internacional Universitario: "La evaluación de competencias en la educación superior"



El motivo de la línea gráfica del III Encuentro Internacional Universitario: “La evaluación de competencias en la educación superior”, está inspirado en nuestra imagen de la rosa de los vientos de las condiciones de un DEA (cfr. Leclercq & Cabrera, 2014, cap. 3) y presenta los aspectos que permiten caracterizar estos desafíos y sugerir algunas rutas en relación a los retos que estos plantean (ver figura 2).

Después de este preámbulo, paso a presentar los desafíos que implica autoevaluar los componentes del currículo, los cuales están ubicados en el “techo” de la arquitectura curricular, sostenido por tres pilares fundamentales: los métodos, los objetivos y la evaluación; esta última, motivo que guía el itinerario de este tercer encuentro.

■ PARTE 1: PREGUNTAS PARA EVALUAR EL PERFIL DEL EGRESADO

Las cuatro preguntas iniciales responden a interrogantes relacionadas con la evaluación de las competencias del perfil de egreso; es decir, de la evaluación a nivel macrocurricular.

PREGUNTA 1: ¿El perfil de egreso, junto con el plan de estudios, permite desarrollar en los estudiantes una arquitectura de cuatro tipos de competencias (DIME): dinámicas (D), instrumentales (I), metacognitivas (M) y específicas (E)?

Esta pregunta supone reflexionar sobre los desafíos que hoy enfrenta la formación en la educación superior. Los jóvenes que formamos hoy en las universidades no solo tendrán que enfrentar retos planetarios (ambientales, políticos, económicos, sociales); sino sus propios problemas personales (Villardón, 2015). Por eso, cada uno de ellos tiene que estar preparado a nivel cognitivo y también en términos de actitudes, relaciones y destrezas (Kozanitis, 2015).

En este contexto, conviene recordar las reflexiones de Piaget sobre la finalidad de la educación. En una entrevista televisada, Piaget³ declaró: “El verdadero problema de la pedagogía es el fin mismo de la educación. ¿Debemos formar individuos que solo son capaces de aprender lo que ya se conoce, de repetir lo que ya se ha aprendido en las generaciones anteriores o hay que formar personas innovadoras, creadoras, que sean capaces de inventar, desde que están en la escuela y, después, durante toda su vida?”.

3 Ver: <<https://www.youtube.com/watch?v=8MIqAbH6Owg>>.

La arquitectura de cuatro tipos de competencia que se muestra a continuación considera una diversidad de competencias que pueden ser desarrolladas en la formación y que permitirán un aprendizaje a lo largo de la vida.

La arquitectura de las competencias

Las competencias y los recursos personales —basándonos en la literatura sobre el tema— deben constituir un conjunto que podemos comparar con un barco de cuatro cubiertas que está permanentemente autoconstruyéndose (ver figura 3). Se desarrollan no solo en la escuela (puerto), sino también durante toda la vida (metáfora del océano), en varios contextos, especialmente cuando el estudiante cambia de especialidad, de ambiente, de objetivos. Las cubiertas del barco corresponden a los cuatro tipos de competencias.

FIGURA 3

La arquitectura de competencias y recursos personales comparados con las cuatro cubiertas de un barco



Las competencias específicas

Estas competencias están vinculadas con un ámbito profesional específico (por ejemplo, la Salud) y a profesiones específicas como Medicina, Farmacia, Enfermería, etc.; o las ciencias de la Ingeniería, de la Administración, de la Historia, de la Lingüística, etc. Muchos de los saberes y de las destrezas aprendidas en estas disciplinas son poco transferibles a otros ámbitos. Por ejemplo, tener conocimientos de Geología no me prepara para trabajar en equipo.

En la metáfora náutica, estas especificaciones (o especializaciones) son representadas por los varios tipos de barcos especializados en funciones precisas; como un rompehielos, un barco para pesca de arrastre, un crucero de guerra o un explorador oceánico. Sus diferencias son evidentes, como son evidentes las diferencias entre las profesiones.

Hay que mencionar que los aprendizajes son de desigual naturaleza, pertenecen a varios órdenes (para los cuales se han propuesto diversas clasificaciones o taxonomías⁴): cognitivos, afectivos (actitudinales), de implicación, sensorio-motriz, relacionales, etc.

Las competencias instrumentales o genéricas

Estas competencias también pueden llamarse “generales” porque se pueden utilizar en un gran número de ámbitos. Son numerosos los ejemplos de tales competencias: saber hablar, leer, escribir, utilizar instrumentos (un diccionario, un teléfono, un correo electrónico, *softwares* como Word, Excel, PowerPoint, etc.). En términos de la metáfora náutica, se trata de los componentes que se encuentran en todo tipo de barco: los motores, la quilla, el ancla, la hélice, el radar, las rejillas de aire, etc.

Este nivel de la arquitectura merece relevar su importancia. Cada universidad define su propia lista de competencias genéricas prioritarias (ver, por ejemplo, en México, Barrón Tirado, 2015; en Barcelona, López, 2015, láms. 3 a 5). En Europa, en la década de 2000, en el marco del proyecto Tuning, se confeccionó una lista de ellas para ser adoptadas y desarrolladas por los centros de educación superior europeos. América del Sur se inspiró en esta lista (ALFA Tuning)⁵, aceptando veintisiete de ellas (Pey, 2007). En la Universidad de Deusto, que ha jugado un rol importante en Tuning (Villa & Poblete, 2007), se priorizó cinco competencias genéricas de tal

4 Las más famosas se pueden consultar en el libro *Ideas e innovación* (Leclercq & Cabrera, 2011, pp. 35-46) y en Leclercq, 2007, láms. 25-30.

5 Ver: <<http://tuning.unideusto.org/tuningal/index.php?option=content&task=view&id=168&Itemid=196>>.

TABLA 1

Las veintisiete competencias genéricas de ALFA Tuning para América Latina y las cuatro más desarrolladas en los veinticinco cursos del pregrado de la UCH en 2006

1. Conocimiento sobre el área de estudio y la profesión.
2. Habilidades en el uso de las tecnologías de la información y de la comunicación.
3. Capacidad de abstracción, análisis y síntesis.
4. Capacidad para organizar y planificar el tiempo.
5. **Capacidad de comunicación en un segundo idioma.**
6. Capacidad para identificar, plantear y resolver problemas.
7. Capacidad de comunicación oral y escrita.
8. Habilidades para buscar, procesar y analizar información procedente de fuentes diversas.
9. Capacidad para tomar decisiones.
10. Compromiso ético.
11. **Capacidad de trabajo en equipo.**
12. Capacidad crítica y autocrítica.
13. Habilidad para trabajar en contextos internacionales.
14. Habilidades interpersonales.
15. Responsabilidad social y compromiso ciudadano.
16. Valoración y respeto por la diversidad y multiculturalidad.
17. **Capacidad de aprender y actualizarse permanentemente.**
18. Capacidad de motivar y conducir hacia metas comunes.
19. Compromiso con la preservación del medioambiente.
20. Compromiso con su medio sociocultural.
21. Capacidad de aplicar los conocimientos en la práctica.
22. Capacidad de actuar en nuevas situaciones.
23. Capacidad creativa (generar nuevas ideas).
24. Habilidad para trabajar en forma autónoma.
25. Compromiso con la calidad.
26. **Capacidad para formular y gestionar proyectos.**
27. Capacidad de investigación y desarrollo.

forma que todo perfil de egreso debe incluirlas: comunicación verbal, comunicación escrita, sentido ético, trabajo en equipo y uso de la tecnología (Villa & Poblete, 2011).

Kozanitis (2015, p. 73) menciona que, entre las doce competencias genéricas que el CEAB (*Canadian Engineering Accreditation Board*) hizo obligatorias, el Polytechnique de Montreal ya había privilegiado dos de ellas: la comunicación oral y escrita y el trabajo en equipo.

En 2005, la Universidad de Chile inició una reforma del pregrado en catorce facultades, incorporando en el currículo de cada facultad cuatro cursos “a la carta” que el estudiante podía elegir libremente de entre veinticinco cursos ofrecidos. La tabla 1 presenta la lista de las veintisiete competencias y resalta las cuatro más frecuentemente desarrolladas en veinticinco cursos de pregrado (Chauriye, 2007).

Las competencias autoreflexivas o las metacognitivas

Estas competencias se refieren al autoconocimiento (“lo que domino”, “lo que ignoro”, “lo que hago bien”, “el tiempo que necesito para hacerlo bien”, “mis necesidades de sueño, de descanso”, “mis preferencias”, etc.). Al contrario de las dos categorías precedentes de competencias, no hay un libro en el cual cada uno de nosotros pueda aprender quién es o en quién se transforma. Hay que aprenderlo viviendo y los docentes tienen que organizar eventos que favorezcan estos aprendizajes. En términos de la metáfora náutica, esto se puede comparar a la sala de los instrumentos de navegación, con los paneles de control, que permiten conocer permanentemente el rumbo, la velocidad, la cantidad de energía, la ubicación del barco en los mapas de navegación, etc.

Las competencias dinámicas o motivacionales

También se pueden llamar “de toma de decisiones”. Estas competencias consisten en la priorización de valores y acciones que permiten manejar los proyectos personales y la perseverancia en cumplirlos.

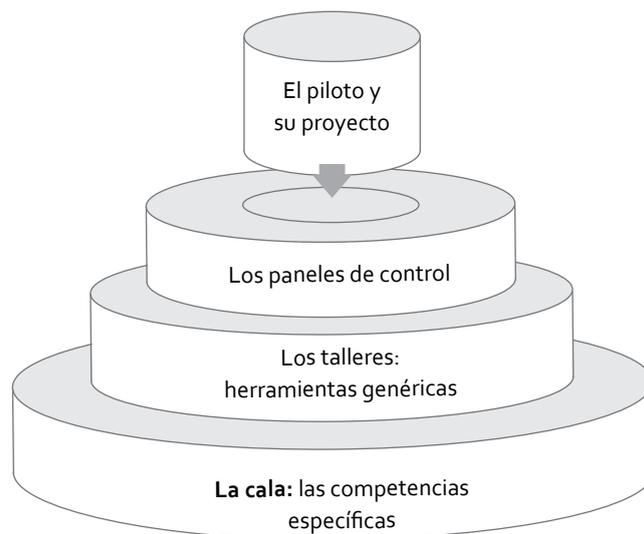
En términos de la metáfora náutica, se refiere a la posición de capitán, de piloto en el puesto de mando, que da las órdenes, que fija los rumbos, la velocidad, las etapas, las paradas y escalas a tierra, los conflictos internos y externos.

Las interacciones entre los niveles

La figura 4 ilustra que cada nivel no solo se apoya sobre los inferiores, sino que los incluye.

Por ejemplo, “comunicarse eficazmente” (competencia genérica) se hace sobre contenidos precisos (competencias específicas). Además, las competencias de los dos niveles inferiores de la arquitectura son monitoreadas por los dos niveles superiores: “Los paneles de

FIGURA 4
La integración de competencias entre varios niveles



control (competencia metacognitiva) son inútiles si nadie los lee o los utiliza (competencia dinámica)".

Una confirmación de esta propuesta de cuatro tipos de competencias o "intenciones educativas" se puede encontrar también en Knight (2000) de la Open University (UK). Knight ha estudiado el tema de la empleabilidad de los egresados de las universidades, entrevistando empleadores y preguntándoles cuáles son sus criterios prioritarios para contratar a un candidato egresado de una universidad. Sintetizando sus observaciones, ha destacado cuatro criterios:

- **Comprensión profunda:** se refiere a la competencia profesional; es decir, si se requiere a un químico, se buscará uno bueno.
- **Manejo de destrezas:** diferentes habilidades, como expresarse bien en varios idiomas, utilizar las TIC (tecnologías informáticas y de comunicación), manejar un proyecto, dirigir un equipo, etc.
- **Habilidades de metacognición:** conocer (y comprender) sus fuerzas y sus debilidades, así como regularse a corto plazo.

- **Sentido de autoeficacia:** el sentimiento (la convicción) de autoeficacia (descrito en Bandura, 1977a, 1977b, 1986, 1993, 1995 y 1997)⁶, hace referencia a la motivación, el compromiso profesional y social, las ganas de mejorar permanentemente.

Siguiendo otro itinerario de investigación, Knight (2000) propone categorías similares a las de la literatura propuesta por mí. Sobre esta base, Knight ha desarrollado un programa de formación llamado "Skills +", que significa "más allá de las destrezas", porque considera que las dos últimas categorías de competencias han sido (y siguen siendo) muy poco desarrolladas en las instituciones de enseñanza superior.

PREGUNTA 2: ¿Las competencias están definidas como "actuación en situación compleja (E), movilizando recursos internos y externos de modo integrado (I), reflexivamente (M) y de modo responsable (D)"?⁷

Esta definición, adaptada de Tardif, puede integrarse con la arquitectura de competencias DIME en lo que he llamado "definición de competencia a geometría variable"; porque, tal como se muestra en la tabla 2, es posible complementarla utilizando todas o algunas de las especificaciones entre paréntesis de la columna derecha en tinta azul.

TABLA 2

Definición del concepto competencia entendida como "actuación en situación"

DEFINICIÓN DE COMPETENCIA DE J. TARDIF (2006)	
La capacidad de actuar	(automáticamente)
en situaciones complejas	(no vividas previamente en forma idéntica, que pertenecen a una familia de situaciones)
movilizando e integrando recursos internos y externos	(saberes, actitudes, destrezas) (documentos, personas, instrumentos)
con ética y reflexividad	(metacognición)

⁶ Ver también Prieto Navarro (2001).

⁷ Esta pregunta es la consecuencia de la definición de una competencia propuesta en el capítulo 1 del libro *Ideas e innovación* (Leclercq & Cabrera, 2014, p. 27).

He adoptado las palabras de Jacques Tardif (2006), quien ha sugerido centrarse en dos aspectos al hablar de competencias:

- Los grados de “dominio” de capacidades precisas (como calcular un porcentaje, comprender un texto, orientarse en el terreno con un mapa, etc.), por lo cual es posible lograr un 20/20 en una prueba.
- Los niveles de “desarrollo” de la competencia, porque no hay límite en el nivel de excelencia que se puede alcanzar en el desarrollo de una competencia. Los niveles de desarrollo habituales son: novato (estudiante al ingresar a la universidad), en construcción (estudiante avanzado en su formación), competente (estudiante recién graduado), experimentado (egresado que ha ejercido su profesión durante años), experto (egresado que ha profundizado, escribiendo artículos o libros, dando conferencias)⁸.

La necesidad de concebir situaciones complejas movilizando recursos múltiples obliga a los profesores a concebir tareas concertadas —por ejemplo, proyectos (individuales o en equipo) o actuaciones en el ámbito profesional—, lo que hace necesaria la integración de los aprendizajes de los cuatro niveles de la arquitectura y de naturaleza cognitiva, actitudinal, sensorial, motriz y relacional.

Eso implica que un currículo se concibe “de atrás hacia adelante”, empezando con el perfil de egreso (las competencias requeridas al egreso), lo que es una tarea colectiva. Solo después se pueden definir las asignaturas, especificando la contribución de cada una a la formación común.

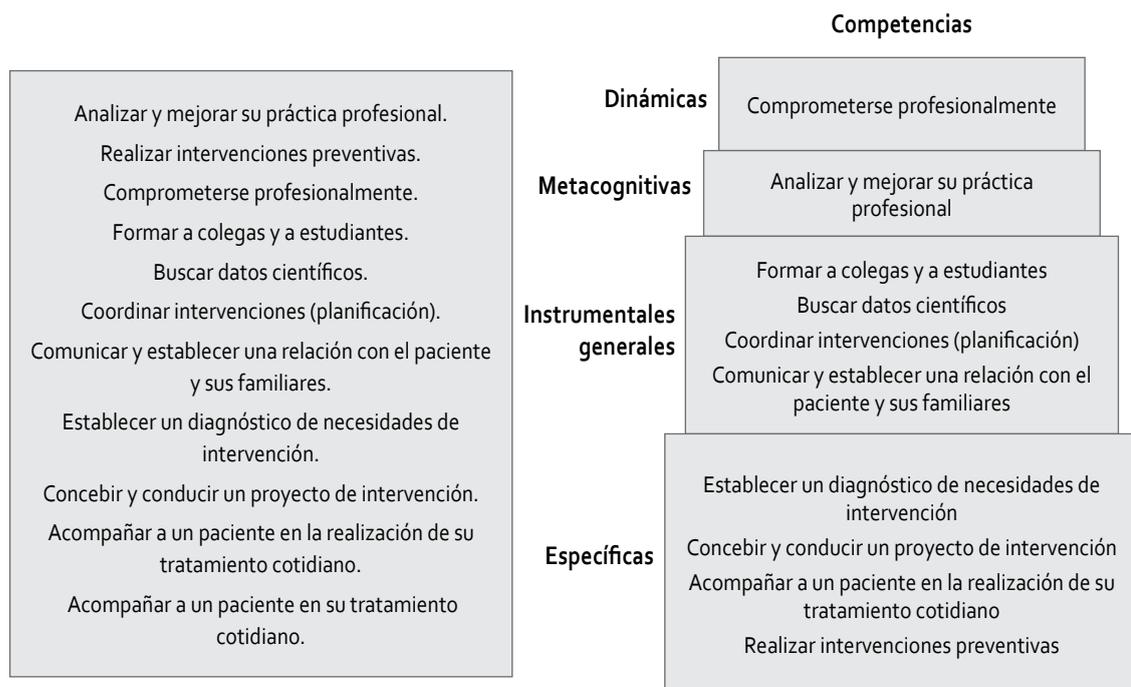
PREGUNTA 3: ¿El número de competencias desarrolladas por el currículo es limitado (no más de diez) y propone altas expectativas de modo que nunca se pierden de vista?

En un programa europeo de formación de enfermeras, se enumeran diez competencias, presentadas en la figura 5 (versión original a la izquierda) y que han sido organizadas según las cuatro categorías de la arquitectura de competencias (a la derecha), para darles mayor estructura.

8 Pueden verse ejemplos en López, 2015, láms. 12, 13 y 17.

FIGURA 5

Dos modos de presentar las competencias clave de una formación en Enfermería



PREGUNTA 4: ¿El currículo es evaluado a cuatro niveles de profundidad?

Donald Kirkpatrick (1983, 2006) describe cuatro niveles de profundidad para los modos de evaluación de los resultados de un programa de formación: satisfacción, aprendizajes, comportamiento en terreno y evaluación en instancias de último impacto. Estos niveles de profundidad se esbozan en la figura 1. A continuación, se hace una breve referencia a cada uno de estos niveles.

Evaluación de la satisfacción

Es el modo más fácil y más frecuentemente utilizado. Un ejemplo de este tipo de evaluación sería las encuestas de fin de ciclo que responden los estudiantes al terminar un semestre académico.

Evaluación de los aprendizajes

Es lo que se hace clásicamente en las instituciones de formación a través de los exámenes escritos u orales, de los portafolios de trabajo supervisado en el terreno y de las tesis de fin de formación.

Evaluación de los comportamientos en el terreno

Eso se hace más escasamente por varias razones. Primero, porque los egresados están dispersos en actividades profesionales diferentes, en lugares a veces alejados del sitio de formación. Segundo, porque este tipo de evaluación ocurre meses o años después de la formación. Tercero, porque los evaluadores (a menudo los supervisores locales de los empleados) son diferentes y es difícil evitar la subjetividad. Cuarto, porque las funciones profesionales varían, de modo que los criterios de evaluación deben variar también y eso hace difícil las comparaciones. Una posibilidad es la de utilizar portafolios “de desempeños y de desarrollo”, de “trayectorias personales” o “de itinerarios de formación” (ver Leclercq & Cabrera, 2014, cap. 12). Esto se hace cada vez con mayor frecuencia en la formación en salud (Medicina, Enfermería, etc.).

Evaluación del impacto sobre los beneficiarios últimos

Los beneficiarios últimos de la formación de médicos o de enfermeras son los pacientes; en el caso de los profesores, lo son sus estudiantes; de los ingenieros, son los usuarios de puentes o de represas hidroeléctricas, etc. Sin embargo, es difícil establecer la importancia del factor “formación” en estos desempeños últimos porque intervienen otros factores. Por ejemplo, la salud de los pacientes no depende únicamente de los médicos; sino de factores genéticos, ambientales, de recursos técnicos y sociales, etc.

Un modo de evaluación (pero no el único) consiste en entrevistar a los beneficiarios mismos o de estar atento a sus requerimientos de formación profesional. Por ejemplo, hoy en día las opiniones de los estudiantes sobre cada asignatura universitaria son sistemáticamente recogidas en operaciones llamadas EEE (Evaluación de la Enseñanza por los Estudiantes). Otro ejemplo es la investigación que realizan Tamblyn y otros (2007), quienes han calculado la correlación entre las quejas relativas a 3425 egresados de las facultades de dos provincias de Canadá entre 1993 y 1996 (seguidos hasta 2005) y las calificaciones en varios tipos de pruebas cuando estaban en la universidad⁹.

9 Los resultados de esta investigación se muestran en el capítulo 8 de *Ideas e innovación* (Leclercq & Cabrera, 2014, pp. 166 y 167).

■ PARTE 2: EVALUAR AL NIVEL DE UN PROGRAMA

Las preguntas 5, 6 y 7 desarrollan temas que corresponden al nivel meso o intermedio curricular; es decir, a la evaluación del programa o del plan de estudios por competencias.

PREGUNTA 5: ¿Se cuenta con un “enfoque de programa” que permite a cada docente especificar cuándo y cómo contribuye su curso a los tipos de competencias que el programa desarrolla (DIME)?

Intentaré responder en parte a esta pregunta haciendo referencia al dilema de las dos trampas.

El dilema de las dos trampas

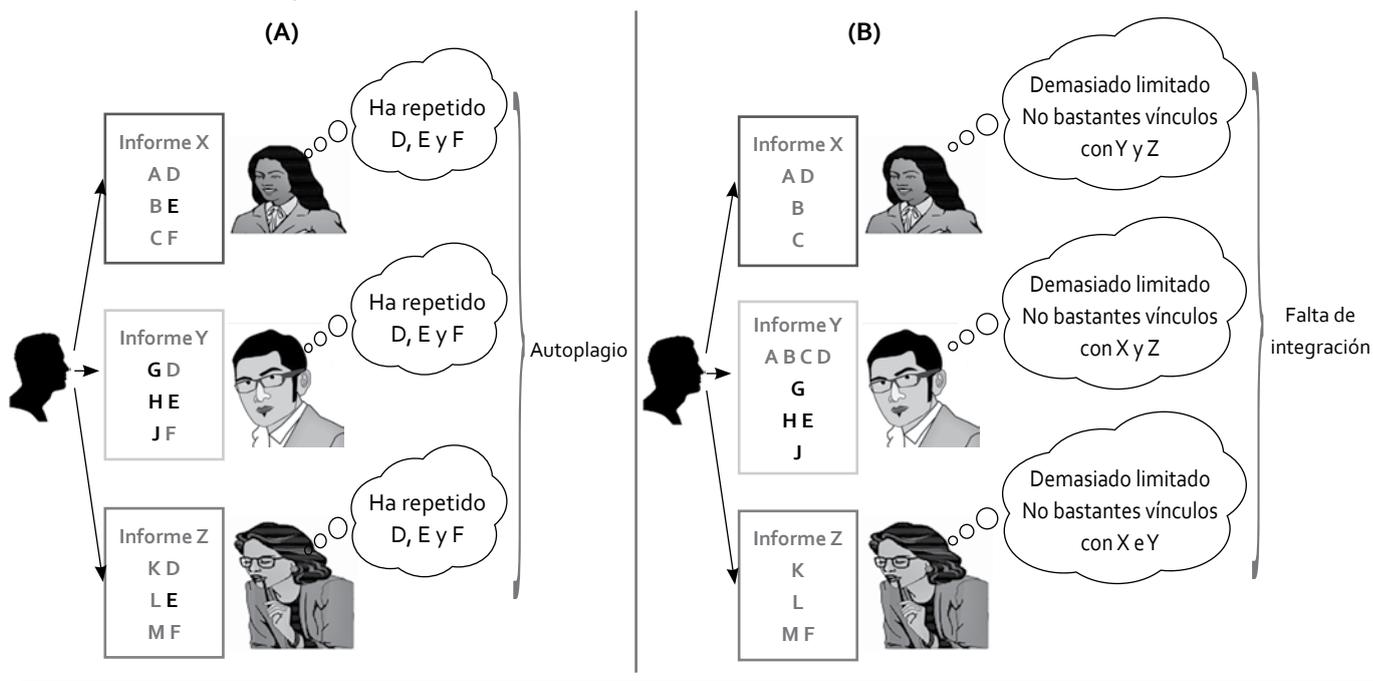
Frecuentemente, los estudiantes formulan la siguiente queja: “No veo para qué me sirve este contenido”. Una de las causas de ello es que algunas evaluaciones vinculadas a una asignatura aislada no están integradas en un desempeño complejo significativo. Además, cuando se trata de entregar informes distintos a varios docentes, los estudiantes enfrentan lo que se puede llamar “el dilema de las dos trampas” (ver Cabrera, 2014, p. 420).

La figura 6A describe la trampa 1, en la cual el estudiante ha repetido unos elementos (D, E y F) en tres informes distintos. En este caso, los docentes pueden acusarlo de haber intentado engañarles, presentando lo mismo a los tres. La figura 6B describe la trampa 2, en donde el estudiante no ha repetido contenidos. En este caso, los docentes podrían quejarse de que el estudiante no ha vinculado los contenidos, reprochándole el no ser capaz de relacionar las cosas.

Una posible solución a este dilema puede ser que los tres profesores se pongan de acuerdo en la realización de una tarea compleja y la califiquen con un solo informe en el cual el estudiante ha destacado (con criterios previamente convenidos) los asuntos o temas que se derivan de la asignatura de cada docente. Sin embargo, deberá ser parte del informe también el establecer claramente un problema, un cuerpo argumentativo y unas conclusiones, las que además deberán ser calificadas por cada docente. La posible solución admitida se enmarca dentro del llamado “enfoque de programa”:

El Enfoque de programa, conocido también como Enfoque de carrera, es un modelo de organización de la formación universitaria que promueve que las competencias definidas en el perfil

FIGURA 6
El dilema de las dos trampas



del egresado se desarrollen en los cursos del plan de estudios de manera gradual y secuencial, lo que asegura la articulación entre los tres niveles curriculares. Este enfoque, de alguna manera, le exige al cuerpo docente un trabajo colaborativo pues requiere la concertación de diferentes aspectos que forman parte de la carrera: desde el establecimiento del perfil del egresado hasta las estrategias de enseñanza y evaluación de los cursos vistos en conjunto y teniendo como meta el logro de las competencias del perfil (Kozanitis, 2015).

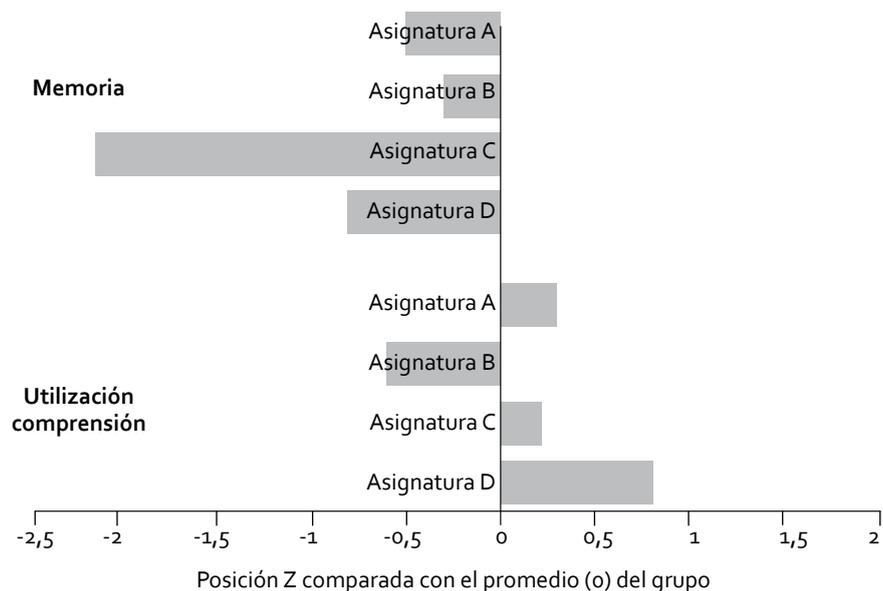
Mediciones repetidas en varios cursos

Cuando se busca recoger información relacionada con el ejercicio o utilización de diversos procesos mentales (por ejemplo, la capacidad de recordar lo memorizado o la capacidad de comprender y utilizar el saber), es recomendable que varios profesores midan estas capacidades y comparen sus resultados por cada estudiante —esto lo hice en 2003 (Leclercq & Detroz,

2003) con otros tres colegas—. Los estudiantes recibieron, de manera individual, un gráfico llamado “radiografía Z de procesos mentales”, en el cual registraban sus fortalezas utilizando una escala (comparada a la media del grupo) en el uso de la memoria y en cada una de las cuatro asignaturas. La figura 7 presenta la “radiografía Z” de un estudiante (G) que es mejor en comprensión que en memorización. Sus resultados se ubican a la izquierda de la línea vertical, que representa la media del grupo (para tres de las cuatro asignaturas, sus resultados se ubican a la derecha del 0).

FIGURA 7

La radiografía Z del estudiante G en la operación RESSAC (2003)¹⁰



Nota: la operación RESSAC es llamada así porque consiste en los resultados de exámenes estandarizados al servicio del aprendizaje en candidaturas (por ejemplo, pregrado).

10 En el capítulo 3 de *Ideas e innovación* (Leclercq & Cabrera, 2014, pp. 59-61), se muestra adicionalmente gráficos personales, entrevistas a estudiantes sobre lo que habían aprendido sobre sí mismos gracias a este gráfico, el impacto que estos resultados habían tenido sobre los cambios de estrategia para estudiar y, finalmente, las tasas de éxito (en cada asignatura) del grupo de los que habían modificado sus métodos de estudio y del grupo de los que no.

Verificación de las intenciones de formación y las evaluaciones efectivas

Knight (2000) preguntó a quince docentes (colegas suyos) que participaban en un mismo programa (plan de estudios) si su asignatura contribuía a desarrollar las doce capacidades del programa. En la tabla 3, las asignaturas aparecen en columnas con un código numérico (100 T1, etc.) y las capacidades, en filas (A.1, A.2, etc.). En las celdas, un *check* (✓) indica una respuesta “sí” y “x” una respuesta “no”. Luego, Knight preguntó a estos mismos docentes si podían entregarle evidencias (instrumentos y resultados) de evaluación de estas capacidades; las celdas sombreadas corresponden a las respuestas positivas.

TABLA 3
Resultados de la encuesta de Peter Knight (2000)

The contribution to programme learning outcomes made by key courses															
Skill outcomes (Programme Specification, section 10.2)	100 T1	100 T2 [1]	100 T3	200	205	211	225	227	228	231	232	234	235	236	300 [1]
A.1. Critical capabilities	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
A.2. Argumentation	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
A.3. Open-mindedness	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
A.4. Tolerance of Ambiguity	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
B.1. Information-handling	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
B.2. Research skill	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
B.6. Presentational	✓	✓	✓	✓	✓	✓	✓	✓	X	✓	✓	✓	✓	✓	✓
C.1. Reflectiveness	✓	✓	✓	✓	✓	✓	✓	✓	X	✓	✓	✓	✓	✓	✓
C.2. Independence	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
C.3. Problem-working	✓	✓	✓	✓	✓	X	✓	✓	✓	✓	✓	✓	✓	✓	✓
C.5. Interpersonal	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
C.6. Groupwork	✓	✓	✓	X	X	X	✓	✓	✓	✓	✓	✓	✓	✓	X

En la tabla 3, se puede observar que, de entre las 330 celdas, solo siete tienen una “x”. Eso significa que, en el 98% de los casos, estos profesores consideraban que su asignatura contribuía al desarrollo de estas doce capacidades del programa. De las 324 celdas marcadas con *check*, solo cuarenta y cuatro están sombreadas (14%); lo cual significa que solo un pequeño porcentaje de estas capacidades fue efectivamente evaluado. Sabemos que los estudiantes no invierten energía

ni esfuerzo en capacidades que no son evaluadas. Entonces, ello nos lleva a dudar de la hipótesis de que las contribuciones esperadas hayan sido las que se registraron en primera instancia.

En 2002, se repitió esta experiencia en un programa con once docentes y diez capacidades genéricas de la universidad de Lieja y se obtuvo resultados semejantes. A la primera pregunta, el 81% de las respuestas fue positiva y solo para el 27% de ellas se declararon evidencias de evaluación (ver Leclercq & Cabrera, 2014, p. 132).

PREGUNTA 6: ¿Cada asignatura evidencia la triple concordancia entre objetivos¹¹, métodos y evaluaciones (O-M-E)?

Lo que la literatura en inglés llama *alignment*, lo llamo “triple concordancia” para insistir en los tres componentes que deben tener una asignatura o un currículo. La figura 8 esquematiza la situación de un currículo ficticio en el cual hay cuatro métodos a la izquierda y tres modos de evaluar a la derecha. En esta figura, las líneas indican las concordancias entre los objetivos, que es el pilar central del currículo, y cada uno de los componentes de los otros dos pilares (métodos y evaluaciones). La ventaja de este modo de presentación (Leclercq & Cabrera, 2014, p. 32) es que, cuando un objetivo se vincula con un método M₁ y con una evaluación E₃, resulta, por transitividad, que M₁ y E₃ se vinculan entre sí. Este principio permite mantener simple el esquema. En la figura 8, se ha señalado tres lugares en los cuales falta concordancia: la señal 1 indica objetivos enseñados pero no evaluados; la señal 2 indica un objetivo evaluado pero no enseñado; y la señal 3, un método que no aporta a ningún objetivo.

Un ejemplo de cómo se informa sobre la triple concordancia _____

La tabla 4 presenta el formato que solía entregar a los trescientos estudiantes de mi curso “Introducción a las ciencias de la educación” en la Facultad de Psicología de la Universidad de Lieja. Asociados a las competencias dinámicas y metacognitivas que se desarrollan en el curso, se presentan los objetivos, los métodos y las evaluaciones respectivas.

¹¹ La palabra “objetivos” hace referencia a las intenciones pedagógicas en la formación. Estas intenciones pedagógicas, en el ámbito macrocurricular, son las competencias del perfil de egreso; mientras que, en el ámbito microcurricular, son los resultados de aprendizaje del curso.

FIGURA 8

Tres ejemplos de ausencia de triple concordancia en un currículo ficticio

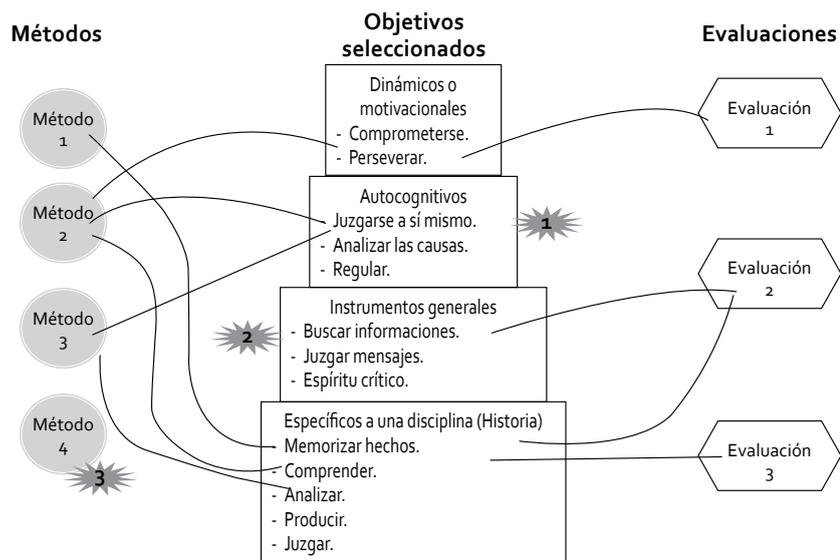


TABLA 4

Documento entregado y comentado oralmente durante la primera clase a los estudiantes del curso "Introducción a las ciencias de la educación" en la ULgll

Competencias	Métodos	Objetivos	Evaluaciones (DEA)
Dinámicas	Películas y debates (para modelización).	Que el estudiante tenga creencias epistemológicas, como por ejemplo: "Soy capaz de reflexionar", "las verdades cambian", "puedo ser parte del debate".	Huellas electrónicas de participación en foros de internet del curso.
Auto o metacognitivas	Test espectrales metacognitivos. Apuntes de autodiagnóstico.	Que el estudiante sea consciente de sus métodos de trabajo, de sus capacidades de lograr el éxito en cada curso y en el programa y comprenda por qué.	Realismo en el uso de grados de certeza. Buenos índices de confianza, imprudencia y discriminación.

Competencias	Métodos	Objetivos	Evaluaciones (DEA)
Instrumentales	Test espectrales metacognitivos con prueba de selección múltiple con soluciones generales e implícitas (ninguna, todas, faltan datos, absurdo)	Que el estudiante practique la vigilancia cognitiva.	Examen escrito a libro abierto (LA) con prueba de selección múltiple con soluciones generales e implícitas y grados de certeza.
Generales	Unas clases dadas en inglés (pero con mucha preparación).	Que el estudiante sea capaz de entender una presentación oral en inglés, después de una preparación específica.	Examen a libro cerrado (LC) sobre la traducción de vocabulario técnico anunciado (memoria).
Específicas	Ponencias orales presenciales y lectura de los capítulos del libro de base de la asignatura.	Que el estudiante comprenda el contenido de algunos elementos anunciados en una lista escrita.	Examen a libro cerrado (llenar un espacio vacío con una palabra) + grado de certeza.

Un ejemplo de triple concordancia utilizando una tabla de Excel durante la preparación del curso

Cabrera (2014, p. 200) presentó la tabla de triple concordancia (tabla 5) de su asignatura “Gestión de proyectos artísticos” en la Universidad de Chile; además, ha expresado sus métodos en términos de los “ocho eventos de aprendizaje”. La aparición de una columna vacía en la parte “Métodos” o en la parte “Evaluación” significaría una ausencia de concordancia. La presentación en Excel le facilitó la verificación de la triple concordancia a medida que desarrollaba su asignatura. Los números dentro de las celdas corresponden a una escala de importancia. Se puede ver que un objetivo puede ser desarrollado por varios métodos y evaluado por varias técnicas de evaluación.

TABLA 5

Tabla de triple concordancia de Cabrera

OBJETIVOS	Recursos	RA#1								RA#2				RA#3				RA#4				
					R	R							R	R								
		Concibe y postula un proyecto artístico.	Planifica un proyecto artístico.	Produce (escenifica) un proyecto artístico.	Conoce la estructura de formato de postulación FONDART.	Reconoce los componentes de cadenas técnicas.	Usa el formato de postulación Fondart online.	Usa Excel.	Se comunica por e-mail (grupos, etc.).	Produce Internet 2.0.	Reflexiona sobre cómo las artes contribuyen al desarrollo humano.	Reflexiona sobre su propia contribución al desarrollo humano.	Reflexiona sobre las políticas chilenas para Cultura.	Se expresa eficazmente en respuestas escritas.	Explica el concepto de desarrollo humano.	Explica las principales políticas chilenas para Cultura.	Realiza las tareas en los plazos requeridos.	Participa de forma activa en los encuentros grupales.	Colabora en la definición y distribución de tareas.	Toma en cuenta los puntos de vista de los demás y es constructivo.	Analiza desempeño de pares en planificación y producción en equipo.	Analiza su propio desempeño en la plan. y prod. en equipo.
Importancia	3	3	3	2	2	2	1	1	2	3	2	2	1	2	2	2	2	2	2	1	2	2
Imita						2	1	1	2													
Recibe	3	3		2	2					3		2		2	2							
Practica	3	3		2	2	2	1	1	2				1									
Explora		3								3		2				2	2	2	2			
Experimenta					2				2													
Crea	3	3	3							3						2	2	2	2			
Debate										3	2	2		2	2							
Metaaprende											2									1	2	2
Proyecto	3	3	3		2				2							2	2	2	2			
Portafolio										3	2	2	1	2	2						2	
Problemas	3	3		2	2	2	1	1														
Pruebas				2	2								1	2	2							2
Reporte escrito en equipo		3					1						1				2					
Reporte oral en equipo		3															2					
Tareas tecnológicas					2	1	1	2														
Evaluación de pares		3	3										1			2	2	2	2	1		
Autoevaluación		3	3									2	1			2	2	2	2		2	

Fuente: Cabrera, 2014.

PREGUNTA 7: ¿Las evaluaciones aseguran que los logros que se desean evaluar se integran en una situación compleja relacionada mediante problemas, casos, simulaciones, proyectos, portafolios, etc.?

El modelo de Miller relacionado con las evaluaciones en Medicina

Miller (1990) hizo la distinción entre cuatro contenidos de evaluación y los representó en forma de triángulo, probablemente para indicar que para hacer algo se necesita muchos saberes. En la figura 9, a la derecha de su triángulo, se hace una descripción de las evaluaciones y, en la columna siguiente, se indican los capítulos del libro *Ideas e innovaciones* en donde se detallan estos conceptos e instrumentos.

Los exámenes clínicos objetivos y estructurados o ECOE (simulaciones o juego de roles)

El principio de un ECOE es observar, a partir de casos concretos, a un estudiante simulando durante un tiempo fijo (7 minutos, por ejemplo) el papel de un profesional de los ámbitos de Medicina, Farmacia, Enfermería u otra profesión afín. A menudo, en estos casos los otros roles son representados por pacientes o clientes simulados (profesores o actores). Durante tales situaciones complejas, el estudiante debe movilizar muchos recursos cognitivos, actitudinales, relacionales, destrezas e integrarlos. A menudo, para un ECOE de Medicina o Farmacia, se reserva consultorios de hospital para esta actividad (la figura 10 grafica catorce consultorios de un corredor de un hospital que se reservaron por todo un día para este tipo de prueba).

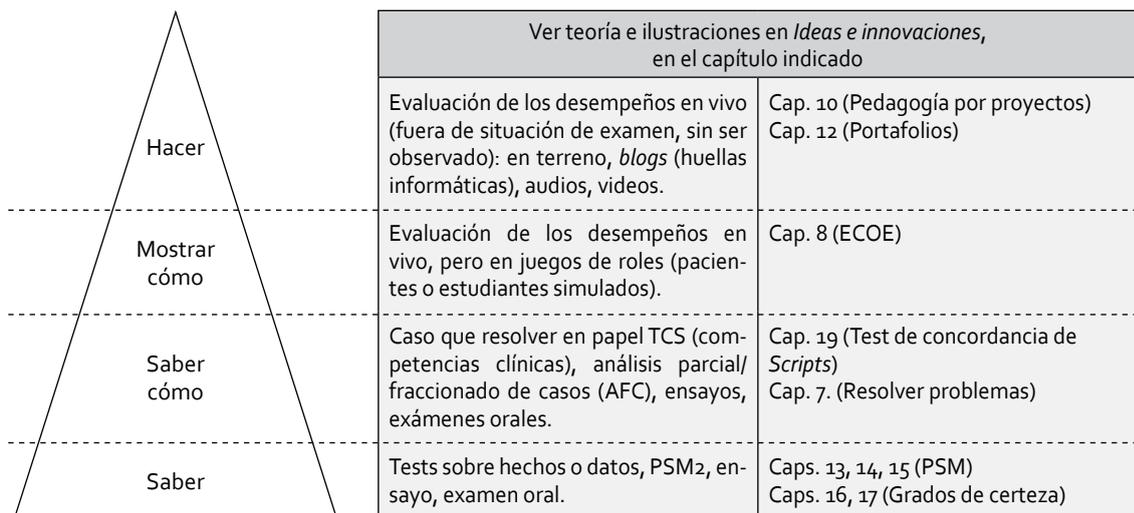
En cada consultorio (llamado “estación”), un interrogador/evaluador juega su rol y califica al estudiante según una pauta de evaluación (ver figura 11). Algunos criterios de evaluación son comunes a todas las estaciones; otros, son específicos a la estación (al caso)¹². La duración de los juegos de roles es la misma en todas las estaciones. Una campana anuncia que el tiempo ha acabado y que cada estudiante debe salir de la oficina y entrar a la siguiente estación.

La figura 11 presenta una situación típica de ECOE en Farmacia¹³, en la cual dos farmacéuticas profesionales se reparten los dos roles: evaluadora (con la pauta) y cliente (que ha aprendido de memoria su papel).

¹² Algunos ejemplos de tipos de pautas son presentados en Philippe, Leclercq & Bourguignon, 2014, pp. 156-161.

¹³ Otras ilustraciones de ECOE se encuentran en Leclercq & Poumay, 2007a, láms. 24-35.

FIGURA 9
Modelo piramidal de Miller



Nota: el nivel "Hacer" corresponde a los niveles 3 y 4 de Kirkpatrick (ver pregunta 4 de la parte 1). La ilustración de los ejemplos de resolución de problemas complejos son presentados en Leclercq & Poumay, 2007a, láms. 11-23; mientras que los de proyectos complejos en equipo, en el mismo documento (láms. 46-72).

FIGURA 10
Las catorce estaciones (tres son de descanso) de un ECOE en Medicina

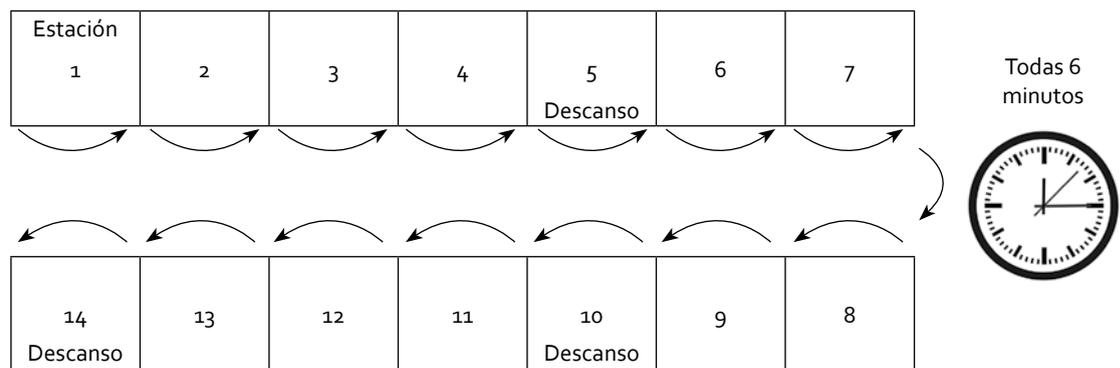


FIGURA 11

Simulación (juego de rol) en una estación ECOE en Farmacia



■ PARTE 3: EVALUACIÓN A NIVEL DEL AULA

Las interrogantes que se responden a continuación plantean propuestas en el nivel microcurricular del aula.

PREGUNTA 8: ¿Presenta el dispositivo de evaluación de aprendizajes (DEA) varios criterios de validez o calidad?

Tal como definió Cabrera en una ponencia¹⁴, un DEA constituye un grupo coherente de herramientas de evaluación que interactúan de forma coordinada para observar y medir los avances de los estudiantes y emitir juicios fundados sobre el aprendizaje producido en una unidad de trabajo, un curso, un conjunto de cursos o un programa curricular completo.

¹⁴ Véase el artículo de Álvaro Cabrera "Evaluación de competencias a nivel microcurricular: aplicación en un curso real" en este mismo libro.

Estos instrumentos de evaluación, en tanto interactúan entre sí, deben contener algunos criterios de validez o calidad. A continuación, se mencionan ocho criterios de calidad.

Los ocho criterios de validez (ETICPRAD) de un DEA ---

La tabla 6 presenta:

- Los ocho criterios de calidad, en la columna izquierda (más desarrollados en Leclercq & Cabrera, 2014, cap. 4).
- Una escala con los tres niveles posibles de calidad (insuficiente, mejorable, suficiente), en las tres columnas de la derecha.
- Unas marcas circundando el nivel de cada criterio en el contexto del análisis del DEA de un curso hipotético.

TABLA 6

Síntesis de las evaluaciones de un DEA ficticio que toma en cuenta los ocho criterios ETICPRAD

Criterios de validez examinados	Escala utilizada para evaluar estos criterios en un curso X		
E = Validez ecológica (autenticidad, proximidad con la realidad). Por ejemplo, terreno, simulaciones, juegos de roles, proyectos.	Insuf.	Mejorable	Sufic.
T = validez teórica (se apoya en una teoría sólida). Por ejemplo, basada en la arquitectura de las competencias, soluciones generales implícitas y grados de certeza.	Insuf.	Mejorable	Sufic.
I = validez informativa (diagnóstica o detallada). Por ejemplo, test espectrales metacognitivos (ver Leclercq & Cabrera, 2014, sección 10).	Insuf.	Mejorable	Sufic.
C = validez consecuencial (referida a las consecuencias de la evaluación en los actores: los estudiantes, sobre los docentes, sobre la institución). Por ejemplo, evaluaciones que refuerzan o atentan en contra de la motivación (ver pregunta 8 en este artículo).	Insuf.	Mejorable	Sufic.
P = validez predictiva (los resultados se correlacionan con otros criterios aceptados como válidos: el éxito escolar o profesional (ver pregunta 8 en este artículo y Leclercq & Detroz, 2014, p. 149-150).	Insuf.	Mejorable	Sufic.

Criterios de validez examinados	Escala utilizada para evaluar estos criterios en un curso X		
R = validez de replicabilidad (fiable: pocos errores de medición). Por ejemplo, el uso de pruebas de selección múltiple (con soluciones generales e implícitas) aumenta la replicabilidad de la calificación con el número de preguntas.	Insuf.	Mejorable	Sufic.
A = validez de aceptabilidad (es bien recibida por los estudiantes, por los Docentes, en tiempo, en costos). Por ejemplo, los test espectrales metacognitivos (ver pregunta 10 en este artículo) necesitan poco tiempo de preparación y de corrección si los datos son procesados por un <i>software ad hoc</i> (ver Leclercq & Cabrera, 2014, cap. 22).	Insuf.	Mejorable	Sufic.
D = validez deontológica (estandarización, exenta de sesgo); ver Leclercq & Cabrera, 2014, cap. 6.	Insuf.	Mejorable	Sufic.

Los términos medios

Típicamente este análisis se aplica a un conjunto de herramientas de evaluación utilizado, por ejemplo, en un curso semestral, ya que cada instrumento o herramienta de evaluación sirve para evaluar propósitos diferentes. Adicionalmente se debe considerar los recursos (materiales, humanos, de tiempo) tanto del lado de los docentes como de lado de los estudiantes. Por ello, se plantea que se deben buscar términos medios (concesiones) que permitan aplicar un DEA, sabiendo que se necesitarán años para mejorar varios aspectos de un dispositivo de evaluación de un curso o un plan de estudios. En contextos académicos, la planificación cuidadosa de una mejora puede requerir todo el año siguiente para implementarla y ajustarla.

Un ejemplo de validez consecuencial

A menudo se escucha decir: "Hay que motivar a los estudiantes". Viau (2006) señala al respecto: "Primero, hay que tener cuidado de desmotivarlos". Esta opinión se apoya en sus observaciones a través de cuestionarios aplicados a los estudiantes de dos universidades en Canadá. El autor observa que la motivación de los ingresantes era buena, pero que bajaba a lo largo de los años de formación (aunque los que habían fracasado no se encuentran en los semestres finales). ¿Cuáles son las causas de este fenómeno? Pueden originarse en los métodos o en las evaluaciones, o incluso en los objetivos propuestos¹⁵.

¹⁵ A no ser que se origine también en el esfuerzo que el estudiante realiza en el contexto de una formación universitaria y se vuelve más exigente conforme avanza en edad porque entran en competencia con otros objetivos (ganarse la vida, formar una pareja, criar niños, etc.). Leclercq y Cabrera (2014, pp. 44-45) presentan la teoría de Viau (2009)

La validez deontológica y las notas subjetivas

Investigadores franceses han desarrollado la “docimología crítica” (ciencia de los exámenes), que consiste en investigar los sesgos de calificación que resultan de la subjetividad en las calificaciones de los desempeños complejos¹⁶.

PREGUNTA 9: ¿Los componentes de las evaluaciones (condiciones, instrumentos, niveles cognitivos, requisitos) reflejan altas expectativas que corresponden a un nivel alto de exigencia?

El currículo oculto

El currículo oculto se define como “lo que nadie enseña, pero que todos aprenden”. Se manifiesta en reglas implícitas; por ejemplo: (i) toda pregunta tiene una respuesta y a toda pregunta se debe contestar de inmediato; (ii) los problemas están siempre bien planteados puesto que están siempre muy bien presentados en el aula; (iii) existen suficientes datos disponibles para resolver los problemas; y (iv) siempre existe una y solo una solución correcta, etc.

Sabemos que la vida real es diferente: los problemas están muchas veces mal planteados, ya que comúnmente en estos faltan datos y además pueden existir varias soluciones.

El déficit de medición de la comprensión

Ya en 1956, Bloom había señalado que en los EEUU se privilegiaban las preguntas que evalúan la memoria en comparación con aquellas que miden la comprensión. El tipo de preguntas que se prioriza, especialmente en las pruebas de selección múltiple clásicas (ver Leclercq & Cabrera, 2014, cap. 13), es solo una de las causas. Otra explicación es la predominancia de pruebas a libro cerrado sobre las pruebas a libro abierto. Otra característica de los test espectrales metacognitivos (ver pregunta 10) es su focalización en la comprensión a libro abierto, en donde el principio rector es que hay que comprender antes que memorizar.

en cuanto a la motivación en un contexto académico, dependiendo de tres percepciones del estudiante: del valor de la actividad académica, de sus propias competencias y de su poder de control sobre la situación. Los criterios ETIC PRAD están más ampliamente ilustrados en Leclercq, 2007, lám. 39-60.

16 Leclercq y Cabrera (2014, cap. 6) describen estos sesgos y los modos de reducirlos, en particular gracias a la utilización de rúbricas.

Las pruebas de selección múltiple (PSM) tóxicas: un ejemplo de validez consecuencial catastrófica

Decimos que las PSM tóxicas son catastróficas puesto que siempre asumen que una de las “k” soluciones presentadas es correcta. En primer lugar, las PSM clásicas acostumbran a los estudiantes a no preguntarse si es obligatorio contestar (porque la pregunta es absurda indiscreta, inmoral, etc.); en segundo lugar, los acostumbran a no buscar sus propias soluciones (ver Leclercq & Cabrera, 2014, pp. 88 y 89). La tabla 7 presenta un ejemplo de una pregunta típica con opciones múltiples (invito a responderla).

TABLA 7

Ejemplo de una prueba de selección múltiple de consigna clásica

Consigna (clásica): En las preguntas siguientes, hay siempre UNA sola solución correcta.

Pregunta:

¿Cuál es el perímetro de un triángulo de lados 2 cm, 3 cm y 6 cm?

a) 9 cm

b) 10 cm

c) 11 cm

d) 12 cm

e) 13 cm

La mayoría de quienes responden a una pregunta así, eligen como opción la respuesta c (11 cm), no dándose cuenta del hecho de que no es posible construir un triángulo con tales medidas (la suma de b y c es inferior a 6). Caemos así en las trampas a las que nos ha habituado la consigna clásica.

Las soluciones generales implícitas (SGI)

La figura 12 presenta, en la parte superior izquierda, el proceso mental reforzado por la consigna (clásica) de las PSM tóxicas y, en la parte inferior derecha, los procesos mentales esperados, con las respuestas que se pueden anunciar al inicio en las instrucciones y que hemos llamado “soluciones generales implícitas” (ver Leclercq & Cabrera, 2014, pp. 266-268).

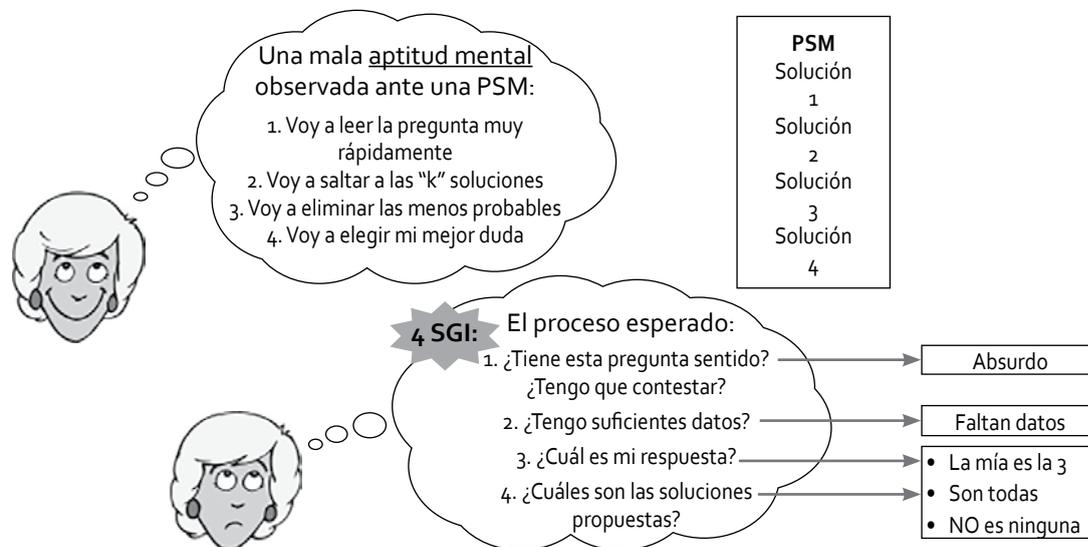
Son generales porque se aplican a todas las preguntas de una PSM y son implícitas porque no se repiten en cada pregunta, existen como posibilidad de respuesta, lo que obliga a los estudiantes a la “vigilancia cognitiva”.

Leclercq y Cabrera (2014, cap. 13) presentan un modelo llamado “SurVie”, porque destaca los varios modos de aproximarse a las preguntas según un eje bipolar horizontal que va desde un extremo —que es la sugerencia de soluciones, como en el caso de las pruebas de selección

múltiple tóxicas (en dónde la menos recomendable es la que plantea las preguntas verdadero-falso)—, pasa por el requerimiento de respuestas que exige explicar cómo contestar y va hacia el otro extremo —que es el de la “vigilancia cognitiva” y la “espontaneidad”.

FIGURA 12

El proceso mental reforzado por las pruebas de selección múltiple tóxicas y los procesos mentales reforzados por las soluciones generales e implícitas



Fuente: Leclercq & Cabrera, 2014.

PREGUNTA 10: ¿El tiempo dedicado a la metacognición es suficiente, en sus componentes de autoreflexión, autodiagnóstico, autorregulación, para entrenarla como lo ilustran los test espectrales metacognitivos (TEM)? En otras palabras, siendo un componente clave de los aprendizajes, ¿son los dispositivos de evaluación de los aprendizajes (DEA) una parte importante de los cursos del programa?

El efecto “evaluación”

Hattie (2009) ha sintetizado varios metaanálisis sobre los efectos de diversos métodos. Ha evidenciado que la evaluación con intención formativa es el método más eficiente. Veremos a continuación que hay ciertas condiciones que deben cumplirse para considerarla realmente formativa para que no se quede solo en buenas intenciones. Lamentablemente, las

investigaciones realizadas por Hattie no han medido de manera sistemática las competencias de los cuatro niveles de la arquitectura de objetivos (ver pregunta 1) y debemos ser conscientes de que la eficacia del método de un nivel (por ejemplo, el de las herramientas específicas) no significa que este sea eficaz también para los otros tres niveles de la arquitectura.

Preocuparse de los cuatro niveles de la arquitectura de las competencias, en especial las metacognitivas y las dinámicas (en particular, las epistemológicas), es un rasgo particular de los TEM.

Definición de metacognición

Se han propuesto varias definiciones de “metacognición” a partir de la primera definición dada por Flavell (1976, 1979). El desafío que hemos enfrentado es proponer una definición que sea operacional; es decir, que inspire la construcción de dispositivos concretos de formación y de evaluación. La definición de Leclercq y Poumay de las operaciones metacognitivas consiste en los siguientes elementos:

3 operaciones: auto-juzgar, auto-diagnosticar y auto-regular, aplicadas a

2 objetos: los procesos (mentales) y los productos (resultados de los procesos), en

3 momentos: antes (PRE), durante (PER) o después (POST) para

2 tipos de situaciones: aprendizaje o de desempeño (en pruebas de saberes específicos o en situaciones de la vida real) (Leclercq & Cabrera, 2014, pp. 42-43).

Los grados de certeza (probabilidad de que la respuesta sea correcta)

De nuestra definición, se sigue que los métodos de evaluación deben sistemáticamente dedicar atención al autojuicio. Eso se puede realizar con “grados de certeza” añadidos por el estudiante a cada respuesta que dé a las preguntas de una prueba y que son descritas en los capítulos 16 y 17 del libro de Leclercq y Cabrera (2014). Los grados de certeza han sido utilizados en varios ámbitos: pilotaje de aviones, enfermería, medicina de urgencia, transfusión de sangre, radioprotección, atención a pacientes diabéticos, etc.; y, por supuesto, en estudiantes de los niveles de primaria, secundaria y de universidad.

Es posible usar diferentes escalas para expresar el grado de certeza. Actualmente recomiendo el uso de seis grados:

Significado	5%	20%	40%	60%	80%	95%
Código	0	1	2	3	4	5

A continuación, veremos cómo los grados de certeza se utilizan en los TEM.

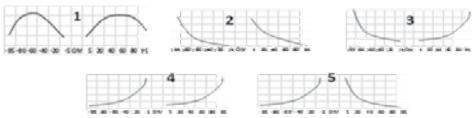
El formulario del test espectral metacognitivo

Un formulario de test espectral metacognitivo (tabla 8) tiene varias particularidades:

- En la primera y sexta columnas se presenta, respectivamente, el espectro de calidades de cada respuesta. Cuando se califique el test, se marcará el grado de certeza (consignado por el estudiante) en la primera columna si la respuesta fue incorrecta o en la sexta columna si la respuesta fue correcta.

TABLA 8

Las preguntas 22 a 24 de un test espectral metacognitivo sobre los grados de certeza

COLUMNAS																													
1						2			3	4	5	6																	
CALIDAD DE LAS RESPUESTAS INCORRECTAS						PREGUNTA			R	C	RC	CALIDAD DE LAS RESPUESTAS CORRECTAS																	
95	80	60	40	20	5	P.22. Si el índice de confianza es igual al 60%, el índice de imprudencia es igual a 1. 24% 2. 40% 3. 100%						5	20	40	60	80	95												
95	80	60	40	20	5	P.23. A continuación se muestra los índices de imprudencia y confianza de 3 estudiantes. ¿Cuál estudiante tiene la mejor configuración? <table border="1" style="margin: 10px auto;"> <thead> <tr> <th>Estudiante</th> <th>Imprudencia</th> <th>Confianza</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>46</td> <td>78</td> </tr> <tr> <td>2</td> <td>33</td> <td>76</td> </tr> <tr> <td>3</td> <td>66</td> <td>82</td> </tr> </tbody> </table>			Estudiante	Imprudencia	Confianza	1	46	78	2	33	76	3	66	82				5	20	40	60	80	95
Estudiante	Imprudencia	Confianza																											
1	46	78																											
2	33	76																											
3	66	82																											
95	80	60	40	20	5	P.24. ¿Cuál de las configuraciones espectrales siguientes es la mejor? 						5	20	40	60	80	95												
ÍNDICE DE IMPRUDENCIA						ÍNDICE DE CONFIANZA - ÍNDICE DE IMPRUDENCIA					ÍNDICE DE CONFIANZA																		
Nota antes del bono											Nota después del bono																		

- La segunda columna está destinada para la pregunta que desea evaluar el profesor. La tercera columna (**R**) está reservada para la respuesta del estudiante, la cuarta columna (**C**), para que exprese el grado de certeza de la respuesta que da, mientras que la quinta columna (**RC**) consigna la respuesta que el profesor considera correcta.
- Entre preguntas hay un espacio vacío reservado para las notas personales (apuntes) de autodiagnóstico por cada pregunta. Para tal efecto, el estudiante deberá realizar el análisis metacognitivo teniendo en cuenta dos preguntas: ¿por qué estaba tan seguro de mi respuesta aunque era incorrecta? o ¿por qué dudé de mi respuesta aunque era correcta?
- Al final del TEM existe una zona que permite al estudiante calcular el índice de confianza (sumando cada grado de certeza de las respuestas correctas y dividiendo el resultado entre el total de respuestas correctas), el índice de imprudencia (sumando cada grado de certeza de las respuestas incorrectas y dividiendo el resultado entre el total de respuestas incorrectas) y la diferencia entre el índice de confianza y el índice de imprudencia. Esta diferencia, que puede ser negativa, cero o positiva, le indicará al estudiante (y al profesor) qué tan capaz fue en diferenciar lo correcto de lo erróneo.

El proceso del test espectral metacognitivo

Este proceso, que se puede hacer a libro abierto o cerrado (Leclercq & Cabrera, 2014, cap. 9), se desarrolla en cuatro fases:

1. Los estudiantes contestan las preguntas (20 en 40 minutos, contenidas en pruebas de selección múltiple con cuatro soluciones generales e implícitas) sobre el formulario del test espectral metacognitivo.
2. Los estudiantes copian sus respuestas y certezas sobre una hoja especial que entregan al docente que, de este modo, tiene la posibilidad de analizar las respuestas antes de la comunicación de las respuestas correctas.
3. El docente organiza el debate (ilustrado en la figura 13) y, para cada pregunta, sigue el siguiente proceso:
 - Comunica la respuesta que considera correcta.
 - Invita a los estudiantes a formular objeciones, a proponer respuestas alternativas y, por supuesto, a argumentar a favor de sus puntos de vista.

- Según los casos, acepta o rechaza las propuestas de los estudiantes, también con argumentos. De este debate, a menudo resulta que el profesor indicará al programa informático que hay preguntas que tienen más de una respuesta correcta. Es el precio a pagar cuando se utiliza pruebas de selección múltiple con soluciones generales e implícitas que inducen a debate.
 - Invita a los estudiantes a rodear la posición de su respuesta sobre el eje del espectro de calidades (a la izquierda en caso de respuesta incorrecta, a la derecha en caso de respuesta correcta).
 - Invita a los estudiantes a autodiagnosticarse (reflexionar sobre) las causas de su respuesta (y redactar sus apuntes por escrito), haciéndose una de las dos preguntas clave escritas al fin del test espectral metacognitivo: "¿Por qué estaba tan seguro de mi respuesta aunque estaba equivocado?" o "¿Por qué dudé aunque mi respuesta era correcta?".
 - Pasa a la pregunta siguiente y repite el proceso.
4. Los estudiantes preparan un informe metacognitivo retrospectivo.

FIGURA 13

Ejemplo de debate en grupos de trescientos estudiantes



Las huellas metacognitivas

La figura 14 da un ejemplo de una parte de test espectral metacognitivo utilizado por un estudiante (con un formulario de TEM un poco diferente al de la tabla 8), con sus respuestas y grados de certeza en tinta azul y con la respuesta correcta en tinta roja. También se ven las posiciones rodeadas (en tinta roja) sobre el eje espectral de calidades. He pegado (en forma de llave) la pregunta que se debe plantear el estudiante.

FIGURA 14

Ejemplo de la pregunta metacognitiva (en forma de llave) que plantea un estudiante para buscar las causas de su repuesta a la pregunta 2 de la prueba

NOMBRE.....

Test Espectral Metacognitivo (Dieudonné Leclercq, setiembre de 2011, UCH, Santiago de Chile)

6 = Ninguna

7 = Todas

8 = Faltan datos

9 = Absurda

CALIDAD DE LAS RESPUESTAS INCORRECTAS						PREGUNTAS	CALIDAD DE LAS RESPUESTAS CORRECTAS					
100	80	60	40	20	0		0	20	40	60	80	100
						P.1. Un triángulo tiene una base de 5 cm y una altura de 4 cm, por tanto, tiene un área igual a: 1) 8 cm ² 2) 12 cm ² 3) 15 cm ² 4) 20 cm ²						
						6 6 80						
						P.2. ¿Cuál es el perímetro de un triángulo cuyos lados son 2, 3 y 6 cm respectivamente? 1) 10 cm 2) 11 cm 3) 12 cm 4) 36 cm						
						100 80 60 40 20 0						
						9 2 100						
						P.3. Un triángulo tiene una base de 6 cm, por tanto, tiene un área igual a: 1) 3 cm ² 2) 6 cm ² 3) 12 cm ² 4) 24 cm ²						
						8 8 60						
						100 80 60 40 20 0						
						0 20 40 60 80 100						

La figura 15 presenta el autodiagnóstico que dio el estudiante en cuanto a su respuesta a la pregunta 2 y la reflexión que se debe plantear en cuanto a su respuesta a la pregunta 3.

FIGURA 15

Ejemplo de la pregunta metacognitiva (en forma de llave) que plantea un estudiante para buscar las causas de su repuesta a la pregunta 3 de la prueba

CALIDAD DE LAS RESPUESTAS INCORRECTAS						PREGUNTAS	CALIDAD DE LAS RESPUESTAS CORRECTAS					
100	80	60	40	20	0	P.2. ¿Cuál es el perímetro de un triángulo cuyos lados son 2, 3 y 6 cm respectivamente? 1) 10 cm 2) 11 cm 3) 12 cm 4) 36 cm	0	20	40	60	80	100
						<i>Hubiera dibujado el triángulo</i>						
100	80	60	40	20	0	P.3. Un triángulo tiene una base de 6 cm, por tanto, tiene un área igual a: 1) 3 cm ² 2) 6 cm ² 3) 12 cm ² 4) 24 cm ²	0	20	40	60	80	100

El informe metacognitivo retrospectivo

Hubo un tiempo en que solía ofrecer (en realidad imponer) a mis estudiantes la oportunidad de verificar su nivel de dominio de contenidos y del realismo de su autojuicio, tres veces por semestre (uno por mes, una de cinco sesiones). Al final, les exigía (y era una parte del examen) que los estudiantes rindieran un informe retrospectivo basándose en sus apuntes metacognitivos en los tres TEM del semestre¹⁷.

Los impactos de los test espectrales metacognitivos

En la obra de Leclercq y Cabrera (2014, pp. 190-194; 2011, p. 50), se encuentra los impactos del TEM no solo a nivel cognitivo, sino también a nivel metacognitivo y epistemológico. Pienso

17 Las consignas y las partes de tal informe son presentadas en Leclercq & Cabrera, 2014, pp. 185-190.

que los beneficios importantes observados resultan de haber dedicado mucho tiempo a este tipo de evaluación: 10% del tiempo presencial de los estudiantes y un tiempo importante en casa que tuvieron que dedicar para preparar los TEM mucho antes del examen final, además de que eran sometidos a una cantidad apreciable de pruebas de selección múltiple con soluciones generales e implícitas, tal como lo requiere la triple concordancia.

Los test de progreso cubren (incluyen) todo el contenido cognitivo _____

Los contenidos de los test espectrales metacognitivos se limitan a conceptos precisos (un capítulo, los temas estudiados durante un mes, etc.). La facultad de Medicina de la Universidad de Maastricht ha desarrollado un sistema en el cual, cuatro veces por año, los estudiantes deben rendir un test (de más de 200 preguntas) sobre toda la Medicina general. Por supuesto, el test del trimestre siguiente es diferente (pero paralelo) al anterior. Como este test es el mismo para todos los estudiantes (de primer o de sexto año), se pueden evidenciar los progresos individuales y de cohorte. Hoy en día, la mayoría de las facultades de Medicina de Holanda han adoptado el principio y comparten un banco de miles de preguntas¹⁸. Este esquema es particularmente pertinente en el ámbito del “aprendizaje basado en problemas” (ABP), en el cual la enseñanza se reemplaza por el aprendizaje autónomo (Leclercq & Cabrera, 2011, pp. 20-23).

CONCLUSIÓN

La formación de seres humanos en los cuatro niveles de la arquitectura de las competencias es un desafío muy difícil de enfrentar. Tengo la convicción que para tal reto se necesita más que unos principios generales: tenemos que desarrollar métodos, procesos, instrumentos para implementar estos principios y medir rigurosamente los impactos (en los cuatro niveles) de nuestros experimentos y, a partir de eso, compartirlos con nuestros colegas del mundo entero. Es por esta razón que con Álvaro Cabrera hemos reunido, en el libro *Ideas e innovaciones*, muchos ejemplos de conceptos, instrumentos y resultados de diversas formas de evaluación, algunas de ellas muy poco conocidas.

¹⁸ Este tipo de test, que es otra aplicación del “efecto evaluación”, es descrito en Leclercq & Cabrera, 2014, pp. 397-408 y Leclercq. 2007a, láms. 44-47.

Agradezco a los organizadores del III Encuentro Internacional Universitario por haber permitido, los días del encuentro y en el futuro, encontrarme con mis colegas y poder discutir con ellos acerca de estos asuntos.

Bibliografía

- Bandura, A. (1977a). *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall.
- Bandura, A. (1977b). Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191-215.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall (edición en castellano: *Pensamiento y acción. Fundamentos sociales*. Barcelona: Martínez Roca, 1987).
- Bandura, A. (1993). Perceived Self-Efficacy in Cognitive Development and Functioning. *Educational Psychologist*, 28, 2, 117-148.
- Bandura, A. (ed.) (1995). *Self-efficacy in changing societies*. New York: Cambridge University Press.
- Bandura, A. (1997). *Self Efficacy: the exercise of control*. Englewood Cliffs, NJ: Prentice Hall.
- Barrón Tirado, C. (2015). Modelos curriculares y competencias genéricas. Algunas experiencias en la educación superior en México. En C. Díaz Villavicencio (ed.), *Las competencias genéricas en la educación superior. IIº Encuentro Internacional Universitario* (pp. 101-124). Lima: PUCP.
- Cabrera, A. (2014). Evaluar los aprendizajes en la Pedagogía Por Proyectos (PPP). En D. Leclercq & A. Cabrera, *Ideas e Innovaciones. Dispositivos de evaluación de los aprendizajes en la educación* (pp. 197-220). Santiago de Chile: Universitaria.
- Chauryie, S. (2007). Plataforma para la innovación del pregrado (conferencia en el Seminario en Evaluación de Competencias). Pucón, Chile: Universidad Católica de Temuco.

- Flavell, J. (1976). Metacognitive aspects of problem solving. En L. B. Resnick (ed.), *The nature of intelligence* (pp.231-236). Hillsdale, NJ: Erlbaum.
- Flavell, J. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34, 906-911.
- Hattie, J. (2009). *Visible learning*. London: Routledge.
- Kirkpatrick, D. (1983). *A Practical Guide for supervisory Training and Development*. Reading, Mass.: Addison-Wesley.
- Kirkpatrick, D. (2006). *Evaluating-Training-Programs-Four-Levels* (tercera edición). San Francisco: Berrett-Koehler Publ.
- Knight, P. (2000). The Outcomes-Based Education. Conférence à la chaire UNESCO en pédagogie universitaire à Louvain-La Neuve-UCL-Belgium.
- Kozanitis, A. (2015). Integración y desarrollo de competencias genéricas en Polytechnique Montreal. Ponencia del II Encuentro Internacional Universitario, llevado a cabo del 8 al 18 de setiembre en la PUCP, Lima. Disponible en: <http://cdno2.pucp.education/academico/2015/07/22235526/3_Anastassis-Kozanitis_II-Encuentro-Internacional-Universitario_PUCP.pdf>.
- Leclercq, D. (2007). Cinco problemáticas de la evaluación de las competencias: retos, definiciones, principios, funciones y validez (conferencia en el Seminario en Evaluación de Competencias). Pucón, Chile: Universidad Católica de Temuco. Disponible en: <<http://www.uct.cl/docencia/seminario/presentaciones/dleclerq1.pdf>>.
- Leclercq, D. & Cabrera, A. (2011). Conceptos y modelos para concebir, analizar y evaluar innovaciones curriculares basadas en competencias. En *Redes de colaboración para la innovación en la docencia universitaria. II Encuentro de Centros de Apoyo a la Docencia (ECAD)* (pp. 13-60). Talca, Chile: Universidad Católica del Maule. Disponible en: <<http://hdl.handle.net/2268/105423>>.
- Leclercq, D. & Cabrera, A. (2014). *Ideas e Innovaciones. Dispositivos de evaluación de los aprendizajes en la educación*. Santiago de Chile: Universitaria.

- Leclercq, D. & Detroz, P. (2003). Liens entre caractéristiques de départ (dont les résultats aux check-up) et les réussites en 1re candidature. En D. Leclercq (ed), *Diagnostic cognitif et métacognitif au seuil de l'université. Le projet MOHICAN mené par les 9 universités de la Communauté Française Wallonie-Bruxelles* (pp. 127-154). Liège: Editions de l'Université de Liège.
- Leclercq, D. & Detroz, P. (2014). Los roles de un SMART: servicio metodológico de apoyo a los docentes en la realización de tests. En D. Leclercq & A. Cabrera, *Ideas e innovaciones. Dispositivos de evaluación de los aprendizajes en la educación* (pp. 455-476). Santiago de Chile: Universitaria.
- Leclercq, D. & Poumay, M. (2007a). Concebir tareas complejas para evaluar (inferir) las competencias (conferencia en el Seminario en Evaluación de Competencias). Pucón, Chile: Universidad Católica de Temuco. Disponible en: <<https://uct.cl/docencia/seminario/presentaciones/dleclercq.pdf>>.
- Leclercq, D. & Poumay, M. (2007b). La métacognition. En D. Leclercq, *Méthodes de Formation et Théories de l'Apprentissage*. Liège: Editions de l'université de Liège.
- López, D. (2015). El caso de la Universitat Politècnica de Catalunya. Ponencia del II Encuentro Internacional Universitario, llevado a cabo del 8 al 18 de setiembre en la PUCP, Lima. Disponible en: <http://cdno2.pucp.educa-tion/academico/2015/07/22235526/2_David-Lopez_II-Encuentro-Inter-nacional-Universitario_PUCP.pdf>.
- Miller, G. A. (1956). The Magical Number Seven, Plus or Minus Two, Some Limits on Our Capacity for Processing Information. *Psychol. Review*, 63, 81-97.
- Miller, G. E. (1990). The Assessment of clinical Skills/Competence/Performance. *Academic Medicine*, 65(9), S63-S67.
- Pey, R. (2007). Proyecto Tuning América Latina (conferencia en el Seminario en Evaluación de Competencias). Pucón, Chile: Universidad Católica de Temuco. Disponible en: <<http://www.uctemuco.cl/docencia/seminario/presentaciones/rpey.pdf>>.
- Philippe, G., Leclercq, D. & Bourguignon, J.-P. (2014). ECOEs: exámenes clínicos objetivos y estructurados En D. Leclercq & A. Cabrera, *Ideas*

e innovaciones. Dispositivos de evaluación de los aprendizajes en la educación (pp. 153-170). Santiago de Chile: Universitaria.

- Prieto Navarro, L. (2001). La autoeficacia en el contexto académico: exploración bibliográfica comentada. *Miscelánea Comillas*, 59, 281-292. Disponible en: <<http://www.uky.edu/~eushe2/Pajares/prieto.PDF>>.
- Tamblyn, R. & otros (2007). Physician scores on a national clinical skills examination as predictors of complaints to medical regulatory authorities. *JAMA*, 298(9): 993-1001.
- Tardif, J. (2006). *Evaluer les compétences. Documenter le parcours de développement*. Montréal: Chênelière Education
- Viau, R. (2006). La motivation des étudiants à l'université: mieux comprendre pour mieux agir. Conférence à l'IFRES. Université de Liège.
- Viau, R. (2009). *La motivation en contexte scolaire*. Bruxelles: De Boeck.
- Villa, A. & Poblete, M. (dirs.) (2007). *Aprendizaje basado en competencias. Una propuesta para la evaluación de las competencias genéricas*. Bilbao: Mensajero.
- Villa, A. & Poblete, M. (2011). Evaluación de competencias genéricas. Principios, oportunidades y limitaciones. *Bordón*, 63(1), 147-170.
- Villardón, L. (2015). El porqué y el cómo de las competencias genéricas en educación Superior. Ponencia del II Encuentro Internacional Universitario, llevado a cabo del 8 al 18 de setiembre en la PUCP, Lima. Disponible en: <http://cdno2.pucp.education/academico/2015/07/22235526/1_Lourdes-Villard%C3%B3n_II-Encuentro-Internacional-Universitario_PUCP.pdf>.

Evaluación de competencias a nivel microcurricular: aplicación en un curso real

■ *Álvaro Cabrera Maray*

INTRODUCCIÓN

Quienes practicamos la enseñanza en universidades y otras instituciones de educación superior, estamos llamados a convertirnos en expertos en evaluación (Leclercq & Cabrera, 2014), en verdaderas autoridades en “la ciencia” (Piéron, 1963) y “el arte” (Schön, 1987) de evaluar los múltiples aprendizajes de nuestros estudiantes. Esto debido a la importancia que el “factor evaluación” tiene en la vida académica de los estudiantes, lo que se expresa por lo menos de dos formas. Primero, en el vínculo entre evaluación y aprendizaje. Por un lado, la evaluación implementada con intención formativa propone retroalimentar al estudiante, haciendo de esta retroalimentación un factor clave en su aprendizaje. Por otro, lo que implementemos como evaluación en nuestros cursos —nuestros fundamentos epistemológicos, el nivel de procesos mentales que pretendamos evaluar— tendrá la capacidad de “focalizar los esfuerzos de los estudiantes” (Leclercq & Cabrera, 2014, p. 15). Les indicará cuáles son las expectativas que tenemos de su aprendizaje, qué es y qué no es importante aprender y cuáles son las formas apropiadas de expresar lo aprendido. En segundo lugar, la evaluación es importante por las repercusiones que tiene en el avance académico de los estudiantes. En múltiples ocasiones, la evaluación se usa para generar un juicio que se traduce en una sanción, en el sentido de una definición sobre aprobación, promoción o egreso. En estos casos, que las evaluaciones sean ecuanimes y justas, que estén libres de sesgos, se transforma en una cuestión de ética profesional y personal. El peso que el “factor evaluación” tiene hace necesario que asumamos su diseño e implementación como una importante tarea académica, que demanda toda nuestra dedicación y rigor intelectual.

Este artículo aborda la evaluación de competencias a nivel microcurricular; es decir, a nivel de un curso específico. Para graficar la evaluación a este nivel, he decidido basarme, principal mas no únicamente, en una experiencia implementada en un curso real: el curso Gestión de Proyectos Artísticos que me tocó impartir el segundo semestre del año 2009 en la Universidad de Chile. Hago esto con un doble propósito. Primero, ejemplificar conceptos clave en evaluación —como triple coherencia, dispositivo de evaluación de los aprendizajes, condiciones y criterios de calidad de este dispositivo (Leclercq & Cabrera, 2014)— en un nivel que es cercano a los profesores. Al mostrar una materialización posible de estos conceptos, al nivel de un curso, busco que cada profesor y profesora que lea este artículo intente construir su propia versión de estas abstracciones a partir de su conocimiento previo (Bransford, Brown, & Cocking, 1999), en tanto docente a cargo de toda o parte de la evaluación en un curso. Segundo, mi intención es mostrar que la reflexión profunda y la sistematización de lo realizado en un curso puede dar pie a producción académica. El análisis posterior y la búsqueda de mejora pueden derivar en producción intelectual susceptible de publicación y difusión en audiencias interesadas. La investigación en la sala de clases (Angelo & Cross, 1993) es una extensión lógica del practicante reflexivo de Schön (1987) y sus resultados pueden llevar a mejorar la docencia, lo que a su vez debiese traducirse en más y mejor aprendizaje. Sin pretender resolver la tensión que existe en la vida de los académicos, entre privilegiar la investigación o la docencia (Lang, 2005; O’Meara, Terosky & Neumann, 2008), la investigación en la propia docencia abre un camino donde ambas se entrelazan y potencian.

En las siguientes secciones, abordaré brevemente el contexto en que este curso fue impartido y algunos elementos de diseño, necesarios para situar la información siguiente. El grueso del texto estará dedicado al “dispositivo de evaluación de los aprendizajes” (DEA) del curso, desarrollando en detalle tres de las herramientas de evaluación que fueron utilizadas. Finalmente, analizaré el DEA del curso utilizando el modelo ETICPRAD de criterios de calidad (Leclercq, 2014a), para responder preguntas acerca de la validez de las inferencias que es posible hacer a partir de la información entregada por las evaluaciones. Terminaré con una discusión sobre algunos de los aprendizajes extraídos de esta implementación del curso.

CONTEXTO

El curso Gestión de Proyectos Artísticos es un curso transversal a las nueve carreras de la Facultad de Artes de la Universidad de Chile. El segundo semestre del año 2009 dicté una de sus secciones, compuesta por un pequeño grupo de nueve estudiantes de Licenciatura en Artes Visuales.

La estrategia pedagógica elegida para el curso fue la conocida como “aprendizaje basado en proyectos” (ABP) o “pedagogía por proyectos” (PPP). La PPP (Kilpatrick, 1918) “provee a los estudiantes una oportunidad concreta de construir, colectivamente, un conocimiento compartido en una situación contextualizada” (Cabrera, 2014, p. 197) y supone que los estudiantes desarrollen habilidades y aprendan conceptos en los tiempos apropiados según las necesidades del proyecto, de acuerdo al principio pedagógico de aprendizaje situado o contextualizado (Anderson, Reder & Simon, 1996; Greeno, 1997; Lave & Wenger, 1991). Para que esto ocurra —para que los estudiantes realicen un proyecto—, como profesores debemos entregar gran parte del control de esta actividad a ellos, para aumentar potencialmente su motivación. La iniciativa recae en ellos y también las expectativas sobre el proceso y producto final, enviando un mensaje de confianza sobre sus capacidades (Cabrera, 2014; Leclercq & Cabrera, 2011). En el curso que describo, el proyecto estudiantil consistió en organizar y llevar a escena una actividad artística con público. En términos de competencias a fomentar, este curso estaba diseñado para contribuir con el desarrollo de tres competencias genéricas de la Universidad de Chile que, junto a otras, se espera estén presentes en los egresados de todas las carreras de la institución: ser socialmente responsable y comprometido con la ciudadanía, trabajar en equipo y ser autocrítico. Además, el curso era una de las actividades académicas que debían contribuir al desarrollo de la competencia “gestionar proyectos artísticos”, transversal a todas las carreras de la Facultad de Artes.

DISEÑO DEL CURSO

Además de las competencias genéricas a cuyo desarrollo debía contribuir, el curso poseía ciertos contenidos que debían ser tratados, ciertos saberes que debían ser dominados por los estudiantes. La definición de competencia que comparto con muchos autores (por ejemplo, Beckers, 2002; Leclercq & Pourmay, 2008; Scallon, 2004; Tardif, 2006) es la que expresáramos con Leclercq como una “actuación eficaz en situaciones complejas, movilizando recursos internos (saberes, actitudes o destrezas) y externos” (como redes y otros recursos) (Leclercq & Cabrera, 2011, p. 13). Justamente, parte de esos “recursos internos” a ser movilizadas son saberes y destrezas que constituyen la “materia” del curso. En este sentido, no hay competencia sin contenidos; pues la actuación compleja no es una repetición automatizada, sino una actuación que demanda un repertorio de saberes a los cuales recurrir. Que ciertos contenidos sean necesarios para el despliegue de la competencia implica que los profesores somos responsables de impartir estos saberes, de poner a los estudiantes en contacto con

fuentes donde estén presentes y de evaluar que hayan sido realmente incorporados (Leclercq & Cabrera, 2014).

RESULTADOS DE APRENDIZAJE

Estos fueron creados al “filtrar” las competencias a través de los contenidos del curso, al traducir las competencias al contexto que establece la “materia” del mismo. Este proceso de interpretación dio origen a cuatro resultados de aprendizaje (RA) que orientaron el diseño del curso:

- RA1: producir una actividad artística con público para vivenciar las lógicas y necesidades de la producción en Chile.
- RA2: expresar por escrito sus propias reflexiones y opiniones sobre la relación entre las expresiones artísticas y el concepto de “desarrollo humano”, para enriquecer su respuesta frente a la pregunta: ¿para qué producir proyectos artísticos?
- RA3: trabajar en equipo para organizar y llevar a escena un proyecto artístico con público en vivo.
- RA4: estimar la seguridad en sus propias respuestas, analizar su confianza y prudencia (o imprudencia), para fomentar la metacognición y la reflexión sobre sus propios procesos de aprendizaje.

Comparto las visiones que proponen a las competencias como orientadoras de un programa completo y a los resultados de aprendizaje como orientadores de los cursos individuales. Existen pocas oportunidades, en el curso de un programa de formación, para observar la competencia de egreso; es decir, observar al estudiante en plena actuación en una situación cercana a la realidad (profesional/académica) y que le demande la movilización integrada de amplios recursos. Fuera de situaciones cercanas al egreso —como, por ejemplo internados, pasantías, prácticas, o proyectos—, es difícil observar directamente al estudiante desplegando la competencia. Lo que hacemos en el resto de las actividades de aprendizaje es inferir que la competencia está siendo desarrollada a partir del logro de ciertos resultados de aprendizaje (RA), que han sido a su vez creados como una traducción de la competencia a los contenidos específicos del curso. Los cuatro RA fueron luego expresados (desagregados, divididos) en indicadores y recursos: los primeros son desempeños observables —incluso medibles que utilizan distintos instrumentos de evaluación—, mientras que los segundos son saberes fundamentales, necesarios para la competencia. Indicadores y recursos en conjunto describen cada resultado de aprendizaje. Este proceso originó 22 indicadores o recursos, que en conjunto describían a los cuatro RA. Desde la

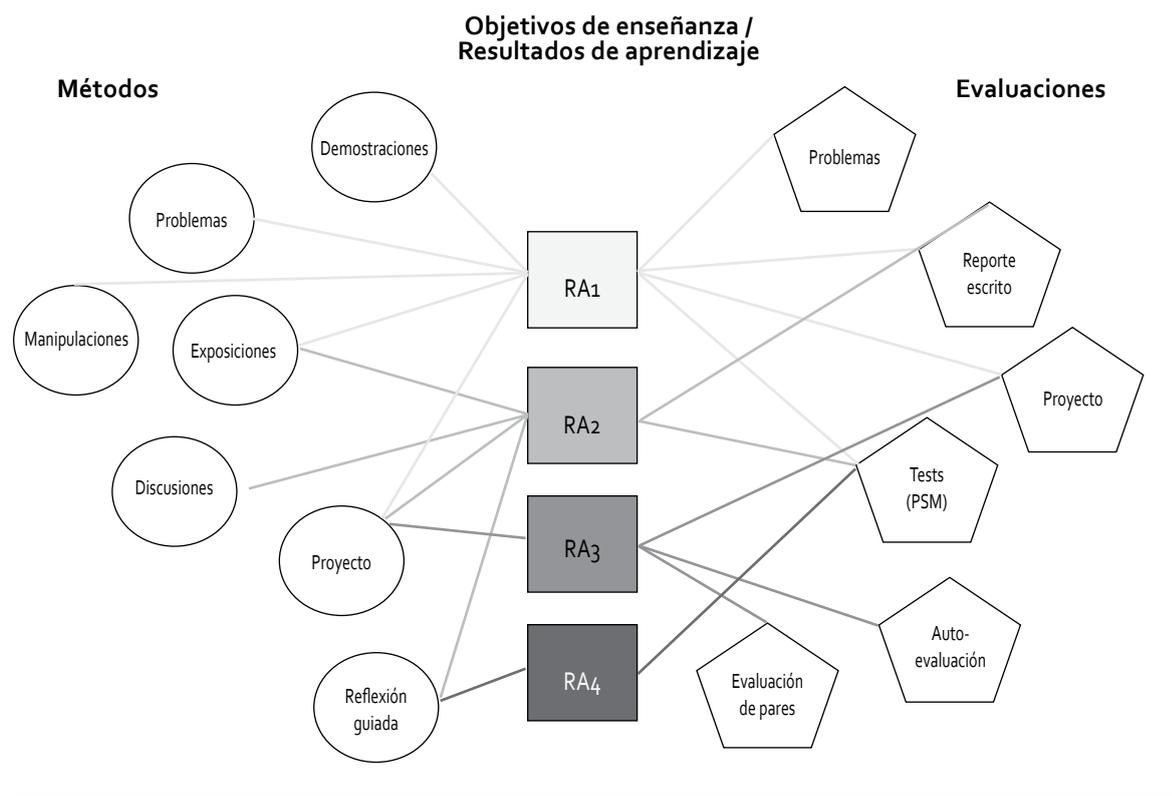
perspectiva de la evaluación, los indicadores nos muestran qué evidencia recoger para decidir si el RA ha sido o no logrado.

TRIPLE CONCORDANCIA

Propuesto por Leclercq (1995) —debido a que el término “alineamiento” (*alignment*) no lograba expresar la complejidad de las relaciones existentes—, el concepto de “triple concordancia” plantea verificar que todos los métodos empleados fomenten y que todas las evaluaciones evalúen por lo menos uno de los resultados de aprendizaje (RA). Una forma de verificarlo es mediante una representación, como en la figura 1. En esta, los RA están representados en la

FIGURA 1

Triple concordancia en el curso Gestión de Proyectos Artísticos



Fuente: elaboración propia.

columna central. A la izquierda se muestran las principales estrategias metodológicas utilizadas y a la derecha se ven los instrumentos de evaluación que reunieron las evidencias necesarias para decir que los aprendizajes se lograron. La figura 1 también muestra las conexiones entre estos componentes, indicando métodos que contribuyen al desarrollo y evaluaciones que miden el logro de determinados RA. Lo que buscamos es que no exista ningún vacío; es decir, que no haya: a) RA que no estén siendo desarrollados por ninguna metodología o no estén siendo evaluados por ningún instrumento; y b) metodologías que no estén contribuyendo al desarrollo de ningún RA o instrumentos que no estén evaluando ningún RA. Este chequeo de consistencia también puede ser intentado al nivel de indicadores y recursos. Sin embargo, la representación de la figura 1 se hace insuficiente y debemos recurrir al formato que llamamos “radiografía de la triple concordancia del curso” (Cabrera, 2014; Leclercq & Cabrera, 2011) que se muestra en la tabla 1: en la mitad superior se ven los indicadores y recursos; la mitad inferior está dividida entre metodologías y evaluaciones; los símbolos en las celdas indican que en el diseño existe relación entre los componentes; el desafío de lograr que no existan vacíos se expresa en que no hayan columnas vacías (indicadores que ningún método fomenta o ningún instrumento evalúa), ni filas vacías (métodos o evaluaciones sin conexión con ningún indicador o recurso).

ESTRATEGIAS METODOLÓGICAS

En este curso, me propuse tener una gran variedad de metodologías. Para lograrlo, utilicé el Modelo de 8 Eventos de Aprendizaje (Leclercq & Poumay, 2005). El modelo nos propone ocho tipos de situaciones en las cuales una persona (un estudiante) puede aprender y ocho clases de relación estudiante-educador en que el aprendizaje puede ser facilitado, construido, afianzado. Así, el estudiante puede aprender cuando imita, cuando observa y repite una actuación, mientras que el profesor —o alguien provisto por él— sirve de modelo a imitar. El estudiante también aprende cuando recibe, cuando escucha, cuando lee, cuando ve un video y el rol del profesor es ser la fuente que transmite. El estudiante también aprende cuando practica, cuando realiza una tarea que requiere repetición (como dominar un instrumento musical) y el rol del profesor es guiar, retroalimentar, corregir y motivar la práctica. El estudiante también aprende cuando explora, cuando busca en distintas fuentes y sintetiza, y la tarea del profesor es proveer un conjunto de fuentes relevantes, significativas, que permitan una búsqueda sustantiva. El estudiante también aprende cuando experimenta, cuando puede manipular un micromundo que reacciona y lo retroalimenta, y el rol del profesor es proveer ese micromundo y asegurar su reactividad. El estudiante también aprende cuando crea, cuando genera algo nuevo, y la tarea del profesor

TABLA 1

Radiografía del curso Gestión de Proyectos Artísticos

	Recursos	RA#1								RA#2								RA#3				RA#4				
						R	R									R	R									
	Concibe y postula un proyecto artístico.																									
	Planifica un proyecto artístico.																									
	Produce (escenifica) un proyecto artístico.																									
	Conoce la estructura de formato de postulación FONDART.																									
	Reconoce los componentes de cadenas técnicas.																									
	Usa el formato de postulación FONDART online.																									
	Usa Excel.																									
	Se comunica por e-mail (grupos, etc.).																									
	Produce Internet 2.0.																									
	Reflexiona sobre cómo las artes contribuyen al desarrollo humano.																									
	Reflexiona sobre su propia contribución al desarrollo humano.																									
	Reflexiona sobre las políticas chilenas para Cultura.																									
	Se expresa eficazmente en respuestas escritas.																									
	Explica el concepto de desarrollo humano.																									
	Explica las principales políticas chilenas para Cultura.																									
	Realiza las tareas en los plazos requeridos.																									
	Participa de forma activa en los encuentros grupales.																									
	Colabora en la definición y distribución de tareas.																									
	Toma en cuenta los puntos de vista de los demás y es constructivo.																									
	Analiza desempeño de pares en planificación y producción en equipo.																									
	Analiza su propio desempeño en la planificación y prod. en equipo.																									
	Evalúa su grado de seguridad en sus propias respuestas.																									
MÉTODOS	Imita						X	X	X	X																
	Recibe	X	X		X	X					X		X		X	X										
	Practica	X	X		X	X	X	X	X	X			X													
	Explora		X								X		X				X	X	X	X						
	Experimenta					X					X															
	Crea	X	X	X							X						X	X	X	X						
	Debate										X	X	X		X	X										
	Metaaprende											X										X	X	X		
EVALUACIONES	Proyecto	X	X	X		X				X							X	X	X	X						
	Portafolio										X	X	X	X	X	X								X		
	Problemas	X	X		X	X	X	X	X																	
	Pruebas				X	X							X	X	X										X	
	Reporte escrito		X					X						X					X							
	Reporte oral		X															X								
	Tareas tecnológicas						X	X	X	X																
	Evaluación de pares		X	X										X			X	X	X	X	X					
Autoevaluación		X	X									X	X			X	X	X	X			X				

Fuente: elaboración propia.

es escuchar, estar atento, detectar las necesidades del estudiante en su proceso de creación y proveer los apoyos justos. El estudiante también aprende cuando debate, cuando discute con otros, y el rol del profesor es organizar y animar ese debate. Finalmente, el estudiante aprende cuando reflexiona acerca de su propio aprendizaje (metarreflexión), de sus procesos, sus logros y sus luchas; en este caso, la tarea del profesor es acompañar esta reflexión, reflexionar con el estudiante. El Modelo de 8 Eventos de Aprendizaje nos ubica un nivel por encima de la discusión sobre estrategias metodológicas específicas (como el aprendizaje basado en problemas, o en proyectos, o la clase magistral, con todos sus promotores y detractores) y nos permite ver que todos estos tipos de situaciones son válidas para lograr el aprendizaje de los estudiantes. Es tarea de cada profesor definir cuáles son aquellos eventos más fecundos a la hora de desarrollar los aprendizajes esperados en sus propios cursos.

DISPOSITIVO DE EVALUACIÓN DE LOS APRENDIZAJES

El resto de este artículo estará dedicado al componente “evaluación” del curso. Leclercq y Cabrera (2014) propusieron el concepto de “dispositivo de evaluación de los aprendizajes” (DEA) como una forma de pensar sistemáticamente este aspecto del diseño del curso. Los autores definieron un DEA como un grupo coherente de herramientas de evaluación que interaccionan de forma coordinada para observar y medir los avances de los estudiantes y emitir juicios fundados sobre el aprendizaje producido, en una unidad de trabajo, un curso, un conjunto de cursos o un programa curricular completo (Leclercq & Cabrera, 2014, p. 26).

Leclercq (2014b) define el DEA a partir de sus componentes. Primero, las finalidades, el para qué vamos a evaluar. ¿Vamos a evaluar para certificar o sancionar algo? (no en un sentido punitivo, sino para expresar que la evaluación tiene una determinada sanción —una calificación—, que a su vez tiene consecuencias), ¿o vamos a evaluar para formar y mejorar? (para estimar el nivel de avance de los estudiantes y entregarles retroalimentación; o para evaluar la calidad de la enseñanza), ¿o ambas? En función de estas finalidades, que generalmente son múltiples, el DEA permite elaborar síntesis de observaciones (como una nota final, por ejemplo, que sintetiza el conjunto de evaluaciones/observaciones en un curso), construidas sobre la base de hipótesis lógicas (puntajes, promedios, ponderaciones, estándares), que utilizamos para tomar decisiones (como aprobar o reprobar un curso, u obtener o no una beca). Estas síntesis buscan expresar la evaluación de distintos objetos —en mi caso, resultados de aprendizaje e indicadores—, cuyo logro permite inferir que los estudiantes están desarrollando las competencias asociadas. Para

realizar observaciones que permitan llegar a las síntesis que se busca, se debe definir las condiciones de la evaluación, lo que Leclercq (2014c) llama el prisma de las características de un DEA. Para esto debemos tomar múltiples decisiones, sobre qué evaluar, cuándo evaluar, quién cumplirá el rol de evaluador, etc., y sobre los instrumentos de evaluación a usar. Finalmente, el último componente de un DEA son ocho criterios de calidad contra los que contrastaremos los instrumentos y características del DEA del curso. Estos criterios expresan las dimensiones de validez del DEA —por ejemplo, validez ecológica, teórica, informativa y deontológica— y serán discutidos hacia el final de este artículo.

La tabla 2 (Leclercq, 2014c, p. 53) muestra, en una visión panorámica, el conjunto de decisiones que debemos tomar en uno de los componentes del DEA: el prisma de características. En la parte superior, vemos las facetas prioritarias: las funciones, el tipo de referencias y la precisión. Las funciones pueden ser formativa o sancionadora, lo que fue discutido más arriba. Las referencias pueden ser criteriosales o normativas. En el primer tipo, el desempeño de los estudiantes es contrastado con un criterio fijo —un estándar que marca la condición de logro— y cada estudiante es medido de acuerdo a esos criterios sin importar el rendimiento del resto del grupo. En el segundo tipo de referencias (normativa), el rendimiento de un grupo es ordenado de acuerdo a una distribución normal (una curva con forma de campana) y se establece un punto de corte que define el éxito. La posición de un estudiante en la curva sí es afectada por el rendimiento del grupo. En educación, nuestra posición es que privilegiemos evaluaciones que usen referencias criteriosales, con umbrales que representen las condiciones de logro y que nos esforcemos para que ojalá todos nuestros estudiantes alcancen esos niveles. La precisión de la evaluación puede ser sumativa o diagnóstica. La primera implica que la evaluación nos entrega solo una imagen de la suma de todas las observaciones/evaluaciones (lo que es necesario para ciertas finalidades). La precisión diagnóstica nos entrega una gran cantidad de información, lo que permite evaluar de mejor forma los aprendizajes de los estudiantes y la enseñanza desplegada por el profesor. La tabla 2 muestra el resto de decisiones que es necesario tomar: entre otras, qué vamos a evaluar (¿productos o procesos?), cuáles dimensiones de la respuesta (¿solo la exactitud o la exactitud más algo (rapidez, concisión, certeza?)) y quiénes van a evaluar (¿solo el profesor, o además otros expertos (jurado), o los propios estudiantes en evaluación de pares y autoevaluación?).

TABLA 2

Plantilla para describir las características de un DEA

		Componentes del DEA
A. Las 3 facetas prioritarias		
A1. Función (PARA QUÉ)	Sancionadora	
	Formativa	
A2. Referencias	Criteriales (absolutas)	
	Normativas (relativas)	
A3. Precisión	Sumativa	
	Diagnóstica	
B. Objeto (QUÉ). B1. Foco	Productos	
	Procesos	
B2. Dimensiones	Exactitud + Concisión	
	Exactitud + Certeza	
	Exactitud + Rapidez	
	Exactitud +	
C. Sujetos (QUIÉN)	Individuos	
	Grupos/equipos	
D. Destinatarios (PARA QUIÉN)	Confidencial	
	Intermedia	
	Pública	
E. Agentes (POR QUIÉN) y en cuál ámbito (360°)	E1. Ámbito escolar docente/experto/ipso/pares	
	E2. Ámbito profesional	
	E3. Ámbito Intermedio	
E4. en una relación	Arbitraria	
	Transaccional (confrontada)	
F. Tiempo (CUÁNDO) periodicidad	F1. Continua	
	F2. Puntual	
	F3. Única	
	F4. Repetida	
	F5. Mejorable	
G. Estandarización (CÓMO)	G1. De la corrección subjetiva / objetiva	
	G2. Del contexto: natural vs. estandarizado	
	G3. Adaptativo según contenido, exigencias, desempeños anteriores, lugar, momento.	
H. Previsibilidad de los criterios	H1. Anunciados	
	H2. Escondidos	
	H3. Intermedios	

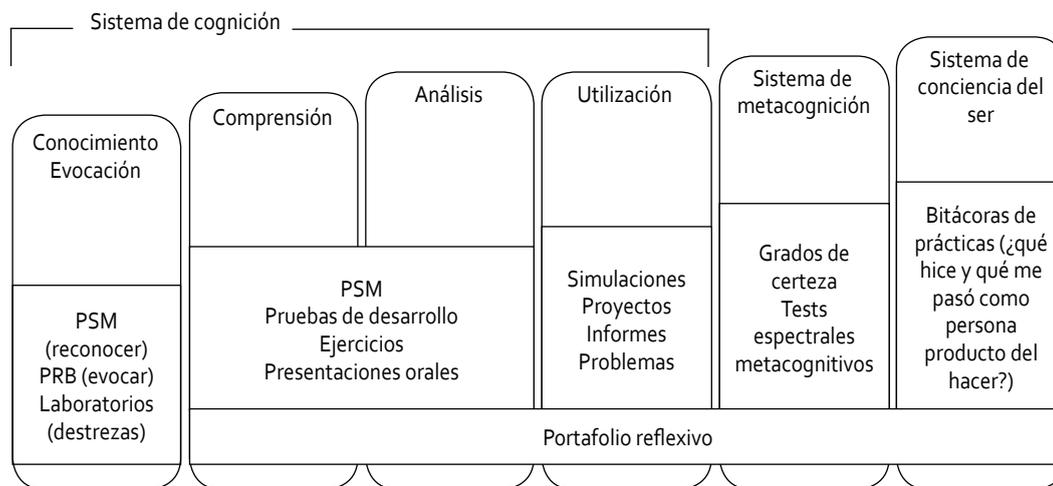
Fuente: Leclercq, 2014c, p. 53.

INSTRUMENTOS DE EVALUACIÓN

El último grupo de decisiones que debemos tomar se relaciona con los instrumentos de evaluación que vamos a incluir en nuestro DEA. Proponemos el uso de una taxonomía de objetivos educacionales como marco orientador para la selección de estos instrumentos. La figura 2 muestra la propuesta de calce entre la taxonomía de objetivos educacionales planteado por Marzano (2001) y diversos instrumentos de evaluación que pueden ser usados para medir aprendizajes en los distintos niveles de la taxonomía. La taxonomía de Marzano (2001) plantea un sistema de cognición con cuatro niveles (conocimiento/evocación, comprensión, análisis y utilización), un sistema de metacognición y un sistema de conciencia del ser, donde el autor ubica la evaluación de emociones, eficacia y motivación. Sobre estos sistemas y sus niveles, superponemos los instrumentos de evaluación que consideramos pueden entregar evidencias de logros para objetivos educacionales en el respectivo nivel. Así, el proceso mental de conocimiento/evocación puede ser evaluado por cierto tipo de preguntas de selección múltiple (PSM), que demanden reconocer la respuesta correcta entre varias, o por preguntas de respuesta breve (PRB), que requieren la evocación de una información. Este nivel también podría ser evaluado en situaciones de laboratorio o clínicas donde se entrenan destrezas o desempeños sencillos. Los niveles dos y tres, de comprensión y análisis, pueden ser evaluados con otro tipo de PSM —por ejemplo, con viñetas o casos que demanden procesos mentales por sobre el mero reconocimiento de la respuesta correcta—. Evidencias de comprensión y análisis también pueden ser recogidas a través de presentaciones orales, pruebas de desarrollo y ejercicios complejos. El nivel de utilización, de acuerdo a cómo lo concibe Marzano (2001) —con toma de decisiones, resolución de problemas, experimentación e investigación—, demanda ser evaluado mediante instrumentos como simulaciones, proyectos, problemas complejos e informes en formatos cercanos a la realidad profesional/académica —por ejemplo, un artículo de investigación escrito como si fuera a ser enviado a revisión para publicación—. Algunos aspectos del sistema de metacognición pueden ser evaluados usando grados de certeza en las respuestas dadas en las pruebas (Leclercq, 2014d) y la estrategia que Leclercq (2014e) denomina “tests espectrales metacognitivos”. Para evaluar transformaciones al nivel del sistema de conciencia del ser, proponemos un instrumento del tipo bitácoras de práctica, en donde el estudiante registre no solo lo hecho durante la práctica, sino también los cambios que experimentó como persona producto de ese hacer. Por último, un instrumento de evaluación como el portafolio reflexivo (Poumay & Maillart, 2014) permitiría reunir información proveniente de evaluaciones en todos los niveles de la taxonomía, junto a las reflexiones del estudiante acerca de lo que estas evidencias pueden significar en términos de su aprendizaje.

FIGURA 2

Instrumentos de evaluación ordenados de acuerdo a la taxonomía de objetivos educacionales de Marzano (2001)



Fuente: Marzano, 2001; elaboración propia.

Hasta aquí hemos recorrido un camino marcado por la gran cantidad de decisiones que debemos tomar al diseñar e implementar el componente de evaluación de un curso. Luego de leer las secciones anteriores, es fácil abrumarse al momento de decidir frente a la gran variedad de elementos a ponderar. Mi sugerencia es no asustarnos, pues estas son decisiones que de todos modos ya tomamos. El desafío del profesor universitario como practicante reflexivo (Schön, 1987) es hacer que todas estas decisiones sean explícitas y conscientes, y tener justificación para esas decisiones. Al mismo tiempo, es importante no engañarnos y pensar que evaluar los aprendizajes de nuestros estudiantes, sobre todo cuando esos aprendizajes son desempeños complejos, es una tarea sencilla. No lo es y pensar que es más simple de lo que es no nos ayudará a abordarla en toda su complejidad. Por último, incluso en medio del cúmulo de decisiones que se deben tomar, sugiero siempre recordar las que considero tres preguntas claves en la evaluación (adaptadas de Bain, 2004):

- ¿Cuáles son los cambios intelectuales y personales que pretendo que ocurran en mis estudiantes producto de mi curso, como contribución a su formación global/integral?

- b) ¿Cuáles son las informaciones que consideraré como evidencia de que esas transformaciones están ocurriendo?
- c) ¿Cuáles son los instrumentos de evaluación (que domino/conozco, tengo la posibilidad de implementar) que mejor me proveen ese tipo de informaciones?

La evaluación es la recolección de evidencias que me permiten decir (inferir) que los aprendizajes están ocurriendo.

DEA DEL CURSO

El DEA que resultó producto de las decisiones tomadas se ve en la figura 3, representado en forma tradicional. El principal instrumento de evaluación es el proyecto, con 30% de la nota final dependiente de la puesta en escena de la actividad artística. Además incluye pruebas de selección múltiple (PSM), resolución de problemas, evaluación de pares y auto-evaluación. El 60% de la calificación final del curso proviene de evaluaciones grupales y 40% de evaluaciones individuales.

FIGURA 3

DEA del curso *Gestión de Proyectos Artísticos (representación 1)*

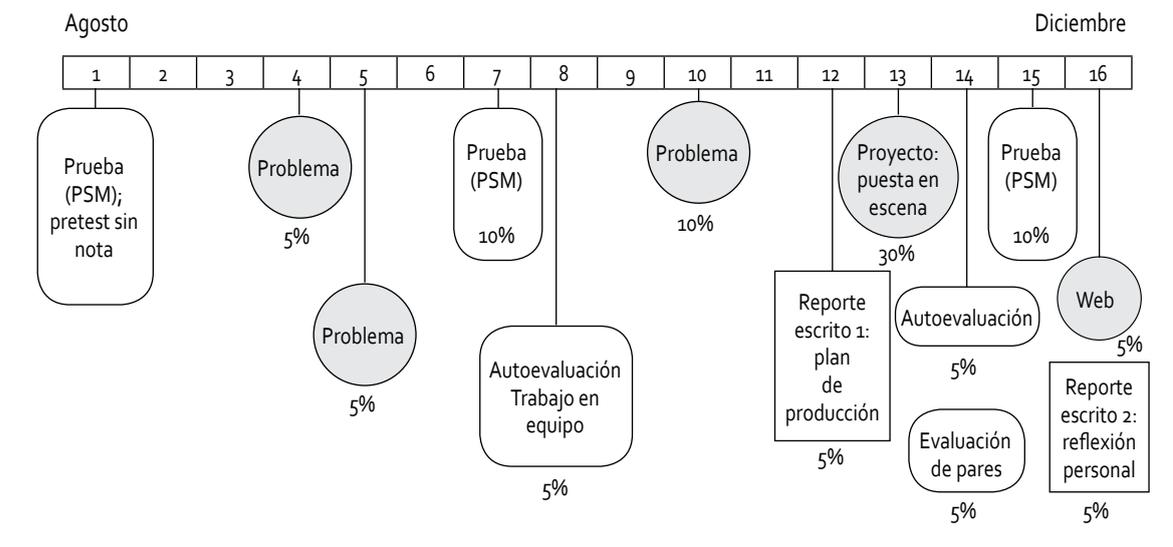
Evaluaciones del curso	
Proyecto (puesta en escena + registro web) 35%
Reporte escrito 10%
Pruebas (PSM) 20%
Resolución de problemas 20%
Autoevaluación del trabajo en equipo 10%
Evaluación de pares del trabajo en equipo 5%

Fuente: elaboración propia.

La figura 4 muestra el mismo DEA del curso, agregando la dimensión temporal, pues le indica al estudiante la semana del semestre en que ocurrirá cada evaluación programada. En las siguientes secciones abordaremos tres instrumentos de evaluación de los utilizados en este DEA, describiremos su implementación y analizaremos sus fortalezas y debilidades. Como propuesta para superar estas últimas mostraremos ejemplos de evaluaciones que otros académicos han implementado en sus cursos.

FIGURA 4

IDEA del curso *Gestión de Proyectos Artísticos* (representación 2)



Fuente: elaboración propia.

■ INSTRUMENTO 1. PRUEBAS CON PREGUNTAS DE SELECCIÓN MÚLTIPLE (PSM)

Confeccionamos una prueba de 28 PSM para evaluar que los estudiantes poseyeran ciertos recursos (saberes) considerados necesarios para el logro de tres de los cuatro resultados de aprendizaje (RA) del curso (RA 1, 2 y 4). La prueba fue aplicada completa como pretest (sin calificación) y dividida en dos ocasiones posteriores como postest (con calificación). Utilizamos la que se conoce como “consigna simple o clásica” (Leclercq & Cabrera, 2014, p. 281), donde se le indica al estudiante que dentro de las “k” soluciones hay una correcta y debe seleccionar solo una. La consigna clásica presenta varias debilidades que hemos detallado en otros trabajos (Leclercq & Cabrera, 2014), como adormecer la vigilancia cognitiva al simplificar los problemas hasta un nivel en que existe una única respuesta correcta. En términos del currículum oculto, se transmite, por ejemplo, la noción de que las preguntas están siempre bien planteadas, que siempre hay una (única) respuesta correcta o que siempre se deben contestar. Una alternativa más auténtica, más cercana a la realidad profesional, es la propuesta por D’Ivernois y Chabot (2003) en sus preguntas en forma de “cascada divergente”. La figura 5 muestra un ejemplo de estas preguntas —desarrollado en un taller para profesores que impartimos el año 2013 en la Facultad de Medicina de la

Universidad de Chile—. Para ser usadas en evaluaciones anteriores, los profesores habían creado cuatro PSM que debían ser contestadas a partir de una viñeta clínica. La información que obtenían eran cuatro respuestas, correctas o incorrectas, y el comportamiento de los estudiantes frente a los distractores (en el caso que algún distractor atrajera muchas de las respuestas de los estudiantes, por ejemplo). Utilizando el mismo caso ya elaborado, los profesores lograron transformarlo en la pregunta tipo cascada divergente que muestra la figura 5. La consigna pide al estudiante escribir “Sí” o “No” en los paréntesis junto a una serie de aseveraciones que se hacen dentro del caso. La información resultante son 12 respuestas. Estas evidencian lo que el estudiante entiende no solo respecto a lo que antes eran las alternativas correctas, sino también respecto a muchos de los que eran distractores. Aumenta así la precisión diagnóstica de la evaluación. La corrección puede hacerse sumando puntos positivos por respuestas correctas, no sumando puntaje en el caso de errores que no afectan la salud del paciente descrito en el caso y restando puntos cuando se trata de errores (respuestas incorrectas) que afectan la condición del paciente. Puede incluso existir una escala de gravedad de los errores. Sin dejar de ser una evaluación en formato “lápiz y papel”, aumenta su autenticidad, pues refleja más cercanamente problemas que el estudiante podría

FIGURA 5

Ejemplo de pregunta tipo “cascada divergente”

CONSIDERE EL CASO CLINICO Y CONTESTE SI O NO A LAS 12 PREGUNTAS:

Don Julio ingresa a la unidad de cirugía operado de colecistectomía, con vía venosa permeable, sonda nasogástrica, sonda Foley, un drenaje y requerimientos de oxígeno.

Por indicación médica se le administra suero fisiológico 1500 cc a goteo lento, Profenid EV cada 8 hrs en bolo de 100 cc de suero fisiológico, durante las primeras horas presenta emesis cuantificada en 200 cc y diuresis por sonda Foley de 600 cc.

En el caso clínico anterior, son ingresos:

Emesis (1_), Suero fisiológico (2_), Diuresis (3_), Drenaje (4_), Profenid (5_).

Al medir el contenido del drenaje se obtienen 180 cc. Por lo tanto, el resultado del balance hídrico es -80 ml (6_), lo que implica un balance hídrico equilibrado (7_).

Con el resultado obtenido ¿Concluiría usted que Don Julio se encuentra en riesgo de Hipervolemia (8_) y debe avisar al médico residente (9_)? ¿O concluiría que se encuentra bien, pero en observación (10_)?

Al controlar posteriormente los signos vitales de Don Julio ve que su temperatura corporal es de 39°C. Por lo tanto, al hacer el balance hídrico usted debe considerar 12 ml menos de agua metabólica (11_) y 12 ml más de pérdidas insensibles (12_).

Fuente: elaboración propia. Pregunta desarrollada en el taller impartido en el año 2013 en la Facultad de Medicina de la Universidad de Chile.

encontrar en su futura realidad profesional. En términos del modelo ETICPRAD de criterios de calidad de un DEA (Leclercq, 2014a), que será explicado y aplicado más adelante, diríamos que esta alternativa de transformación de PSM mejora la validez informativa y ecológica de la evaluación.

Grados de certeza

Leclercq (2014d) declara compartir el siguiente postulado de De Finetti: “El conocimiento parcial existe. Detectarlo es necesario y factible” (De Finetti, 1965, p. 109; citado en Leclercq, 2014d, p. 331). Leclercq plantea que el aprendizaje consiste en pasar de estados de conocimiento parcial a otros de conocimiento menos parcial, pero rara vez de la ignorancia total al conocimiento perfecto, ambos “casos particulares (y extremos) del conocimiento parcial” (2014d, p. 331). Sin embargo, la evaluación tradicional considera al estudiante como en uno de esos dos extremos: o bien lo obliga a contestar sin expresar sus dudas (el corrector considera cada respuesta como dada con máxima certeza) o bien, si las dudas son muchas, lo obliga a omitir (el corrector considera al estudiante como totalmente ignorante). Leclercq (2014d) postula que los grados de certeza son una forma de permitir al estudiante expresar su estimación acerca de cuán parcial o completo es su conocimiento en cada respuesta. Si el estudiante es realista, esta será información valiosa para el profesor, le permitirá acceder a información acerca de la seguridad que el estudiante tiene en su conocimiento, que es una parte constitutiva del conocimiento mismo y determina su usabilidad. Probablemente, solo seamos capaces de utilizar el conocimiento del que estamos seguros. Los grados de certeza piden al estudiante indicar, al momento de contestar las preguntas de una prueba, cuán seguro está de la respuesta que acaba de dar. Para esto, Leclercq (2014d) propone una escala de seis niveles expresados en porcentajes: “Estoy [5% - 20% - 40% - 60% - 80% - 95%] seguro de la respuesta que acabo de dar”. En caso de ignorar la respuesta, se le pide al estudiante no omitir, sino que contestar indicando el nivel de seguridad más bajo (5%). No existe descuento de puntaje por respuestas incorrectas. Leclercq (2014f) dedica un capítulo completo al problema de cómo incluir la información obtenida con los grados de certeza en la calificación (el puntaje, la nota) de los estudiantes. Ese punto no será tratado aquí y recomendamos la lectura del capítulo mencionado a quienes se interesen en este punto. Más allá de su inclusión en la calificación, la información obtenida con los grados de certeza puede tener múltiples interpretaciones y usos. Leclercq los llama “un microscopio sobre el pensamiento” (2014d, p. 331), que suma a la información sobre exactitud de las respuestas (correctas vs. incorrectas) una capa extra de información: la seguridad en el propio conocimiento. Así, al analizar las ganancias entre pre y posttest de un estudiante o una cohorte podemos incluir ambos niveles. Un estudiante que contesta correctamente una pregunta en ambos pre y posttests, pero en el pretest estaba 20%

seguro y en el postest está 80% seguro de su respuesta, presenta un 75% de ganancia relativa en su seguridad. Probablemente antes no se sentía seguro de aplicar ese conocimiento y ahora sí lo esté; pero esos avances son invisibles si no usamos una herramienta como los grados de certeza. La tabla 3 (Leclercq, 2014d, p. 343) muestra una forma de interpretar la información obtenida, expresando tanto la exactitud como la seguridad en un plano con dos hemisferios. Al centro del espectro hay un tramo que Leclercq (2014d) llama "ignorancia reconocida", donde se ubica la omisión y las respuestas correctas o incorrectas entregadas con un 5% de certeza. En el plano derecho, se ubica las respuestas correctas. Al principio, con dudas (conocimiento no utilizable) y luego con confianza, ubicándose al extremo las respuestas perfectas, es decir correctas y con el máximo de seguridad. Cada disciplina y profesión tendrá que decidir cuál es el umbral de seguridad apropiado para la enseñanza en su campo (por ejemplo, 95% de seguridad para pilotos de avión, neurocirujanos, ingenieros estructurales; 80% para disciplinas que no impliquen riesgo para la vida humana). En el espectro izquierdo, se ubica las respuestas incorrectas. En el caso de aquellas entregadas con poca seguridad, decimos que el estudiante ha sido prudente, que está consciente de sus dudas y de la parcialidad de su conocimiento, lo cual es positivo; pues la consciencia de una duda gatilla acciones de búsqueda de información, por lo que la toma de consciencia y la expresión de las dudas debería "ser fomentada y no inhibida en el proceso formativo" (Leclercq, 2014d, p. 329). En el extremo izquierdo, se ubica las respuestas que consideramos peligrosas: incorrectas y con un estudiante muy seguro de ellas. Podemos estar aquí frente a supuestos errados, comprensión errónea y, en general, conocimiento difícil de ser modificado.

TABLA 3

Distribución de las respuestas en términos de dominio de los recursos cognitivos

Confusión o mal conocimiento					Ignorancia reconocida ?			Dominio o buen conocimiento				
Imprudente o ignorada (idea falsa)			Prudente o reconocida					Con duda		Confiado		
-95	-80	-60	-40	-20	-5	OM	5	20	40	60	80	95
Total	Alta, temeraria	Baja	Alta	Máxima				Máxima	Alta	Baja	Alta	Total
Peligrosas			Inutilizables					Utilizables				
Porcentaje de imprudencias peligrosas (concepciones erróneas)			Porcentaje de errores prudentes = Prudencia			OM					Porcentaje de confianzas utilizables = Firmeza	
Porcentaje promedio de certeza = Imprudencia (media)								Porcentaje promedio de certeza = Confianza o firmeza (media)				

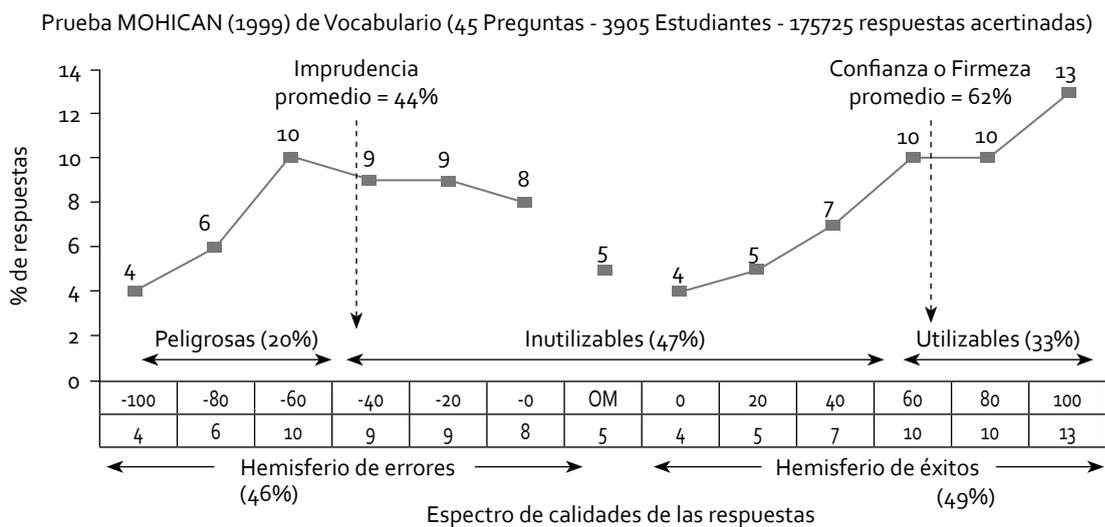
Fuente: Leclercq, 2014d, p. 343.

Leclercq (2014e) propone un mecanismo sistemático para entregar esta información a cada estudiante y pedirle que la use para completar un nivel extra en su proceso de metacognición. De acuerdo a la definición operativa de metacognición de Leclercq y Poumay (2008), está formada por tres tipos de operaciones: autojuicio, autoanálisis (diagnóstico) y autoregulación (cambio para mejorar). Los grados de certeza se circunscriben a la operación metacognitiva de autojuicio (estimación sobre la seguridad en una respuesta). Leclercq (2014e) propone sus “tests espectrales metacognitivos” como una forma de fomentar el autodiagnóstico y la autoregulación. Luego de cada test, el profesor dedica la mitad de una clase a entregar retroalimentación sobre la prueba, discutiendo las respuestas correctas y considerando alternativas propuestas por los estudiantes. En un formato especial, los estudiantes recogen sus impresiones sobre su exactitud y su certeza, enfocándose en dos tipos de situaciones: cuando contestaron correctamente, pero con poca seguridad (¿por qué dudé?) y cuando contestaron erróneamente, pero muy seguros (¿qué entendí mal?). Al final del semestre, Leclercq (2014e) pidió a sus estudiantes analizar sus registros, tomados después de cada test, y escribir una reflexión metacognitiva describiendo esta información, incluyendo tendencias o patrones si los hubiere y propuestas de mejora para un futuro semestre.

Las figuras 6, 7 y 8 grafican otros usos posibles de la información conjunta sobre exactitud y certeza. La figura 6 muestra el rendimiento de una cohorte completa (n = 3905) en una sección (vocabulario) de un test masivo aplicado a los estudiantes que ingresaron en 1999 a las universidades de la Bélgica francófona (Leclercq, 2014d). El perfil que forman las 175 725 respuestas incluidas es positivo. Se observa una gran cantidad de respuestas correctas con altos niveles de seguridad y pocas respuestas incorrectas en los niveles de seguridad más altos. La figura 7 muestra el uso que De la Fuente (Leclercq, De la Fuente & Cabrera, 2014) hizo de los grados de certeza para analizar el comportamiento de determinadas preguntas en sus tests. Detectó preguntas que llamó “sin problemas” (donde las respuestas incorrectas presentan bajos niveles de seguridad) y también preguntas “con problemas” (donde muchos estudiantes contestaban de forma incorrecta y a la vez muy seguros). Junto a sus colegas, determinaron que varias de estas preguntas se concentraban en un grupo específico de contenidos (mecánica de fluidos en un curso de física para futuros médicos) y que estudiantes que contestaban correctamente la mayoría de las otras preguntas, en estos contenidos contestaban mal y muy seguros. Estos profesores fueron lo suficientemente autocríticos como para aceptar que existía un problema en la forma en que estaban enseñando esa área de contenidos y tomaron medidas para mejorar. La figura 8 muestra los perfiles de un estudiante en un pre y un postest. La ganancia es evidente en la disminución de la omisión, el aumento de las respuestas correctas y la seguridad en estas respuestas.

FIGURA 6

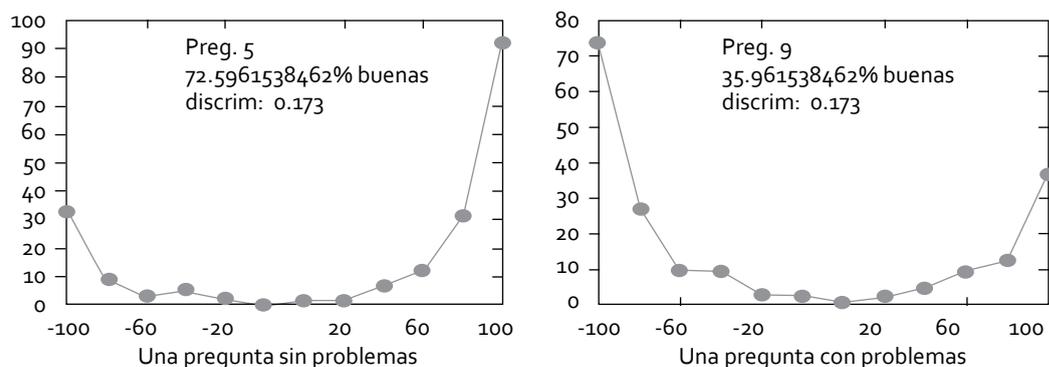
Distribución de las respuestas "acertadas" (es decir, que incluyen exactitud y seguridad) de una cohorte en un test (n = 175 725) respuesta)



Fuente: Leclercq, 2014d, p. 342.

FIGURA 7

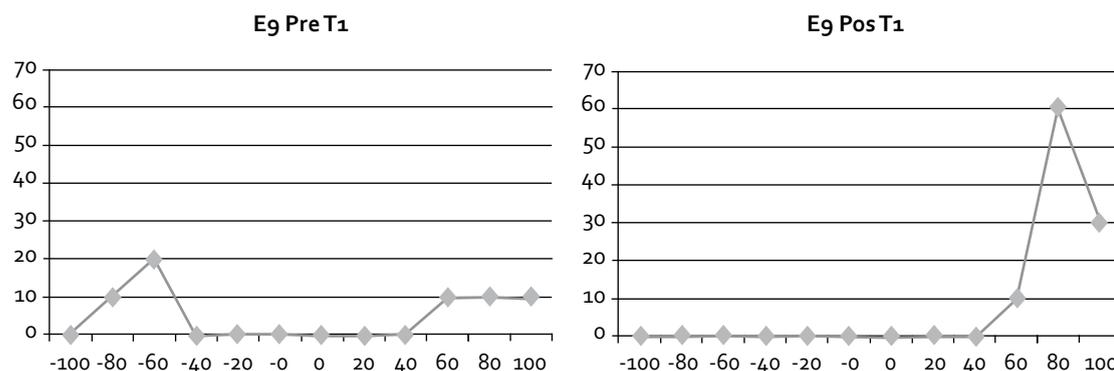
Distribución de las respuestas a dos preguntas: una sin problemas (preg. 5) y otra con problemas (preg. 9)



Fuente: Leclercq, De la Fuente & Cabrera, 2014, p. 445.

FIGURA 8

Distribución de las respuestas de un mismo estudiante (est. 9) en pretest y postest



Fuente: Cabrera, 2014, p. 213.

■ INSTRUMENTO 2. PROYECTO ESTUDIANTIL: LA PUESTA EN ESCENA COMO INSTRUMENTO DE EVALUACIÓN

Cohérente con el hecho de escoger a la pedagogía por proyectos como la principal estrategia del curso, el proyecto estudiantil fue la actividad de aprendizaje con mayor peso en la calificación final. El 40% de esta se constituía de evaluaciones asociadas al proyecto: 30%, la puesta en escena de una actividad artística con público; 5%, un registro del evento que los estudiantes debían hacer disponible en una WEB 2.0; y 5%, un plan de producción que formaba parte del reporte escrito. El proyecto estudiantil contribuyó a evaluar dos de los resultados de aprendizaje del curso (RA1: “producir un proyecto artístico” y RA3: “trabajar en equipo”). En el uso que de este instrumento, se hace evidente la distinción entre evaluación con intención formativa y evaluación con función sancionadora (calificación). En la evaluación de la puesta en escena, tanto en el montaje previo como en el espectáculo propiamente dicho, se consideró una gran cantidad de aspectos; entre estos, la calidad de los artistas que los estudiantes —actuando como productores— lograron convocar, la atmósfera global lograda con el público, diversos aspectos técnicos de organización, de recursos (escenario, refuerzo sonoro, iluminación) y la puntualidad. Como evaluadores, observamos y tomamos nota de estos aspectos, incluyéndolos en la sesión de retroalimentación posterior a la puesta en escena. Sin embargo, la evaluación detallada de estos aspectos no fue incluida en la calificación. Para la función sancionadora de

este instrumento de evaluación, decidimos utilizar un criterio único: “el grupo logra la puesta en escena de una actividad artística con público”, y solo dos posibilidades: “logrado” o “no logrado”. La condición de logrado estaba definida como “cada aspecto de la puesta en escena, bajo el control de los estudiantes, ocurre como fue anticipado en la planificación (o mejor) y complicaciones menores son resueltas con rapidez” (Cabrera, 2014, p. 204). Esta condición suponía la máxima calificación en la escala chilena (7.0) para todo el grupo. Tenía, por lo tanto, una alta probabilidad de ocurrir. El propósito de esta forma de calificar era indicar a los estudiantes que el desafío del proyecto no estaba en obtener la más alta calificación (lo que era muy probable), sino en organizar y llevar a escena una actividad artística que satisficiera las expectativas de un público real. El curso se dividió en dos grupos y dos eventos de gran nivel se realizaron en noviembre de 2009. Todos los estudiantes aprobaron el curso. Entre las fortalezas de la evaluación, se cuenta la coherencia entre la pedagogía por proyectos (PPP) y esta forma de calificar—con un criterio único y condición de logrado suponiendo la calificación máxima—. Esta forma de calificar refuerza el mensaje del curso: el desafío no está en la nota o el grado, sino en la calidad del proyecto, que debe satisfacer las expectativas de una audiencia. De esta forma, se intentó disminuir la “motivación extrínseca” (la presión por la calificación) y aumentar la motivación intrínseca (Deci & Ryan, 1985). Las debilidades principales son dos: 1. La competencia de producción, evaluada grupalmente, contó con un solo evaluador (el profesor), faltando riqueza y variedad de perspectivas; y 2. La competencia individual de producción no fue evaluada (solo se evaluó la competencia grupal). En forma individual, se evaluó la resolución de ciertos problemas y recursos aislados, pero distinto a lo realizado por el estudiante en el proyecto colectivo. Esta debilidad cuestiona en parte la validez teórica del dispositivo de evaluación de los aprendizajes (DEA) del curso. Específicamente, la representatividad de contenidos, pues como profesor no puedo afirmar que todos los estudiantes lograron todos los aprendizajes que consideraba necesarios como evidencias del logro de esta competencia. Las propuestas de mejora se desprenden de las debilidades descritas. Primero, convocar a un jurado experto, conformado por tres o cuatro profesionales o profesores que acuerden una pauta de observación, estén presentes durante el montaje previo y en la puesta en escena y participen de la sesión de retroalimentación. Segundo, generar un nuevo instrumento de evaluación que entregue evidencias de la competencia de producción de cada estudiante—por ejemplo, una bitácora individual de productor, referida exclusivamente al proyecto colectivo, donde cada estudiante deba registrar las principales decisiones del grupo, así como sus aportes personales (tareas completadas, responsabilidades asumidas, ideas propuestas)—. En síntesis, esta forma de usar la puesta en escena como parte de la evaluación funcionó, reforzó el enfoque general del curso y a la vez presentó importantes áreas de mejora donde es necesario corregir o sumar.

■ INSTRUMENTO 3. RÚBRICAS PARA EVALUAR EL TRABAJO EN EQUIPO

El 15% de la calificación final del curso correspondió a la evaluación de la competencia “trabajar en equipo”, asociada a uno de los resultados de aprendizaje del curso (RA3). Esta fue evaluada por los estudiantes en formato de autoevaluación (10%) y evaluación de pares (5%). Para definir operacionalmente lo que debía ser evaluado como parte del trabajo en equipo, nos apoyamos en cuatro rúbricas publicadas por Villa y Poblete (2007). Los autores proponen describir el trabajo en equipo a través de cuatro indicadores y entregan una rúbrica con cinco niveles de logro para cada indicador. Una rúbrica es una escala descriptiva que se usa (se necesita) como apoyo en la evaluación de desempeños complejos (Leclercq & Cabrera, 2014, p. 123). Que la escala sea descriptiva quiere decir que cada nivel está descrito en términos de evidencias de logros de aprendizaje que son esperables u observables en ese nivel (o evidencias de dificultades en el aprendizaje, en el caso de los niveles más bajos). Su uso es una ayuda para los evaluadores o jueces, que deben observar, ponderar y emitir un juicio sobre un desempeño complejo o una producción de los estudiantes, usualmente en tiempos muy acotados. La rúbrica guía a los jueces en la observación y disminuye la dispersión de evaluaciones de los jueces en comparación con otras formas de evaluar que no cuentan con escalas descriptivas (Leclercq & Cabrera, 2014, pp. 121-122). Las rúbricas siguen una lógica incremental, donde en general el escalón superior supone el logro del inmediatamente anterior. De este modo, muestran —a profesores y estudiantes— un itinerario esquemático de estadios de desarrollo de un determinado aprendizaje (Leclercq & Cabrera, 2014, p. 123). Los niveles más altos de logro de un conjunto de rúbricas —por ejemplo, de todas las rúbricas usadas en un programa de estudio, o todas las usadas en un semestre o año particular— pueden entregar una imagen de aquello que para el cuerpo de profesores constituye las máximas aspiraciones formativas, las más altas expectativas del currículum. Las rúbricas además son flexibles: pueden ser diseñadas para evaluar productos específicos (un proyecto, una presentación), o para evaluar competencias genéricas (como el trabajo en equipo). Además, pueden adaptarse para ser usadas por distintos agentes evaluadores, como jueces externos, el profesor y los estudiantes en la forma de autoevaluación y evaluación de pares.

La tabla 4 muestra la rúbrica de Villa y Poblete (2007) para el indicador 4, en formato de evaluación de pares, tal como se usó en el curso. Cada estudiante recibió las cuatro rúbricas (una por cada indicador) y al costado un espacio para indicar los nombres de sus compañeros de equipo y el nivel de la rúbrica en que ubicaba a cada uno respecto a ese indicador. La tabla 5 muestra la rúbrica para el indicador 1, en formato de autoevaluación, tal como se usó en el curso. Nuevamente, a cada estudiante se le entregaron las cuatro rúbricas, ahora redactadas como autoevaluación, junto a una plantilla para que indicaran el nivel de la rúbrica en que se autoubicaba en cada indicador.

TABLA 4

Rúbrica para la evaluación de pares del indicador 4: trabajo en equipo

Indicador 4				
Toma en cuenta los puntos de vista de los demás y retroalimenta de forma constructiva				
1	2	3	4	5
No escucha las intervenciones de sus compañeros y los descalifica sistemáticamente. Quiere imponer sus opiniones.	Escucha poco, no pregunta, no se preocupa de la opinión de los otros. Sus intervenciones son redundantes y poco sugerentes.	Acepta las opiniones de los otros y sabe dar su punto de vista en forma constructiva.	Acepta las opiniones de los otros y sabe dar su punto de vista en forma constructiva.	Acepta las opiniones de los otros y sabe dar su punto de vista en forma constructiva.
			Fomenta el diálogo constructivo e inspira la participación de calidad de los otros miembros del grupo.	Fomenta el diálogo constructivo e inspira la participación de calidad de los otros miembros del grupo.
				Integra las opiniones de los otros en una perspectiva superior, manteniendo un clima de colaboración y apoyo.

Fuente: Cabrera, 2014, p. 211.

TABLA 5

Rúbrica para la autoevaluación del indicador 1: trabajo en equipo

Indicador 1				
Realicé las tareas que me fueron asignadas por el grupo dentro de los plazos requeridos				
1	2	3	4	5
No cumplí con las tareas asignadas.	Cumplí parcialmente las tareas asignadas.	Cumplí las tareas asignadas.	Cumplí las tareas asignadas.	Cumplí las tareas asignadas.
	No cumplí los plazos requeridos.	En los plazos requeridos.	En los plazos requeridos.	En los plazos requeridos.
			La calidad de mi tarea supuso un notable aporte al equipo.	La calidad de mi tarea supuso un notable aporte al equipo.
				Mi trabajo orientó y facilitó el del resto de los miembros del equipo.

Fuente: Cabrera, 2014, p. 210.

El uso de rúbricas presentó varias fortalezas. En primer lugar, permitió implementar autoevaluación y evaluación de pares. Los indicadores que describían lo que se debía entender como trabajo en equipo y las descripciones de evidencias de logro en los niveles materializaron una expresión abstracta y a veces repetida sin mucha reflexión: trabajo en equipo. Esto permitió a los estudiantes pensar y discutir acerca de estos puntos. Según sus palabras, les dio un lenguaje para hablar de cuestiones que no habían podido tratar antes de conocer las rúbricas. La primera aplicación de la autoevaluación gatilló un diálogo sincero entre los estudiantes que culminó con la conformación de dos grupos y la puesta en escena de dos actividades artísticas. Por último, haber usado una misma rúbrica adaptada para ambas evaluaciones nos permitió hacer análisis posteriores comparando la autoevaluación con la evaluación de pares, buscando saber, por ejemplo, cuán realistas fueron los estudiantes; es decir, cuán cerca o lejos estuvieron sus propias evaluaciones de las que sus pares hicieron sobre ellos. O cuáles estudiantes se subestimaron y cuáles se sobreestimaron respecto a la evaluación que de ellos hicieron sus compañeros. La principal debilidad en la evaluación de este resultado de aprendizaje es que la competencia de trabajo en equipo, a nivel individual, no fue evaluada por el profesor; solo fue evaluada por los estudiantes en los formatos que acabo de describir. Las dificultades para realizar esta evaluación son evidentes; ya que, para evaluar la capacidad de trabajar en equipo de cada estudiante, el evaluador debería estar presente mientras ese trabajo ocurre. Y el trabajo en equipo tiene lugar principalmente fuera de la sala de clases, en espacios del campus, las residencias de los estudiantes y, cada vez más, en línea, con el uso masivo de herramientas como los documentos de Google. Gillet y Marée (2005) describen una experiencia diseñada para que la observación de los equipos trabajando fuera posible por parte del profesor. Este curso (“Legislación en el ámbito de la salud”) estaba concebido como pedagogía por proyectos. Durante el semestre, cada grupo de estudiantes debía jugar el rol del gobierno y preparar un proyecto de ley. Cada clase, de tres horas, se dividió en dos mitades. La primera se dedicaba al trabajo de investigación para la redacción de la ley. Este ocurría en la sala de clases y el profesor observaba el trabajo de los grupos, anotando en una pauta comentarios y puntos extra o puntos menos para los estudiantes, de acuerdo a algunos criterios. La tabla 6 muestra una versión esquemática de esta pauta. Durante la segunda mitad de la clase, los grupos se alternaban personificando al gobierno y presentando al Parlamento los avances en su proyecto de ley. Cuando no personificaban al gobierno, los grupos jugaban el rol de bancadas de partidos de oposición que cuestionaban, criticaban y exigían al gobierno más información respecto del proyecto. Al final del semestre, la calificación final del proyecto grupal se veía afectada por los puntos extra o en contra que el estudiante recibió individualmente durante la observación semestral del profesor.

TABLA 6

Criterios de la pauta de Gillet y Ercicum: versión esquemática (las celdas reales permiten tomar apuntes para justificar los descuentos o puntos extras)

		Est. 1	Est. 2	Est. 3	Est. 4	Est. 5
Trabajo en grupo	Cohesión con el equipo	- -	+		+	++
	Asistencias	+	+			
	Cortesía, respeto por los demás					
	Propone ideas	-	++			
	Estructura los argumentos o los datos					
	Total					
Presentaciones orales	Mira al público			-		
	Dicción	++			++	
	Audible			--		
	Mensajes claros					++
	Mensajes precisos					
	Mensajes concisos			+		
	Respeto al tiempo	+			+	+
	No lee las láminas					
	No lee su texto					
	Total					
Debate	Calma					
	Respuestas precisas					
	Respuestas concisas					
	Juega el rol parlamentario					
	Trabajo en equipo (cohesión)					
	Cortesía, Respeto					
	Total					

Fuente: Leclercq, Gillet, Ercicum & Cabrera, 2014, p. 239.

La última sección de este artículo aborda la pregunta por la validez del dispositivo de evaluación de los aprendizajes (DEA) que se implementó en este curso. Chatterji define “validez” como aquello que “refiere a la exactitud de las inferencias e interpretaciones que hacemos a partir de los resultados o ‘puntajes’ de una herramienta de evaluación” (2003, p. 54). La validez puede verse afectada por errores sistemáticos que se nos cuelan en el proceso y distorsionan

lo que estamos evaluando —lo alejan de lo que queremos evaluar—, debido a cómo creamos o a cómo implementamos un instrumento de evaluación (Chatterji, 2003, p. 54). Eliminar al extremo posible las fuentes de error sistemático y asegurar la validez de las inferencias que hacemos de los resultados de la evaluación, es considerada la más importante tarea de los expertos en evaluación (AERA, APA & NCME, 2014; Chatterji, 2003). Las principales asociaciones estadounidenses involucradas en evaluación educacional proponen entender la validez como un concepto unitario (AERA, APA & NCME, 2014). Chatterji (2003) propone que, al determinar la validez de las inferencias que hacemos, nunca perdamos de vista los propósitos de la evaluación (¿para qué estamos evaluando?) y la población específica que está siendo evaluada (¿a quiénes voy a evaluar?). Así, el juicio sobre la validez de las inferencias que podemos extraer de un instrumento de evaluación (o, en nuestro caso, del DEA de un curso), deberá ser una decisión tomada con un enfoque holístico que pondere integralmente las distintas dimensiones de validez del instrumento y las contextualice respecto a los propósitos y poblaciones específicas de la evaluación.

Leclercq propuso un modelo para analizar sistemáticamente las dimensiones que componen la validez de un instrumento de evaluación. Estableció ocho dimensiones de validez y los organizó en el acrónimo ETICPRAD, que representa la validez ecológica, validez teórica, validez informativa, validez consecucional, validez predictiva, “replicabilidad” (fiabilidad), aceptabilidad y validez deontológica (Leclercq, 2006 y 2014a). Lo que sigue es un análisis del DEA del curso Gestión de Proyectos Artísticos de acuerdo a las ocho dimensiones propuestas en el modelo ETICPRAD, finalizando con el juicio integral acerca de la validez de las inferencias que se pudo extraer de este conjunto de instrumentos de evaluación de los aprendizajes.

La primera dimensión de validez propuesta por Leclercq (2006 y 2014a) es la validez “ecológica”. Esta refiere al nivel de autenticidad de la tarea que se evalúa. La pregunta que pretende responder es: ¿qué tan cercanas a la realidad (de una profesión u ocupación) son las condiciones de la evaluación? El supuesto es que tareas auténticas —es decir, tareas realizadas en condiciones que logran representar situaciones que los estudiantes podrían encontrar en sus futuras ocupaciones— incrementan la motivación de los estudiantes al aumentar su percepción de valor acerca de la tarea (Viau, 1994). En el caso del curso, la validez ecológica es alta. Primero, porque utiliza la pedagogía por proyectos como principal estrategia metodológica. La principal actividad de aprendizaje (y de evaluación) del curso fue un proyecto donde los estudiantes actuaron como profesionales (productores artísticos) en una situación muy cercana a la realidad que quienes trabajan en esta área encuentran en el medio (la demanda de convocar público real fue clave para lograr esta autenticidad). En segundo lugar, considero que la forma de calificar

que se escogió para este curso es coherente con este enfoque y refuerza su mensaje central: el desafío de aprendizaje no se centra en la calificación final, sino en satisfacer las expectativas de una audiencia que asiste a una actividad artística.

La segunda dimensión que Leclercq (2006 y 2014a) propone es la validez “teórica”. Esta puede ser dividida en dos tipos: la validez teórica de cobertura o contenido y la validez teórica de constructo. La primera es lo que en inglés se conoce como *content-based evidence of validity* (Chatterji, 2003, pp. 58-61). Alcanzar esta validez de contenido implica que todos los contenidos relevantes sean evaluados (*content relevance*) y que estos estén representados proporcionalmente en la evaluación (*content representativeness*). Por otro lado, la validez teórica de constructo es más alta en la medida que la “definición operacional” de aquello que se evalúa tiene sustento en teorías, en resultados de investigación o en sistematizaciones de conocimientos surgidos desde la práctica (Chatterji, 2003; Leclercq, 2014a). La definición operacional refiere a los indicadores que describen los aprendizajes esperados en el curso y las tareas o preguntas de evaluación que permiten reunir evidencias acerca del logro de esos indicadores (Chatterji, 2003). La validez teórica de constructo busca que indicadores e ítems se elaboren basados en la literatura especializada. Una de las formas de medir esta dimensión de validez es lo que en inglés se conoce como *validity evidence of internal structure* (Chatterji, 2003, pp. 62-63), donde las respuestas dadas a grupos de ítems que han sido creados de acuerdo a constructos o sub-constructos propuestos en la literatura son analizadas en conjunto (usando el coeficiente Alpha de Cronbach, por ejemplo), para determinar si las hipótesis teóricas presentes en la literatura se ven reforzadas en el caso de un curso particular. Para el DEA del curso, se estima el nivel de validez teórica como moderado o parcial. Dentro de las fortalezas, se cuenta con: a) el soporte en teorías, modelos y resultados de la investigación de los diversos componentes del curso —validez teórica de constructo—; y b) el chequeo de triple coherencia realizado como parte del diseño, lo que mejora la validez teórica de contenido. En las debilidades, se cuenta con dos importantes vacíos: a) la competencia individual para producir proyectos artísticos; y b) la competencia individual para trabajar en equipo, no fueron evaluadas por el profesor. Esta omisión impide como profesor asegurar que todos los estudiantes efectivamente desarrollaron todos los aprendizajes esperados, cuestionando en parte la validez teórica de contenido del DEA.

La tercera dimensión de validez es la “informativa” (Leclercq, 2006 y 2014a), que se relaciona con la cantidad y la calidad de las informaciones que nos entregan los instrumentos de evaluación. Estas informaciones pueden ser acerca de los aprendizajes de los estudiantes (idealmente, con detalle suficiente como para permitir diagnósticos precisos acerca de lo que cada estudiante ha logrado y las áreas en que requiere más apoyo para continuar su desarrollo). Estas informaciones

también pueden iluminar nuestra propia docencia, indicándonos los puntos donde se puede mejorar y pueden ser de utilidad a la institución, como parte de la evaluación de sus programas académicos o de las acciones de desarrollo profesional que realizan los docentes, por ejemplo. Mi evaluación es que en el curso la validez informativa es alta. Primero, por el hecho de haber utilizado una gama amplia de instrumentos de evaluación (multimodal), lo que generó una gran variedad de informaciones respecto al aprendizaje de los estudiantes. En segundo lugar, el uso de grados de certeza (Leclercq, 2014d) permitió acceder a un nivel de información sobre el aprendizaje de los estudiantes rara vez disponible: su propia apreciación sobre la seguridad en los conocimientos adquiridos, como una representación de su conocimiento parcial.

La cuarta dimensión del modelo ETICPRAD es la validez “consecuencial” (Leclercq, 2006 y 2014a). Esta refiere a las consecuencias, positivas y negativas, que el uso de determinados instrumentos de evaluación puede tener para los estudiantes (por ejemplo, en sus hábitos de estudio), para los profesores (por ejemplo, inconscientemente modificar su enseñanza para enfocarla en “entrenar” a los estudiantes solo para rendir en un cierto tipo de test), para la cultura de la institución (por ejemplo, recompensar este tipo de enseñanza y no otros) y para el sistema en general (por ejemplo, evaluaciones que resultan en “etiquetas” para instituciones, las que son tratadas con liviandad por los medios de comunicación masiva y a veces por los gestores de política pública). La evaluación de este tipo de validez requiere estudios longitudinales que sigan a los involucrados por tiempos más extendidos que lo relatado en esta experiencia. Por lo tanto, los datos obtenidos no fueron suficientes para realizar una estimación sobre esta dimensión de validez en el DEA del curso. Información anecdótica, referida al devenir de los estudiantes que pasaron por el curso, indica que algunos de ellos encontraron en la organización de proyectos artísticos una veta de trabajo que continuaron cultivando. En ese caso, la evaluación habría tenido al menos la consecuencia positiva de motivar a los estudiantes a continuar esta línea de trabajo (o no abandonarla).

La quinta dimensión propuesta por Leclercq (2006 y 2014a) es la validez “predictiva”, entendida como la capacidad de un instrumento de evaluación de predecir desempeños futuros del estudiante, en evaluaciones por venir o futuras ocupaciones. Por ejemplo, para una prueba nacional de selección universitaria (como la PSU que se usa en Chile), la validez predictiva es de la mayor importancia, pues lo que esperamos de esta prueba es básicamente que nos indique quiénes de aquellos que la rinden son los más aptos para tener éxito en una carrera universitaria. El valor de la prueba radica en su capacidad de predecir un desempeño futuro (el éxito en la universidad, la obtención del grado). Nuevamente, la evaluación de esta dimensión de validez requiere estudios longitudinales y comparación entre los resultados de nuestro instrumento

y otras evaluaciones posteriores. Por lo tanto, esta dimensión no fue evaluada en el caso del curso que referido.

La sexta dimensión de validez propuesta por Leclercq (2006 y 2014a) es la “replicabilidad” o “fiabilidad” (en inglés, *reliability*; en francés, *fidélité*). La fiabilidad refiere a la consistencia de los resultados de una evaluación en el tiempo y entre distintos correctores (Chatterji, 2003; Leclercq, 2014a). Los evaluadores se esfuerzan por aumentar la fiabilidad minimizando al máximo las fuentes de “error de medición aleatorio” (Chatterji, 2003, p. 67). Esto se logra incluyendo una cantidad suficiente de ítems o controlando las condiciones en que ocurre la evaluación para que sean óptimas y equitativas para todos los evaluados. En el caso del DEA de mi curso, las condiciones de evaluación impidieron recolectar datos que permitieran evaluar esta dimensión. La mayoría de las evaluaciones ocurrieron en un momento único, sin repeticiones (como la puesta en escena del proyecto) y tampoco hubo varios jueces. Por lo tanto, la fiabilidad no fue evaluada en el DEA de este curso.

La séptima dimensión del modelo ETICPRAD (Leclercq, 2006 y 2014a) es la “aceptabilidad”. Esta tiene dos significados: primero, un sentido equivalente a lo que en inglés se conoce como *utility* (Chatterji, 2003, pp. 70-73); es decir, refiere a la factibilidad de implementar un instrumento de evaluación en términos de tiempo, recursos, logística, entre otros aspectos. La pregunta aquí es: ¿resulta viable este instrumento de evaluación? El segundo significado de aceptabilidad refiere a si el tipo de evaluación es coherente con los valores, principios y cultura de los estudiantes, profesores y la institución (Leclercq, 2014a). En la experiencia relatada, se evalúa que la validez de aceptabilidad es moderada o parcial. En primer lugar, el DEA propuesto fue aceptado y bien recibido por los estudiantes, quienes lo entendieron y se comprometieron en su cumplimiento. Por otro lado, la variedad de actividades de aprendizaje y de instrumentos de evaluación implementados implicó una alta carga de trabajo para mí, en mi calidad de profesor. Este hecho puede afectar la viabilidad de un DEA como este.

Por último, la octava dimensión de validez que propone Leclercq (2006 y 2014a) es la validez “deontológica”. Esta dimensión aborda cuestiones éticas. Primero, el problema de la justicia en la evaluación, de la ecuanimidad al momento de evaluar a todos los estudiantes. Chatterji (2003, p. 57) lo llama *fairness*. Esta ecuanimidad se materializa al asegurar condiciones de evaluación equivalentes para todos los estudiantes. También cuando tomamos consciencia de los sesgos que pueden afectar nuestra capacidad de juzgar. Los sesgos existen y como profesores debemos intentar detectarlos y eliminarlos de nuestra práctica habitual (incluyendo nuestra evaluación). Entre estos se cuenta: sesgos de “halo” (que preceden al evaluado y prejucian nuestra estimación), sesgos de contraste (que perjudican o benefician a estudiantes por la comparación que hacemos

con los que inmediatamente les antecedieron), o el efecto *Posthumus* (De Landsheere, 1992[1971]), donde los docentes modifican su enseñanza y evaluaciones de modo que se perpetúa una distribución normal (curva de Gauss). También existen sesgos culturales, de género, raciales y de clase que afectan a la evaluación. Estos sistemáticamente perjudican a determinados grupos humanos —muchas veces minoritarios— por su género, raza, etnicidad, lenguaje, clase social o una interacción de estos rasgos. Un ejemplo lo constituyen minorías lingüísticas que no son evaluadas en el idioma que mejor dominan o cuyo bilingüismo no es evaluado y por tanto no es valorado como parte de sus fortalezas académicas (García & Baetens Beardsmore, 2009). En el DEA del curso, se evalúa la validez deontológica como moderada o parcial. Las fortalezas estuvieron en que las mismas condiciones de enseñanza y evaluación se aplicaron a todos los estudiantes y en que, como profesor, realicé una revisión de los instrumentos de evaluación en busca de sesgos sistemáticos. La principal debilidad se relaciona con uno de los vacíos mencionados antes: el aporte individual a la producción grupal no fue evaluado por mí, en mi rol de profesor (solo fue evaluado por los estudiantes como evaluación de pares y autoevaluación). En el desarrollo del proyecto, se pudo haber producido desequilibrios notorios en las cargas de trabajo que los estudiantes asumieron. Esto no fue reconocido en la calificación individual, lo que constituye una injusticia.

Con esta visión panorámica (tres dimensiones sin datos, tres dimensiones con validez moderada y dos dimensiones con validez alta), ¿se puede decir que el dispositivo de evaluación de los aprendizajes (DEA) creado para el curso Gestión de Proyectos Artísticos es válido? Para contestar esta pregunta —es decir, para emitir un juicio global acerca de validez de mi DEA— debo recordar la advertencia de Chatterji (2003): nunca olvidar los propósitos de la evaluación, el para qué estamos evaluando. En el caso del curso, un primer propósito era definir la aprobación o reprobación de un estudiante. Para esto son importantes la validez teórica (moderada), la validez informativa (alta), la validez deontológica (moderada) y la fiabilidad (sin datos). Un segundo propósito del DEA fue reforzar el enfoque de la pedagogía por proyectos. Para esto, era importante la validez ecológica (alta). Un tercer y último propósito fue mejorar mi propia docencia, por lo que la validez informativa (alta) era importante. La validez predictiva era menos relevante para mis propósitos. En cuanto a la validez consecuencial, no tenemos evidencias que sugieran consecuencias negativas e información anecdótica sugiere consecuencias positivas o neutras. Chatterji (2003) también nos llama a nunca olvidar que la evaluación se crea para ser aplicada a una determinada población y para evaluar un particular grupo de características o constructos teóricos. Por lo tanto, la respuesta a la pregunta sobre la validez del DEA del curso es que, para evaluar estos constructos, con estos propósitos y en estos estudiantes, el DEA es válido. Dicho aún con más exactitud, las inferencias que puedo hacer a partir de los resultados

—o “puntajes”— de los instrumentos del DEA son válidas en tanto refieran a estos aprendizajes (constructos) desarrollados por estos estudiantes (Chatterji, 2003, p. 55).

DISCUSIÓN

El DEA del curso descrito fue a propósito concebido y diseñado para incluir una amplia variedad de instrumentos de evaluación. La pregunta que quiero plantear es acerca de la necesidad (y la viabilidad) de desplegar, en un solo curso, la variedad de instrumentos que incluyó este DEA. Comparto la noción de que un estudiante debe experimentar una variedad de situaciones de aprendizaje, que incluso apelen a distintas formas de aprender. Esta variedad debe extenderse también a los instrumentos de evaluación que utilizamos. Sin embargo, lograr esta variedad no puede ser la tarea individual de un profesor o de todos los profesores aisladamente. Lograr variedad en metodologías y evaluaciones debe ser una tarea colectiva y coordinada entre los profesores. Al mismo tiempo, ningún DEA implementado en un curso particular podrá alcanzar, a la vez, los máximos niveles de validez en todas las ocho dimensiones descritas en el modelo ETICPRAD (Leclercq, 2006 y 2014a). Lograr niveles óptimos en todas las dimensiones de validez para el conjunto de evaluaciones del programa es también una tarea que debe ser asumida como una empresa colectiva por los profesores y la institución. Aquí es donde surge un interesante campo de trabajo e investigación futura: dispositivos de evaluación de los aprendizajes (DEA) concebidos, diseñados e implementados en grupos de cursos (por ejemplo, en niveles anuales o semestrales) y en programas académicos completos (como los ejemplos descritos por Cabrera, Van der Vleuten & Leclercq, 2014).

CONCLUSIÓN

En este artículo se describe el dispositivo de evaluación de los aprendizajes (DEA) implementado en el curso Gestión de Proyectos Artísticos dictado en la Facultad de Artes de la Universidad de Chile el segundo semestre del año 2009. De esta forma, se muestra una materialización posible de conceptos clave y modelos en evaluación, como la “triple concordancia” (Leclercq, 1995), el Modelo de 8 Eventos de Aprendizaje (Leclercq & Poumay, 2005) y el Dispositivo de Evaluación de los Aprendizajes (Leclercq & Cabrera, 2014). Además, se describe variados instrumentos de evaluación, como pruebas con preguntas de selección múltiple con grados de certeza (Leclercq, 2014d), un proyecto, evaluación de pares y autoevaluación. Se analiza luego las fortalezas y debilidades del DEA del curso. De las debilidades surgen propuestas de mejora ilustradas con el trabajo que otros académicos han hecho en otros cursos. Finalmente, se evalúa el DEA con el lente de

las ocho dimensiones del modelo ETICPRAD (Leclercq, 2006 y 2014a). Considerando propósitos, constructos y población de la evaluación, se emite un juicio positivo acerca de la validez de las inferencias que pudimos hacer sobre la base del DEA del curso. Este ejercicio de análisis reforzó la convicción de que la docencia universitaria es una tarea compleja que requiere de todo nuestro esfuerzo y rigor intelectual. La docencia —el diseño y la implementación del currículum, los métodos y las evaluaciones de un curso, y el posterior análisis de sus resultados— demanda “serio trabajo intelectual” (Bain, 2004, p. 49). Es una tarea que debemos abordar con el mismo rigor con que enfrentamos nuestra investigación y que debe ser reconocida como tal en los sistemas de evaluación y promoción académica de las instituciones de educación superior.

Bibliografía

- American Educational Research Association (AERA), American Psychological Association (APA) & National Council on Measurement in Education (NCME) (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Anderson, J., Reder, L. & Simon, H. (1996). Situated learning in education. *Educational Researcher*, 25(4), 5-11.
- Angelo, T. & Cross, P. (1993). *From classroom assessment techniques: A handbook for college teachers* (segunda edición). New York, NY: Jossey Bass.
- Bain, K. (2004). *What the best college teachers do*. Cambridge, MA: Harvard University Press.
- Beckers, J. (2002). *Pédagogie des compétences*. Bruxelles: Labor.
- Bransford, J. D., Brown, A. L. & Cocking, R. R. (1999). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.
- Cabrera, A. (2014). Evaluar los aprendizajes en la pedagogía por proyectos (PPP). En D. Leclercq & A. Cabrera (eds.), *Ideas e innovaciones: Dispositivos de evaluación de los aprendizajes en la educación* (pp. 197-220). Santiago de Chile: Editorial Universitaria.
- Cabrera, A., Van der Vleuten, C. & Leclercq, D. (2014). Concebir dispositivos de evaluación de los aprendizajes (DEA) al nivel de un programa. En D. Leclercq

- & A. Cabrera (eds.), *Ideas e innovaciones: Dispositivos de evaluación de los aprendizajes en la educación* (pp. 419-430). Santiago de Chile: Editorial Universitaria.
- Chatterji, M. (2003). *Designing and using tools for educational assessment*. Boston, MA: Allyn & Bacon.
- D'Ivernois, J. E. & Chabot, J. M. (2003). *Guide de construction des cascades (convergentes et divergentes)*. Université de Paris 13, UER de médecine de Bobigny. DPSS.
- De Landsheere, G. (1992[1971]). *Evaluation continue et examens: Précis de docimologie*. Bruselas: Labor. París: Fernand Nathan.
- Deci, E. L. & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York, NY: Plenum Press.
- García, O. & Baetens Beardsmore, H. (2009). Assessment of bilinguals. En O. García (ed), *Bilingual education in the 21st century: A global perspective* (pp. 366-379). West Sussex: John Wiley & Sons.
- Gillet, P. & Marée, M. (2005). Un exemple de méthode active d'apprentissage: la simulation de séances parlementaires. *Puzzle. CIFEN, Université de Liège*, 18, 17-20.
- Greeno, J. (1997). On claims that answer the wrong questions. *Educational Researcher*, 26(1), 5-17.
- Kilpatrick, W. (1918). *The project method*. Teachers College Record.
- Lang, J. M. (2005). *Life on the tenure track: Lessons from the first year*. Baltimore, MD: The John Hopkins University Press.
- Lave, J. & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: University of Cambridge Press.
- Leclercq, D. (1995). *Conception d'interventions et construction de produits pour l'éducation*. Liège: Editions de l'université de Liège.
- Leclercq, D. (2006). L'évolution des QCM. En G. Figari & L. Mottier-Lopez (eds), *Recherches sur l'évaluation en Education* (pp. 139-146). París: L'Harmattan.
- Leclercq, D. (2014a). ETICPRAD: Ocho criterios de validez de un dispositivo de evaluación de los aprendizajes. En D. Leclercq & A. Cabrera (eds.), *Ideas e*

- innovaciones: Dispositivos de evaluación de los aprendizajes en la educación* (pp. 83-91). Santiago de Chile: Editorial Universitaria.
- Leclercq, D. (2014b). Los componentes del dispositivo de evaluación de los aprendizajes. En D. Leclercq & A. Cabrera (eds.), *Ideas e innovaciones: Dispositivos de evaluación de los aprendizajes en la educación* (pp. 35-50). Santiago de Chile: Editorial Universitaria.
- Leclercq, D. (2014c). El prisma de las características de un dispositivo de evaluación de los aprendizajes. En D. Leclercq & A. Cabrera (eds.), *Ideas e innovaciones: Dispositivos de evaluación de los aprendizajes en la educación* (pp. 51-82). Santiago de Chile: Editorial Universitaria.
- Leclercq, D. (2014d). Autoevaluación con grados de certeza: Un microscopio para la evaluación de los aprendizajes. En D. Leclercq & A. Cabrera (eds.), *Ideas e innovaciones: Dispositivos de evaluación de los aprendizajes en la educación* (pp. 329-356). Santiago de Chile: Editorial Universitaria.
- Leclercq, D. (2014e). Metacognición y test espectrales metacognitivos. En D. Leclercq & A. Cabrera (eds.), *Ideas e innovaciones: Dispositivos de evaluación de los aprendizajes en la educación* (pp. 221-242). Santiago de Chile: Editorial Universitaria.
- Leclercq, D. (2014f). Grados de certeza y docimología: Cómo calificar. En D. Leclercq & A. Cabrera (eds.), *Ideas e innovaciones: Dispositivos de evaluación de los aprendizajes en la educación* (pp. 357-386). Santiago de Chile: Editorial Universitaria.
- Leclercq, D. & Cabrera, A. (2011). Conceptos y modelos para concebir, analizar y evaluar innovaciones curriculares basadas en competencias. En *Redes de colaboración para la innovación en la docencia universitaria* (volumen 2, pp. 13-60). Talca: Universidad Católica del Maule.
- Leclercq, D. & Cabrera, A. (2014). *Ideas e innovaciones: Dispositivos de evaluación de los aprendizajes en la educación*. Santiago de Chile: Editorial Universitaria.
- Leclercq, D., De la Fuente, M. & Cabrera, A. (2014). Retroalimentación al estudiante sobre sus evaluaciones. En D. Leclercq & A. Cabrera (eds.), *Ideas e innovaciones: Dispositivos de evaluación de los aprendizajes en la educación* (pp. 431-452). Santiago de Chile: Editorial Universitaria.

- Leclercq, D., Gillet, P., Ercicum, M. & Cabrera, A. (2014). Evaluar la contribución de cada participante a un trabajo grupal. En D. Leclercq & A. Cabrera (eds.), *Ideas e innovaciones: Dispositivos de evaluación de los aprendizajes en la educación* (pp. 221-241). Santiago de Chile: Editorial Universitaria.
- Leclercq, D. & Poumay, M. (2005). *The 8 Learning Events Model and its principles. Release 2005-1*. Liège: Editions de l'Université de Liège.
- Leclercq, D. & Poumay, M. (2008). La metacognition. En D. Leclercq (ed.), *Psychologie éducationnelle de l'adolescent et du jeune adulte* (pp. 235-282). Liège: Editions de l'université de Liège.
- Marzano, R. J. (2001). *Designing a New Taxonomy of Educational Objectives. Experts in Assessment*. Thousand Oaks, CA: Corwin Press.
- O'Meara, K., Terosky, A. L. & Neumann, A. (2008). Faculty careers and work lives: A professional growth perspective. *ASHE Higher Education Report*, 34(3), 1-221.
- Piéron, H. (1963). *Examens et docimologie*. Paris: Presses universitaires de France.
- Poumay, M. & Maillart, Ch. (2014). Los portafolios: Hacia una evaluación más integrada y coherente con el concepto de desempeño complejo. En D. Leclercq & A. Cabrera (eds.), *Ideas e innovaciones: Dispositivos de evaluación de los aprendizajes en la educación* (pp. 243-258). Santiago de Chile: Editorial Universitaria.
- Scallon, G. (2004). *L'évaluation des apprentissages dans une approche par compétences*. Saint Laurent (QC): Editions du Renouveau Pédagogique.
- Schön, D. A. (1987). *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions*. San Francisco: Jossey-Bass.
- Tardif, J. (2006). *Evaluer les compétences: Documenter le parcours de développement*. Montréal: Chenelière Education.

La evaluación de competencias en la educación superior¹

■ José Rafael López Islas

RESUMEN

La universidad contemporánea puede encontrar en las competencias—un término que se originó en el ámbito organizacional— la oportunidad de equilibrar los requerimientos profesionales de un egresado que le permiten desempeñarse con efectividad en el corto y mediano plazo, con las necesidades formativas para un desempeño exitoso en el largo plazo.

La adopción de un modelo de evaluación de la educación universitaria basado en competencias, entendidas en un sentido amplio, tiene implicaciones importantes en el diseño del currículo, la labor de los docentes, los recursos de apoyo para el aprendizaje y la evaluación del estudiante. Medir las competencias de un estudiante requiere transformar el modelo de evaluación de tal forma que se evalúe el desempeño —y no solamente conocimientos— en contextos específicos. Para ello, se requiere establecer definiciones operacionales que traduzcan los conceptos inherentes a una competencia en referentes conductuales que puedan ser evaluados mediante la observación sistemática y registrados con el apoyo de herramientas tales como las rúbricas.

A partir de la experiencia institucional del Tecnológico de Monterrey, esta ponencia revisa las implicaciones de la evaluación de competencias en la educación superior en diferentes aspectos de la vida académica y plantea lecciones y recomendaciones que pudieran ser de interés para otras instituciones educativas.

¹ Las ideas presentadas en esta ponencia se derivan de una experiencia a gran escala realizada en el Tecnológico de Monterrey, bajo el liderazgo del autor en su función de Director de Efectividad Institucional. Aunque la descripción de los conceptos, procesos y resultados se sustenta en la experiencia institucional, las opiniones, posturas y conclusiones son responsabilidad enteramente del autor.

INTRODUCCIÓN

El objetivo primordial de esta ponencia es compartir con la comunidad académica las reflexiones derivadas de una experiencia institucional de largo alcance en materia de evaluación de competencias de egreso a nivel licenciatura (pregrado)² que se llevó a cabo bajo el liderazgo del autor. Si bien la experiencia tiene bases conceptuales derivadas de la literatura, esta ponencia enfatiza la reflexión sobre los aprendizajes institucionales y las posibles implicaciones que pudiera tener para otras instituciones.

El Tecnológico de Monterrey es una institución de educación superior privada, que cuenta con 26 campus a lo largo de la República Mexicana. Atiende a más de 55 mil estudiantes que cursan alguno de los 64 programas de licenciatura que constituyen su oferta académica, y a cerca de 12 mil estudiantes de posgrado tanto de maestría como de especialidad y doctorado. La institución fue fundada en 1943 y desde 1950 ha participado en procesos de acreditación internacionales, como la acreditación institucional de la Southern Association of Colleges and Schools Commission on Colleges (SACSCOC). A lo largo de su historia, ha participado también en diversos procesos de acreditación en México, tanto a nivel institucional como a nivel de programas en lo individual, así como de alcance internacional en varios de sus programas de ingeniería y negocios a través de ABET (conocida anteriormente con el nombre de Accreditation Board for Engineering and Technology) y AACSB (The Association to Advance Collegiate Schools of Business), respectivamente.

La institución se ha caracterizado históricamente por la formación sólida de profesionistas con un alto nivel de competitividad y con un perfil destacado en materia de liderazgo, emprendimiento y visión internacional. Para conseguirlo, ha mantenido de manera continua un enfoque hacia la innovación en materia educativa. Producto de ello, a finales de la década de los noventa, realizó cambios importantes en su modelo educativo, tomando como base la premisa de que la parte central del modelo debe ser el aprendizaje del alumno quien, lejos de ser un receptor de los esfuerzos del profesor, debe asumir un papel activo en su proceso de aprendizaje. El cambio a gran escala del modelo educativo trajo consigo diversas implicaciones, principalmente en cuanto a la labor de los docentes, la definición de los aprendizajes esperados y la evaluación de los mismos.

Paralelamente, en esa época se detonaba a nivel internacional una reflexión crítica acerca de una aparente incapacidad de las instituciones de educación superior para asegurar que sus

² En México, la educación superior ofrece títulos profesionales (nivel de licenciatura, conocida en otros países como pregrado) y grados de maestría, especialidad o doctorado (posgrado).

egresados contasen con un perfil idóneo que lograra satisfacer las nuevas condiciones requeridas por el mercado laboral. Diversas instituciones y, de manera destacada, la comunidad europea misma, iniciaron un gran esfuerzo por replantear el perfil de un egresado, incorporando al ámbito universitario el concepto de competencia que se utilizaba primordialmente en el ámbito laboral.

Asimismo, las agencias acreditadoras de los Estados Unidos se encontraban en una transformación sustantiva, dejando atrás la medición de la calidad académica basada primordialmente en indicadores de capacidad institucional, para privilegiar como criterio de calidad la demostración del logro de resultados, especialmente en cuestión del aprendizaje de los estudiantes. Este cambio implicaba para las instituciones definir los aprendizajes esperados en cada uno de los programas de estudio y desarrollar mecanismos para medirlos.

Así, mientras que el Tecnológico de Monterrey evolucionaba su modelo educativo en respuesta a su necesidad de innovar, el entorno externo le sentaba las bases para concretar y poner a prueba el enfoque de su modelo encaminado a responder a las demandas de la sociedad, que de paso permitiera participar con éxito en los procesos de acreditación internacional. Los primeros programas que participaron fueron aquellos acreditados por ABET y por AACSB; posteriormente fueron incluidos todos los programas, a raíz de que SACSCOC estableció como criterio de acreditación la evaluación de los programas académicos. Hacia finales de la década de 2000, la institución estaba ya inmersa en un proceso de evaluación de sus programas académicos basado en competencias de egreso. Luego, en la revisión de la visión institucional y en la planeación estratégica hacia el año 2020, el Tecnológico de Monterrey estableció, como una de sus iniciativas estratégicas, el desarrollo del Modelo Educativo Tec21, que representa, más que una evolución, una transformación aún más profunda de lo que había desarrollado durante la década anterior.

Las siguientes secciones se enfocan en describir y analizar la experiencia del Tecnológico de Monterrey entre los años 2009 y 2016, período a lo largo del cual la institución incorporó la evaluación de competencias de egreso como parte esencial de su modelo educativo.

EL MODELO Y SU SUSTENTO TEÓRICO

Para poder entender el modelo educativo desarrollado e implementado por el Tecnológico de Monterrey, se requiere conocer los fundamentos en que se sustenta. Esta sección revisa los conceptos y las definiciones pertinentes para el modelo educativo de dicha institución.

Competencias, resultados de aprendizaje y la educación basada en competencias

En esta sección, se examina los conceptos relacionados con las competencias desde tres ángulos: el origen y uso del término competencia en el ámbito laboral; la adopción de los conceptos competencia y resultados de aprendizaje en el campo de la educación como instrumentos para definir los perfiles de egreso de los programas académicos, incorporando además las necesidades formativas que demanda el mercado profesional; y la educación basada en competencias como un modelo educativo que rompe en aspectos clave con el paradigma tradicional de enseñanza.

Desde un punto de vista teórico, el concepto de competencia ha sido y continúa siendo un tema ampliamente discutido. Las definiciones, por un lado, y más aún la discusión sobre su pertinencia en el ámbito de la educación superior, por el otro, han dado pie a una literatura extensa que, sin embargo, no parece llevar a acuerdos entre los especialistas. Conceptos tales como “competencia”, “resultado del aprendizaje”, “objetivo del aprendizaje”, “criterio de desempeño”, tienen tantas definiciones como autores los abordan y no es raro encontrar que se usan de manera intercambiable o incluso contradictoria.

No se pretende con este comentario restar mérito al trabajo teórico ni al debate entre especialistas. Por el contrario, se reconoce que el diseño e implementación de un programa de evaluación de competencias en la educación superior debe contar con un fundamento teórico-metodológico sólido. Sin embargo, se propone, sobre la base de la experiencia institucional del Tecnológico de Monterrey, no esperar a resolver el debate teórico como condición para iniciar un programa de esta naturaleza. Basta con tener un marco conceptual básico que cubra los conceptos esenciales para diseñar y echar a andar el programa y una apertura intelectual para aprender continuamente y generar teoría a partir de la experiencia.

El término “competencia” ha sido usado desde hace tiempo en el campo de los recursos humanos y fue incorporado al contexto educativo hace alrededor de un par de décadas. Los profesionales de los recursos humanos entienden por competencia aquellas habilidades medibles que permiten a una persona obtener resultados en sus labores y avanzar hacia niveles de desempeño superiores (Gosselin, Cooper & Bonnstetter, 2013). De acuerdo con diversos estudios (Lumina Foundation, 2014; ManpowerGroup, 2015, por citar algunos recientes), existe la percepción entre los sectores productivos de que las instituciones de educación superior no están desarrollando en sus graduados las competencias requeridas para el mercado laboral actual. Más aún, se han desarrollado propuestas concretas para que las universidades desarrollen ciertas competencias en todos sus graduados argumentando que, sobre la base de las tendencias

organizacionales, serán más importantes que las competencias propias de una disciplina, pues resultarán indispensables para poderse desempeñar con éxito en el futuro (Partnership for 21st Century Skills, 2011).

En el campo educativo, las competencias son definidas como la combinación de conocimientos, habilidades, actitudes y valores que hacen posible que una persona tenga éxito en una tarea o función en contextos específicos (Blanco Fernández, 2009). Un concepto relacionado es el de resultados de aprendizaje, que son enunciados que establecen de manera específica qué es lo que un estudiante debe ser capaz de hacer, expresados en términos observables, como resultado de su participación en una actividad educativa (Hartel & Foegeding, 2004). La literatura es extensa y existen múltiples definiciones para estos conceptos que podrían ser abordados en otros trabajos académicos.

Una tendencia que va cobrando importancia en el ámbito de la educación, incluida la superior, es la educación basada en competencias. Este modelo educativo no solamente establece los aprendizajes de los estudiantes sobre la base de competencias, sino que cambia la forma en que se lleva a cabo la interacción entre los estudiantes y la universidad. En vez de operar sobre la base temporal de períodos académicos y asignaturas establecidas en una malla curricular, la educación basada en competencias permite que el estudiante avance a su propio ritmo, en la medida que va demostrando tener las competencias requeridas para obtener un grado académico (Johnstone & Soares, 2014; Argudín, 2006). La instrucción contribuye al desarrollo de competencias, pero el estudiante la utiliza con flexibilidad en la medida que la requiere y no bajo un diseño preestablecido. La evaluación se basa, lógicamente, en la medición formal de competencias, que pueden incluso ser certificadas, y no requiere que un estudiante haya cursado asignaturas para poder ser evaluado, toda vez que una competencia también puede ser desarrollada mediante la experiencia laboral (EDUCAUSE, 2014).

Es importante tomar en cuenta que una institución o programa académico puede incorporar la noción de competencia para definir el perfil de egreso sin que deba necesariamente adoptar un modelo educativo basado en competencias, como se describió en el párrafo anterior. Esto ocurre, por ejemplo, con muchas instituciones educativas acreditadas en los Estados Unidos por agencias tales como SACSCOC —que acredita universidades— y ABET —que acredita programas en las áreas de ingeniería y tecnología—. Estas agencias, al igual que otras más, requieren, desde la década de 2000, que los programas académicos hagan lo siguiente: definir sus perfiles de egreso sobre la base de resultados de aprendizaje de los estudiantes (*student learning outcomes*); medir de manera periódica y sistemática qué tanto los estudiantes los han alcanzado

al momento de su graduación; y tomar decisiones de mejora para los programas sobre la base de los resultados, cerrando de esta manera ciclos de mejora continua. Las instituciones o programas no deben adoptar un modelo educativo basado en competencias para ser acreditados, pero sí deben definir los perfiles de egreso sobre la base de *student learning outcomes*.

Típicamente, la forma en que los *student learning outcomes* son expresados es consistente con lo que se ha descrito en párrafos anteriores como competencias, que, para poder ser medidas, requieren ir acompañadas por criterios de desempeño, entendidas como aquellas conductas observables que demuestran que un estudiante cuenta con una competencia. Es decir, los criterios de desempeño, en esta acepción, corresponden a la definición antes referida de resultado de aprendizaje. Esta conceptualización es el fundamento para el modelo de evaluación utilizado por el Tecnológico de Monterrey. Debe mencionarse que, aunque se ha explorado la posibilidad de adoptar un modelo educativo basado en competencias, aún no están dadas las condiciones para ello. Sin embargo, el mero hecho de definir el perfil de egreso sobre la base de competencias tiene implicaciones importantes en diversas áreas del quehacer educativo que se discuten con detalle en secciones posteriores de esta ponencia.

Competencias de egreso y criterios de desempeño

Tomando como base la literatura, pero buscando simplificar la comunicación con la comunidad educativa, el Tecnológico de Monterrey definió dos conceptos que sirven como base para el programa de evaluación. El primer concepto es “competencia de egreso”, entendido como la combinación de atributos —con respecto al conocimiento y su aplicación, a las actitudes y a las responsabilidades— que hacen referencia a las capacidades que un estudiante debe haber desarrollado al momento de su graduación. Se decidió utilizar el término compuesto (“competencia de egreso”) para enfatizar que esta integra componentes cognitivos, procedimentales, actitudinales y de valores, que el estudiante desarrolla a lo largo de la experiencia educativa, de tal forma que pueda demostrar, cuando está próximo a graduarse, que es capaz de aplicarla en diversos contextos.

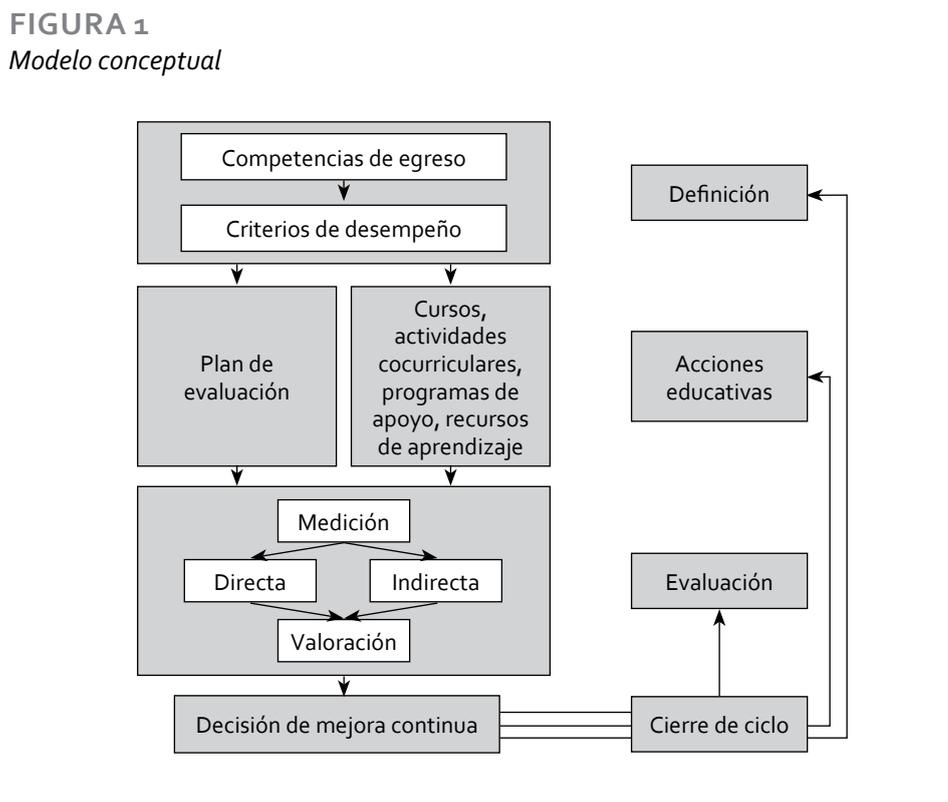
Para poder determinar si un estudiante ha desarrollado una competencia de egreso, se requiere haber definido “criterios de desempeño”; es decir, es necesario contar con enunciados específicos medibles basados en comportamientos observables que evidencian el desarrollo de una competencia. Para cada competencia de egreso, se define al menos dos criterios de desempeño.

Estos dos conceptos —“competencias de egreso” y “criterios de desempeño”— son la base para el modelo de evaluación de competencias del Tecnológico de Monterrey que se describe a continuación.

MODELO DE IMPLEMENTACIÓN

Para incorporar la evaluación de competencias de egreso en todos sus programas, el Tecnológico de Monterrey adoptó un modelo que consta de cuatro componentes: definición, elaboración de plan de evaluación (en paralelo a las acciones educativas), evaluación y cierre de ciclo de mejora (ver figura 1). Este modelo conceptual sirvió como base para la elaboración de un modelo de implementación específico que permitió trabajar con los docentes y directivos. En esta sección, se detalla cada uno de ellos como etapas de un modelo de implementación, discutiendo además algunas consideraciones derivadas de la experiencia de ponerlo en marcha.

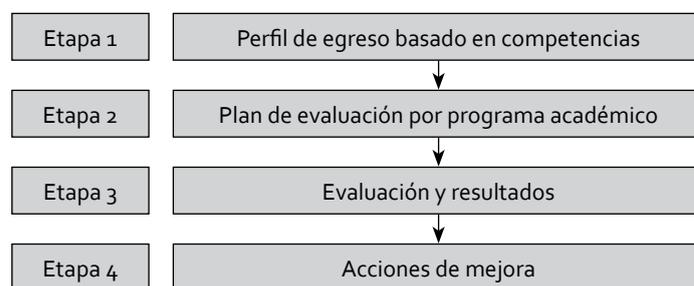
FIGURA 1
Modelo conceptual



Un modelo en cuatro etapas

El modelo de implementación de la evaluación de competencias de egreso adoptado por el Tecnológico de Monterrey consta de cuatro etapas (ver figura 2). En la primera etapa, se definen los resultados esperados en cada uno de los programas (competencias de egreso) con los criterios de desempeño que permiten medirlas. En la segunda, cada programa elabora un plan para realizar sistemáticamente la medición de las competencias de egreso. En tercer lugar, se ejecuta el plan de medición, lo que permite generar información sobre el desempeño de los estudiantes y realizar una valoración del nivel de logro de cada programa con respecto a los resultados esperados. Finalmente, sobre la base de los resultados de la evaluación, se establece y ejecuta acciones de mejora que permitan acercarse de manera progresiva a un mayor nivel de logro de las competencias de egreso. Consistente con la filosofía de mejora continua, el cierre de un ciclo representa el inicio de uno nuevo.

FIGURA 2
Modelo de implementación



■ ETAPA 1: DEFINIR PERFILES DE EGRESO BASADOS EN COMPETENCIAS

El punto de partida para poder evaluar sobre la base de competencias de egreso consiste en definir dichas competencias para cada uno de los programas académicos. Como se discutió anteriormente, estas competencias de egreso reflejan las capacidades que los estudiantes de un programa deben haber desarrollado para el momento de su graduación. Estas capacidades van más allá de lo meramente cognitivo e incluyen también habilidades, así como actitudes y valores, y hacen referencia a un contexto específico de aplicación.

Realizar una definición de esta naturaleza, con competencias de egreso que son transversales al currículo, implica una complejidad muy grande. Por ello, las primeras definiciones de competencias se realizaron con programas que se encontraban en un proceso de acreditación internacional, razón por la cual algunos profesores habían recibido entrenamiento por parte de las agencias acreditadoras en materia de definición de perfiles de egreso basados en competencias. Estas primeras experiencias sirvieron como una experiencia piloto para la institución, cuyos resultados fueron de suma importancia para la implementación a gran escala.

Tomando como base los resultados de estas primeras experiencias y después de analizar las tendencias de las diversas agencias acreditadoras y las experiencias de otras instituciones, se diseñó una estrategia que incluyera a todos los programas académicos. Dado que cada cinco años la institución revisa sus planes de estudio, en la revisión que llevó a nuevos planes entre 2010 y 2011, se incluyó, como primer paso del diseño curricular, la definición de las competencias de egreso de cada programa. Es decir, el perfil de egreso se estableció sobre la base de competencias de egreso que incluían también los criterios para su medición.

Este paso resultó fundamental, pues requirió que los profesores llegaran a acuerdos sobre los resultados esperados de su programa, que van más allá de los objetivos específicos de aprendizaje de las asignaturas que cada quien enseña. Así, al momento de definir la malla curricular, fue posible asegurar la contribución que cada una de las asignaturas debería tener en el logro de las competencias de egreso acordadas para el programa. Es importante mencionar que este ejercicio se enfocó a las competencias de egreso propias de la disciplina y que las competencias de egreso comunes a todos los estudiantes de la institución fueron definidas por equipos multidisciplinarios, bajo la coordinación de la dirección de Efectividad Institucional de la Vicerrectoría Académica.

El ejercicio de definición de las competencias de egreso tuvo como fundamento la opinión de egresados y empleadores que forman parte de los consejos consultivos de los programas³, así como los datos de los estudios institucionales de seguimiento a los graduados y los indicadores de desempeño académico de cada programa. Asimismo, se utilizó como referencia los lineamientos de agencias acreditadoras y sociedades profesionales. Tomando como base esta información, se pidió a los profesores describir las capacidades que debería haber desarrollado un estudiante al momento de su graduación. De manera iterativa, el grupo fue trabajando en un

3 En el Tecnológico de Monterrey, es común que los programas o departamentos académicos cuenten con consejos consultivos, integrados por graduados y empleadores, que tienen una función de asesoría con respecto a las tendencias profesionales y del campo laboral, y opinan sobre decisiones académicas importantes, tales como cambios en el currículo.

documento que describía entre seis y ocho competencias de egreso disciplinares por programa, hasta llegar a un versión final que reflejaba el acuerdo del grupo.

Asimismo, con el propósito de facilitar la evaluación de las competencias de egreso, se pidió a los profesores que, una vez que hubiesen acordado las competencias de egreso, describieran los criterios de desempeño; es decir, aquellos comportamientos observables que permitiesen evaluar si un graduando tiene, y en qué nivel de desarrollo, cada una de las competencias de egreso. Típicamente, para cada competencia de egreso fueron identificados al menos dos criterios de desempeño. Así, al finalizar esta etapa, cada programa contaba con un perfil de egreso expresado en competencias con sus correspondientes criterios de desempeño.

■ ETAPA 2: ELABORAR LOS PLANES DE EVALUACIÓN DE LOS PROGRAMAS ACADÉMICOS

Una vez definidas las competencias de egreso con sus respectivos criterios de desempeño, el comité responsable del diseño curricular procedió al armado de la malla curricular. Una vez que los nuevos planes fueron aprobados, empezaron a impartirse los programas rediseñados. Paralelamente, se pidió a cada director de programa elaborar un plan de evaluación. Este plan consiste en construir una tabla donde se establece, para cada criterio de desempeño, al responsable de la evaluación (generalmente un profesor o el director del programa), así como la fecha y el método con que se debe llevar a cabo la evaluación, además de los criterios para seleccionar a los estudiantes a evaluar. Los planes son específicos y generalmente incluyen ciclos de evaluación anuales para los criterios de desempeño, aunque por lo común se evalúan de manera alternada cada semestre con el fin de balancear la carga de evaluación de los docentes.

Una consideración importante es que aunque los planes de estudio fueron modificados, lo que implicó que algunas asignaturas específicas fueran añadidas o eliminadas, se encontró que la mayoría de las competencias de egreso aplicaban tanto para los planes anteriores como para los rediseñados. Esta estabilidad puede ser atribuida a que la esencia de los programas se mantiene y lo que cambia es la manera en que se busca desarrollar las competencias de egreso. Contar con competencias de egreso estables permitió empezar a realizar la evaluación de las mismas aún y cuando no habían entrado en vigor los nuevos programas en su totalidad.

■ ETAPA 3: LLEVAR A CABO LA EVALUACIÓN DE LAS COMPETENCIAS DE EGRESO

La tercera etapa del modelo de implementación consiste en asegurar que los planes de evaluación establecidos en la etapa 2 se ejecuten tal y como están acordados. Para llevarla a cabo con éxito, se

requiere el compromiso de los responsables de los programas, así como el apoyo de una entidad central que facilite el proceso y asegure su ejecución de manera homogénea entre los programas.

La evaluación se fundamenta en dos tipos de medición: indirecta y directa. La medición indirecta implica hacer una estimación del nivel de dominio de una competencia de egreso mediante una opinión propia (autopercepción) o la de un tercero. Generalmente, estas opiniones se recogen a través de encuestas que pueden ser aplicadas bajo diversas modalidades (en persona, en línea, por teléfono). Estas mediciones proveen información valiosa, pero no son suficientes, pues tienen un sustento principalmente subjetivo.

Por su parte, la medición directa implica constatar de primera mano que un estudiante exhibe los comportamientos observables tal y como se describen en los criterios de desempeño. En el caso de componentes cognitivos, puede hacerse la medición mediante exámenes; pero, para componentes de habilidad y actitud, se requiere realizar observaciones en contextos de aplicación similares a las situaciones en las que se espera que el estudiante se desempeñe profesionalmente. A este tipo de evaluación se le suele conocer bajo el concepto de “evaluación auténtica” (*authentic assessment*) o “evaluación basada en desempeño” (*performance-based assessment*) (Wiggins, 1993; Meyer, 1992; Díaz-Barriga, 2006), en contraposición a una evaluación tradicional basada primordialmente en exámenes.

Para facilitar la observación y aumentar la confiabilidad en la aplicación de los criterios entre diversos evaluadores, típicamente se utilizan “rúbricas” diseñadas específicamente para cada competencia de egreso, que sirven como marco común para la medición, por lo que promueven una mayor homogeneidad en la aplicación de los criterios por parte de los jueces.

En el Tecnológico de Monterrey, se utiliza dos modalidades para la evaluación directa. Una, conocida en inglés como *course embedded assessment* (medición inserta en los cursos), implica utilizar productos del aprendizaje desarrollados en el contexto de asignaturas específicas, como evidencia de que el estudiante que elaboró el producto ha desarrollado alguna competencia de egreso. La medición del nivel de desarrollo de la competencia es independiente de la evaluación que se hace de la tarea en el contexto de la asignatura e incluso, en muchas ocasiones, la medición se realiza una vez que ha terminado el curso, por lo que no afecta a la calificación que obtiene el estudiante. Para que este modelo pueda funcionar, se requiere que los productos del aprendizaje exijan la integración de conocimientos, habilidades y actitudes tal y como están descritas en la competencia a evaluar. Típicamente, las asignaturas más propicias para generar productos del aprendizaje aptos para evaluar competencias de egreso son cursos conocidos como “integradores” en los que los estudiantes realizan proyectos complejos en un contexto de

aplicación realista y para cuya solución requieren hacer uso de los conocimientos y habilidades adquiridos a lo largo de todo el currículo.

Otra modalidad para la medición directa de las competencias es lo que en inglés se conoce como *assessment center* o *assessment day* (centro o día de la evaluación). En esta modalidad, en vez de depender de evidencias generadas en el contexto de ciertas asignaturas, se organiza una sesión en la que los estudiantes deben resolver situaciones complejas para las cuales requieren un nivel adecuado de dominio de las competencias de egreso. A partir de problemas cuidadosamente diseñados, los estudiantes son observados por expertos durante el proceso de solución de los problemas. Esto permite que la evaluación se refiera no solamente a la calidad de la solución que proponen al problema, sino también al proceso que siguen para llegar al resultado, incluyendo su capacidad de trabajar colaborativamente, así como sus habilidades de comunicación oral y escrita.

Una sesión de este tipo es ideal para medir las competencias de egreso, pero requiere varias condiciones. Primero, se debe poder programar un evento en el que participe al menos una muestra representativa de los estudiantes de un programa, próximos a graduarse; para lo cual se requiere encontrar una motivación no solamente para que participen, sino para que hagan su mejor esfuerzo. Segundo, es necesario encontrar un equipo de expertos que incluya no solamente profesores, sino profesionales externos para que participen como jueces en sus áreas de especialidad. Tercero, se debe contar con un diseño de la sesión y de los instrumentos de medición que permitan exponer a los estudiantes a situaciones donde puedan demostrar su nivel de desarrollo de las competencias. Finalmente, se requiere un procedimiento de registro de la información resultante que de manera eficiente y confiable permita realizar una valoración de los niveles de desarrollo de las competencias de los estudiantes de un programa.

En el Tecnológico de Monterrey, la modalidad más común para la evaluación es la medición inserta en cursos, aunque cada vez de manera más frecuente se utiliza la modalidad de centro o día de la evaluación. Para ello, se ha encontrado que un buen incentivo para que los estudiantes participen con interés y para que los jueces externos dediquen una mañana o todo un día a la actividad de evaluar estudiantes, es la percepción de que la sesión se convierte en un evento de reclutamiento de talento. Es decir, los estudiantes saben que están siendo evaluados por profesionales que se desempeñan en el campo de acción, quienes pueden identificar de primera mano posibles recursos humanos para sus organizaciones, lo que también es una motivación para los jueces.

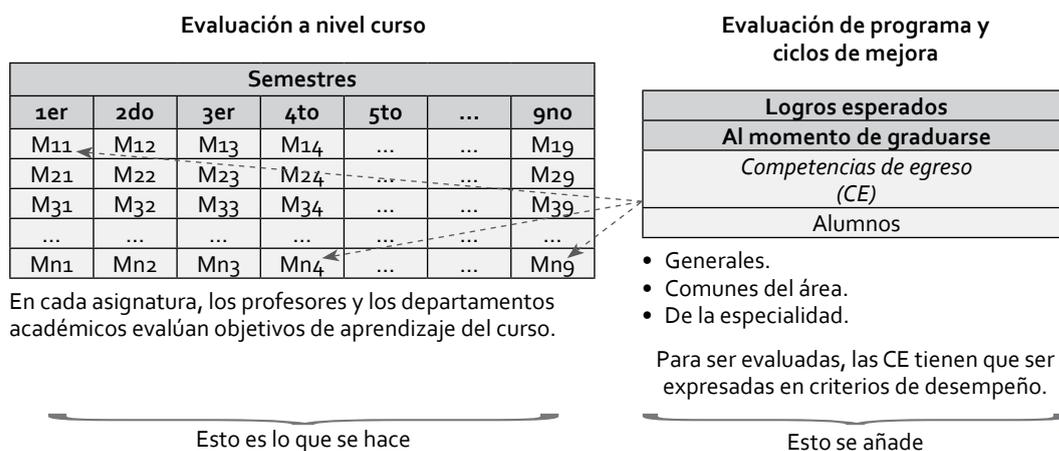
Los resultados de las mediciones individuales no sirven a menos que se haga una valoración del nivel de dominio que exhiben en general los estudiantes de un programa en cada una de las

competencias de egreso. Esta valoración se hace sobre la base de metas establecidas previamente que reflejan los niveles mínimos que un programa establece para cada competencia y que se pueden ir moviendo en la medida en que un programa va mejorando. Así, sobre la base de los resultados de los criterios de desempeño de una competencia de egreso determinada, en la valoración se puede establecer que el programa, en el ciclo actual, obtuvo un nivel de X%, que se encuentra Y puntos abajo o arriba de la meta W establecida al inicio del ciclo.

Un aspecto muy importante a tomar en cuenta es que, en el modelo de implementación del Tecnológico de Monterrey, la evaluación de las competencias de egreso se hace para el programa y no para alumnos en específico. Es decir, si bien se realizan mediciones individuales de los alumnos, la valoración del nivel de dominio de las competencias se hace tomando como unidad de análisis al programa. Por esta razón, es posible realizar la medición con muestras de estudiantes siempre y cuando sean representativas de la población.

La figura 3 muestra cómo la evaluación de los programas es una capa que se añade a la evaluación que regularmente se realiza en cada una de las asignaturas. Es decir, representa un esfuerzo adicional para los profesores, para los programas y para la administración, que se justifica en tanto que genera información sobre el desempeño de los programas que no se puede obtener mediante la acumulación de las evaluaciones de los estudiantes en cada una de las asignaturas que cursan.

FIGURA 3
Evaluación en las asignaturas y a nivel de programa



■ ETAPA 4: IDENTIFICAR Y LLEVAR A CABO ACCIONES DE MEJORA

Una vez que se realiza una valoración del nivel de logro de un programa en cada una de las competencias de egreso (e incluso, en cada uno de los criterios de desempeño), los profesores que forman parte del programa analizan los resultados y acuerdan acciones de mejora. Esta etapa es de suma importancia, pues el modelo se fundamenta en la filosofía de la mejora continua. Las acciones de mejora se pueden referir a cualquier aspecto relacionado con el programa, o incluso al proceso de evaluación mismo. Algunas acciones se encaminan a modificaciones menores en las asignaturas, pero otras pueden representar cambios sustantivos en el currículo. Otras acciones se refieren a la forma de impartir los cursos y a las actividades de aprendizaje utilizadas. En algunos casos, se identifica la necesidad de incluir más actividades de laboratorio u otros contextos de aplicación del conocimiento, especialmente cuando se trata de habilidades más que de conceptos. Es posible también que, como resultado de la evaluación, se identifique la necesidad de ajustar los mecanismos o criterios de la misma. Es decir, las mejoras pueden referirse a cualquier aspecto del programa, incluyendo temas financieros o de gestión si estos implican limitaciones para el aprendizaje de los estudiantes.

Esta etapa es crítica, pues da sentido al esfuerzo de evaluación. Es decir, si no se utilizan los resultados para tomar decisiones de mejora y llevarlas a cabo, todo el trabajo realizado a lo largo de las etapas anteriores se convierte en un mero ejercicio de evaluación *per se* que no conlleva un beneficio tangible para los programas y, muy en especial, para el aprendizaje de los estudiantes. Asimismo, el no cerrar un ciclo de evaluación frena el inicio de un nuevo ciclo que lleve a la institución a un mayor logro de los objetivos que se propone.

Implementar un modelo como el aquí descrito tiene implicaciones importantes en diversas áreas del quehacer institucional. Estas implicaciones se discuten en la siguiente sección.

IMPLICACIONES

La implementación a gran escala de un modelo de evaluación de competencias de egreso, como el del Tecnológico de Monterrey, permite identificar, desde la experiencia, algunos aspectos de interés para el campo y en particular para otras instituciones que deseen incursionar en este camino, que se discuten en las siguientes secciones.

Se requiere encontrar nuevas formas de facilitar el aprendizaje _____

Es común que los perfiles de egreso de los programas académicos hagan referencia a capacidades de los egresados que van más allá del aprendizaje puntual de aspectos

cognitivos, pues son elaborados primordialmente sobre la base de criterios mercadológicos. Sin embargo, cuando se declaran, definen y evalúan las competencias de egreso; es decir, cuando se adquiere el compromiso con el público de evaluar la efectividad de los programas, se vuelve evidente que un modelo de aprendizaje tradicional basado primordialmente en la impartición de cátedras resulta insuficiente. Las competencias de egreso incluyen el aprendizaje a nivel cognitivo, pero requieren también el desarrollo de habilidades y actitudes que difícilmente se pueden aprender si se depende principalmente de un modelo educativo basado en la exposición de los docentes en clase y la consulta de libros de texto y otros recursos bibliográficos.

Esto significa que incorporar la evaluación de competencias de egreso va más allá de aumentar la carga laboral a los docentes por el lado de las actividades de evaluación. Implica, sin lugar a dudas, evolucionar el modelo mismo de enseñanza hacia uno más complejo, toda vez que los aprendizajes se vuelven, en sí mismos, también más complejos. Esta implicación es de la mayor relevancia, pues representa una transformación de la manera en que los profesores diseñan y llevan a cabo las actividades de aprendizaje en sus asignaturas. En particular, el docente debe diseñar actividades que trascienden la exposición de conceptos y la solución de problemas que se refieren a situaciones abstractas, lejanas al contexto de aplicación. El docente requiere acercar al estudiante a situaciones concretas del quehacer profesional, para lo cual es necesario salir del espacio conceptual del salón de clases, lo que demanda del docente una mayor vinculación con su entorno. No se quiere decir con esto que el salón no sea un espacio adecuado para aprender, sino que se requiere enriquecer la experiencia de aprendizaje trayendo al salón de clases situaciones que se asemejan de una mejor manera a la realidad profesional en la que los graduados aplicarán las competencias adquiridas a su paso por la universidad.

Por ello, modelos como el aprendizaje invertido (*flipped learning*) (Fulton, 2014; Goodwin & Miller, 2013) son cada vez más utilizados. En este modelo, se invierte la forma en que se distribuye el tiempo que un alumno destina al aprendizaje. En un modelo tradicional, el estudiante dedica el tiempo en el salón de clases con su docente y sus compañeros para recibir lecciones y aprender conceptos, mientras que la resolución de problemas y otras actividades de aprendizaje se realizan fuera del salón. En el modelo de aprendizaje invertido, el estudiante aprende por cuenta propia conceptos fuera de clase y viene al salón a trabajar en la solución, de manera conjunta con sus compañeros y bajo la tutela del profesor, de situaciones problemáticas donde desarrolla habilidades y forma actitudes, aplicando los conceptos adquiridos fuera de clase.

Se requiere construir un currículo con un mayor nivel de integración

Como se dijo anteriormente, las competencias de egreso no se desarrollan en una sola asignatura. Es decir, su desarrollo ocurre a lo largo del currículo, por lo que es necesario incorporar ejes transversales. El currículo deja de ser una colección de asignaturas impartidas de manera mayoritariamente autónoma por los docentes y refleja el resultado de la colaboración entre estos, que se manifiesta en acuerdos que trascienden las asignaturas específicas que imparten. Conseguir acuerdos de esta naturaleza demanda un liderazgo académico sólido y la capacidad de negociar y facilitar el entendimiento entre expertos cuyas áreas de enfoque son muy diversas.

Como ventaja, sin embargo, se puede llegar a conseguir una apropiación del currículo por parte de los docentes, como un todo, y no solamente por la asignatura de la cual son responsables. Esto no es fácil e implica un cambio sustantivo en la manera en que se considera el currículo y las responsabilidades de los docentes en el logro de los aprendizajes de los estudiantes.

La situación se vuelve un poco más compleja si se toma en cuenta que las competencias de egreso que debe desarrollar el estudiante no son solamente las propias de la disciplina. Si un programa establece, por ejemplo, que sus graduandos deben ser capaces de trabajar en equipo, en algún momento de sus estudios deberían tener un espacio para desarrollar esta capacidad. Generalmente, aunque se consideran importantes las competencias generales, conocidas como *soft skills* (habilidades suaves, en contraposición a las “duras” o propias de la disciplina), no se tiene asignado en el currículo un espacio para su desarrollo. Por ello, cuando una institución empieza a evaluarlas se encuentra con que su nivel de desarrollo es muy dispar entre los estudiantes y que, aunque pasen los años, no se ven mejoras, pues no hay una acción encaminada a aumentar el aprendizaje de los alumnos.

Conseguir que los alumnos desarrollen las competencias generales requiere que los docentes incorporen en sus asignaturas actividades de aprendizaje enfocadas específicamente al desarrollo de dichas competencias. En ese sentido, un modelo como el de aprendizaje invertido permite trabajar con los estudiantes en el desarrollo de las competencias generales, pues el docente tiene oportunidad de verlos en acción, de modelar conductas y de retroalimentar prácticas de los alumnos más allá de los aprendizajes propios de la disciplina. Más aún, la malla curricular debe especificar también líneas transversales para el desarrollo de las competencias generales, identificando aquellas asignaturas específicas en las que se habrá de incorporar actividades de aprendizaje encaminadas a desarrollar cada una de las competencias de egreso.

Las responsabilidades de los profesores se transforman

Una manera diferente de enseñar, de concebir el currículo y de evaluar el aprendizaje no puede implementarse sin cambios importantes en la labor de los docentes. Un modelo curricular basado en competencias de egreso requiere de la apertura, disposición y compromiso del profesorado para abrazar el cambio. Para conseguirlo, la institución debe contar con una estrategia apropiada de administración del cambio. Con el fin de romper la inercia y echar a andar el modelo lo más pronto posible, es recomendable identificar a los docentes que tienden a ser los primeros en adoptar las innovaciones, apoyarles en las etapas iniciales de implementación y recompensarles mediante adecuaciones al esquema de incentivos, de ser necesario. Es importante tomar conciencia de que, en las primeras etapas, el riesgo de cometer errores es alto por lo que se debe blindar a los profesores innovadores y establecer con claridad la expectativa de que, aunque se desea que el modelo sea un éxito, es posible que en sus inicios ocurran fallas que, lejos de ser consideradas un fracaso, deben ser utilizadas como experiencias valiosas de aprendizaje.

Se requiere una forma sustancialmente diferente de evaluar el aprendizaje

Si el aprendizaje es diferente, se requiere una evaluación más compleja. Los componentes cognitivos de las competencias se siguen evaluando con herramientas tales como los exámenes en sus diferentes modalidades. Sin embargo, para evaluar habilidades se requiere diseñar y llevar a cabo pruebas para medir el desempeño en un contexto de aplicación específico, que típicamente utilizan la observación por parte de expertos para estimar el nivel de desempeño del estudiante. Con el fin de aumentar la confiabilidad de la medición, generalmente se usan instrumentos de registro conocidos como rúbricas que, si están bien construidas, reducen el margen de interpretación y aseguran que se evalúan siempre las mismas dimensiones de la habilidad a medir.

La construcción de rúbricas y su utilización para la evaluación de competencias requiere de entrenamiento y práctica. El principio detrás de una rúbrica es sencillo, pero el desarrollo de rúbricas que describan de manera específica, clara y balanceada las conductas observables que debe exhibir el sujeto evaluado requiere de cierto dominio metodológico. Asimismo, aunque el trabajo de un evaluador se simplifica cuando se utiliza una rúbrica, es necesario contar con entrenamiento para utilizar el instrumento de manera apropiada, asegurando la aplicación de criterios homogéneos al momento de evaluar.

Como se explicó anteriormente, a nivel macro, la evaluación de las competencias de egreso representan una actividad adicional que no sustituye la evaluación en cada asignatura, sino que

la complementa para hacer mediciones a nivel programa de competencias de egreso que, como se ha dicho, son transversales al currículo. Es importante mencionar que, en la experiencia del Tecnológico de Monterrey, la evaluación es a nivel del programa y no a nivel del estudiante o del profesor. Es decir, la unidad de análisis del modelo es el programa. Eventualmente, cuando el modelo madure, será posible incorporar la evaluación de cada uno de los estudiantes e incluso generar reportes individuales que certifiquen que el estudiante ha demostrado un nivel apropiado en cada una de las competencias. Esta opción, aunque deseable, implica una complejidad y una madurez del proceso que aún tomará algunos años en conseguir.

Asimismo, aunque el modelo actual incluye el cierre de ciclos de mejora cuando se habla a nivel de los programas, desde la perspectiva de los estudiantes la evaluación es de carácter “sumativo”. Es decir, debido a que se le evalúa primordialmente hacia el final de sus estudios, el estudiante no alcanza a recibir los beneficios de las mejoras que puedan derivarse de la evaluación. En ese sentido, un reto que asumirá la institución en los próximos años es incluir en la evaluación no solamente a todos los estudiantes, como se mencionó en el párrafo anterior, sino incorporar ejercicios de evaluación “formativa” a lo largo del programa y no solamente hacia el final, de tal manera que los estudiantes reciban información sobre su avance y puedan asegurar el cumplimiento de las competencias de egreso en lo que resta de sus estudios.

Se necesita contar con un sólido apoyo informático

El seguimiento a la evaluación de las competencias de egreso puede llegar a ser complejo, no solo a nivel institucional, sino también en cada uno de los programas. Por ello, es recomendable utilizar alguna plataforma tecnológica que documente las competencias de egreso de cada programa, así como los planes de evaluación y los resultados de la misma a lo largo de ciclos de mejora continua. De esta manera, es posible dar seguimiento a los procesos de evaluación de cada una de las competencias de egreso de los programas, desde que se definen hasta que se acuerdan y documentan acciones de mejora resultantes de la evaluación en cada uno de los ciclos. Esto se hace particularmente indispensable cuando se trata de una institución con la complejidad geográfica y del tamaño del Tecnológico de Monterrey, aunque en cualquier caso el uso de herramientas informáticas puede contribuir a simplificar el seguimiento de todo el proceso.

El Tecnológico de Monterrey buscó en su momento una herramienta disponible en el mercado que sirviera para este fin y, dado que no encontró alguna que se adecuara a las características de la institución, desarrolló un sistema propio, conocido como SAEP (Sistema para la Administración de la Evaluación de los Programas). En este sistema están registradas todas las

competencias de egreso de cada uno de los programas, con sus respectivos criterios de desempeño, desde el año en que inició a gran escala el proyecto de evaluación de los programas académicos. Para cada competencia de egreso y criterio de desempeño, se tiene el registro del plan de evaluación correspondiente, así como de los resultados de la evaluación y de las acciones de mejora acordadas, una vez que se analizaron los resultados.

Asimismo, el SAEP cuenta con un módulo conectado con el sistema general de información de los estudiantes, así como con el repositorio de rúbricas de evaluación, que permite que cada profesor pueda evaluar trabajos en línea, quedando la evaluación debidamente registrada en el sistema. El SAEP hace posible también elaborar reportes del avance puntual de la evaluación de las competencias de egreso de cada programa, tanto para cada ciclo como agregado a lo largo de los años.

En la actualidad, existen varias opciones disponibles en el mercado que una institución que recién inicie la evaluación de competencias podría considerar. Al igual que ocurre con cualquier sistema informático, el contar con un desarrollo propio tiene la ventaja de que se adecúa perfectamente al esquema organizacional de la institución. Sin embargo, el costo de desarrollo, mantenimiento y, sobre todo, de continua actualización, puede llegar a ser elevado, por lo que se necesita evaluar la conveniencia de moverse hacia alguna opción comercial siempre y cuando cubra de manera razonable las necesidades institucionales.

Es necesario contar con liderazgo, capacidad de gestión y de manejo del cambio

La implementación de un modelo como el descrito en esta ponencia requiere de un compromiso institucional al más alto nivel. De lo contrario, no es posible motivar a la acción ni acceder a los recursos financieros, tecnológicos y operativos requeridos, que no son pocos. El hecho de crear una capa adicional de evaluación implica destinar incentivos para los profesores, así como un amplio esfuerzo de entrenamiento. En el caso del Tecnológico de Monterrey, en la primera etapa se llevaron a cabo talleres de capacitación para los directores de programa y profesores involucrados en la evaluación, en los que se trabajaba en el desarrollo y aplicación de los instrumentos de evaluación de sus respectivos programas, con el objetivo doble de que el aprendizaje de los profesores fuera experiencial y que el tiempo dedicado a los mismos representase un avance en la tarea a realizar.

Aunque la mayor parte de las actividades necesarias para la evaluación de los programas son llevadas a cabo por los profesores, se requiere de una entidad que diseñe el proceso, organice

las acciones requeridas para su implementación, asegure su correcto funcionamiento y resuelva las situaciones que se presenten. En el caso del Tecnológico de Monterrey, esta entidad es la Dirección de Efectividad Institucional, perteneciente a la Vicerrectoría Académica, que tiene también la responsabilidad de mantenerse al día con las tendencias mundiales en la materia, con el fin de someter al modelo mismo a una mejora continua. Este equipo de gestión es también responsable del seguimiento y registro de las acciones de cada programa, incluyendo el cierre de ciclos de mejora, así como de elaborar los reportes correspondientes tanto internos como para las agencias acreditadoras que lo requieran.

La gestión del modelo implica proveer apoyo en cada etapa del mismo y asegurar el seguimiento del proceso como una actividad continua y no de una sola vez. Para ello, se requiere entrenar docentes y administradores, especialmente en el diseño de competencias de egreso, actividades de aprendizaje complejas y evaluación del desempeño. Se requiere también proveer los recursos financieros y tecnológicos para que el modelo opere sin obstáculos.

Sin embargo, quizá los requerimientos más grandes en materia de gestión son la capacidad de mover al cambio (*change management*) y la creación de una cultura donde se asimile el modelo como parte integral y permanente de la vida institucional. Para el éxito de un programa como el aquí descrito, es necesario contar con un liderazgo sólido que cuente con el compromiso y apoyo institucional al más alto nivel; pues el programa implica transformaciones sustantivas a lo largo de la institución, difíciles de realizar si se deja como responsabilidad exclusiva de algunos profesores y directivos, por más que estos sean entusiastas y tengan las mejores intenciones.

Es una oportunidad para alinear los esfuerzos institucionales con los requerimientos de acreditación

En materia de acreditación, el reto principal es mantener el programa como una actividad continua, viva y no como un mero requisito que hay que retomar cada vez que se aproxima la fecha de renovar la acreditación. Esto implica que el programa debe estar visible en la planeación anual de la institución y que debe dársele seguimiento continuo. Asimismo, es de suma importancia mantener al día el registro de los avances de la evaluación de cada uno de los programas académicos y, si se cuenta con varias acreditaciones, asegurarse de que el modelo institucional permita, con relativamente pocas adecuaciones, cumplir con los criterios de acreditación de cada agencia que, aunque parecidos, no son idénticos.

A MANERA DE CONCLUSIÓN: RETOS Y APRENDIZAJES

Si hubiese que enumerar los principales retos que enfrenta una institución educativa para echar a andar un programa de evaluación de competencias de egreso, tomando como base la experiencia del Tecnológico de Monterrey, habría que destacar los siguientes:

- Diseñar un modelo completo en su alcance, pero sencillo en su aplicación y fácil de comunicar. Este no es un reto menor, pues se requiere desarrollar un modelo conceptual con un sustento teórico-metodológico sólido, pero que pueda ser presentado y ejecutado de la manera más sencilla posible.
- Contar con una estrategia de movilización al cambio. Dado que un modelo de evaluación de competencias tiene implicaciones profundas de diverso orden, es necesario diseñar una estrategia de convencimiento acerca de la necesidad de ejecutar un cambio de tal magnitud. Los integrantes de la institución, particularmente los docentes, deben entender que el cambio no es una ocurrencia del equipo directivo, sino una respuesta a una necesidad impostergable de transformar la educación superior para atender mejor las necesidades formativas de los estudiantes y para responder con evidencias sólidas a las demandas de una sociedad que cada vez más exige a las universidades transparencia y rendición de cuentas. Asimismo, identificar a los docentes dispuestos a innovar y acercarles los incentivos apropiados forma parte de la estrategia de administración del cambio.
- Asegurar los recursos y el apoyo institucional al más alto nivel. Implementar un modelo de evaluación de competencias requiere de una cantidad muy importante de recursos, así como de tiempo y energía de muchos de los integrantes de la institución. Si no se cuenta con el apoyo decidido del equipo directivo de primer nivel, las posibilidades de éxito del proyecto se reducen considerablemente.
- Asegurar el cierre de ciclos de mejora. Uno de los retos más complejos consiste en lograr que el modelo se aplique en su totalidad; es decir, que la evaluación de las competencias no se convierta en un fin en sí misma, sino que sea un componente de un programa permanente orientado hacia la mejora continua. Por un lado, el cierre de ciclos como etapa final del proceso de evaluar trae aparejados cambios encaminados a mejorar los programas, tanto en aspectos muy puntuales como en otros más inclusivos. Por otro lado, el hecho de cerrar un ciclo implica abrir uno nuevo, lo que da continuidad y permanencia al programa.

- Mostrar los resultados y los beneficios de añadir complejidad al modelo educativo tradicional. Relacionado con los puntos anteriores, es importante contar con una estrategia permanente de comunicación que difunda los beneficios obtenidos como resultado de implementar un programa de evaluación de competencias de egreso que, como se ha dicho reiteradamente, implica cambios muy importantes en diversos ámbitos de la institución. ¿Cómo ha mejorado el aprendizaje de los estudiantes? ¿Con qué evidencias se cuenta sobre las capacidades adquiridas y demostradas por los estudiantes en contextos complejos de aplicación del conocimiento? ¿Cuál es el impacto del programa en la empleabilidad de los graduados y en la opinión de los empleadores? Este tipo de beneficios retroalimentan positivamente a la institución y fortalecen el programa de evaluación, asegurando su permanencia.

Asimismo, algunas recomendaciones surgen de la experiencia institucional del Tecnológico de Monterrey, que pudieran ser de utilidad para instituciones con intención de iniciar un programa de evaluación de competencias de egreso:

- Debido a la complejidad teórico-metodológica de la evaluación de competencias de egreso, existe el riesgo de extender por un largo período la etapa de análisis y diseño del programa de evaluación. Es sabido que en un ámbito académico lograr acuerdos sobre aspectos teórico-metodológicos puede tomar un largo tiempo. Por ello, es recomendable asumir una postura pragmática que, basada en una conceptualización apropiada (que pudiera no resolver del todo los debates teóricos sobre el tema, pero que sea consistente con lo que se quiere lograr), permita empezar en un plazo relativamente corto. Es importante considerar que, para que esto tenga éxito, la institución debe estar dispuesta a aprender de la experiencia y a ser muy ágil en los ajustes que pudiera requerir el modelo una vez que esté en funcionamiento. En la experiencia del Tecnológico de Monterrey, los primeros ejercicios de evaluación tenían áreas de oportunidad de diversa índole y, a lo largo de los ciclos subsecuentes, se ha conseguido madurar el proceso y mejorar sustancialmente la calidad del mismo.
- Es recomendable incorporar la definición de las competencias de egreso dentro del proceso regular de revisión curricular. Aunque es posible definir las competencias de egreso en cualquier momento, lo ideal es incluir esta actividad como parte del proceso de revisión de planes de estudio. El contar con un acuerdo de los docentes sobre las competencias de egreso de un determinado programa representa una base muy sólida para desarrollar la malla curricular, lo que permite trazar líneas transversales a lo largo

del currículo, así como fortalecer el sentido de propiedad, por parte de los docentes, del currículo completo y no solamente de las asignaturas que imparten.

- Como en cualquier proyecto innovador, es altamente recomendable conformar un equipo integrado por personas capaces de tomar riesgos y enfrentar un entorno de incertidumbre. Este equipo debe sentir que cuenta con un respaldo institucional sólido, incluyendo la posibilidad de cometer errores siempre y cuando se aprenda de ellos. Por ello, es importante identificar a docentes y directivos innovadores, invitarlos a ser parte integral del proyecto y proveerles de los incentivos apropiados para contar con ellos tanto en un inicio como en las etapas subsecuentes.
- Con el propósito de convencer a la comunidad sobre la necesidad de adoptar un programa de evaluación de competencias, se requiere una estrategia de comunicación que explique adecuadamente las razones para abrazar el cambio. Los comunicados desde lo más alto de la jerarquía organizacional fortalecen la percepción de importancia del proyecto. Para tener un mayor impacto, los mensajes deben, en primera instancia, explicar por qué la situación actual no es sostenible si la institución desea ser parte de las tendencias mundiales orientadas hacia la transparencia y la rendición de cuentas. Asimismo, se recomienda encontrar una mezcla adecuada de mensajes que explican los factores externos a la institución que demandan el cambio, con mensajes que enfatizan los beneficios internos de participar en una iniciativa como esta. Comunicar antes y durante cada etapa del proceso es un elemento clave para el éxito del mismo.

La evaluación de competencias de egreso ayuda a las instituciones de educación superior a medir el éxito de sus programas formativos en el desarrollo del perfil establecido para los graduados de los programas académicos. Permite también identificar áreas de oportunidad para subsanarlas mediante acciones de mejora apropiadas. En ese sentido, contar con un programa de evaluación de esta naturaleza permite tener confianza en que la institución cumple con el compromiso expresado en sus perfiles de egreso. Sin embargo, es importante estar seguros de que las competencias establecidas en los perfiles de egreso son lo suficientemente amplias en cuanto a su alcance que preparen al graduado lo mismo para su primer empleo que para el último; es decir, que el graduado cuente con competencias concretas para una adecuada inserción al mercado laboral y con competencias más amplias que le permitan crecer y alcanzar la cúspide de su carrera profesional después de treinta, cuarenta o más años. A fin de cuentas, la universidad trabaja con el estudiante a lo largo de solamente cuatro o cinco años, pero el impacto de su labor formativa trasciende el primer empleo y dura a lo largo de toda una vida.

Bibliografía

- Argudín, Y. (2006). Educación basada en competencias: Nociones y antecedentes. México: Trillas.
- Blanco Fernández, A. (2009). Desarrollo y evaluación de competencias en educación superior. Madrid: Narcea.
- Díaz-Barriga, F. (2006). La evaluación auténtica centrada en el desempeño: Una alternativa para evaluar el aprendizaje y la enseñanza. En Enseñanza situada: Vínculo entre la escuela y la vida (pp. 125-163). México: McGraw-Hill Interamericana.
- EDUCAUSE (2014). EDUCAUSE Learning Initiative (ELI) - 7 Things You Should Know About Competency-Based Education. Disponible en: <<https://library.educase.edu/~media/files/library/2014/2/eli7105-pdf.pdf>> (recuperado el 7 agosto de 2016).
- Fulton, K. P. (2014). Time for Learning. Top 10 Reasons Why Flipping the Classroom Can Change Education. California: Corwin a Sage Company.
- Goodwin, B. & Miller, K. (2013). Evidence on Flipped Classrooms Is Still Coming In. *Educational Leadership*, 70(6), 78-80.
- Gosselin, D., Cooper, S. & Bonnstetter, R. (2013). Exploring the assessment of twenty-first century professional competencies of undergraduate students in environmental studies through a business-academic partnership. *Journal of Environmental Studies Sciences*, 3, 359-368.
- Hartel, R. W. & Foegeding, E. A. (2004). Learning: Objectives, competencies, or outcomes? *Journal of Food Science Education*, 23, 69-70.
- Johnstone, S. M. & Soares, L. (2014). Principles for developing competency-based education programs. *Change: The Magazine of Higher Learning*, 46(2), 12-19.
- Lumina Foundation (2014). Today's business leaders say, "It's what you know, not where you go" when making hiring decisions, new study shows. Disponible en: <<https://www.luminafoundation.org/news-and-views/>>

- today-s-business-leaders-say-it-s-what-you-know-not-where-you-go-when-making-hiring-decisions-new-study-shows> (recuperado el 5 agosto de 2016).
- ManpowerGroup (2015). 2015 Talent Shortage Survey. Disponible en: <<http://www.manpowergroup.com/talent-shortage-2015/talent+shortage+results>> (recuperado el 5 agosto de 2016).
- Meyer, C. A. (1992). What's the difference between authentic and performance assessment? *Educational Leadership*, 49, 39-40.
- Partnership for 21st Century Skills (2011). Framework for 21st century learning. Disponible en: <http://www.p21.org/storage/documents/1.__p21_framework_2-pager.pdf> (recuperado el 10 agosto de 2016).
- Wiggins, G. P. (1993). *Assessing student performance*. San Francisco: Jossey-Bass Publishers.
- Viau, R. (1994). *La motivation en contexte scolaire. Saint-Laurent*. Québec: Editions du Renouveau pédagogique.
- Villa, A. & Poblete, M. (2007). *Aprendizaje basado en competencias: Una propuesta para la evaluación de las competencias genéricas*. Bilbao: Mensajero ICE / Universidad de Deusto.

Seguimiento de egresados y evaluación de la formación por competencias¹

Juan Fernando Vega Ganoza

RESEÑA

Uno de los rasgos propios del modelo de formación por competencias (MFC) es que propone hacer explícita la conexión entre los programas formativos y las demandas del mundo laboral. Por lo tanto, el ajuste de los egresados al mercado laboral puede ser visto como el criterio final de evaluación por competencias. La ponencia presenta una reflexión sobre esta expectativa a partir de la experiencia y los resultados del Sistema de Seguimiento de Egresados PUCP (SSE PUCP), los servicios que puede brindar al MFC y sus usos a la fecha. Se trata de responder a la pregunta: ¿qué pueden aportar los resultados del SSE a la evaluación de la formación por competencias?

PRESENTACIÓN

La pregunta tiene un responsable. La formuló Dieudonné Leclercq explorando la conexión entre dos iniciativas de trabajo en la Dirección de Asuntos Académicos de la PUCP sobre las que estábamos conversando: la formación por competencias y el seguimiento de egresados.

¹ Esta presentación resume las ideas de un texto mayor ("*Know how* y empleabilidad: evidencias del Sistema de Seguimiento de Egresados PUCP") que capitaliza esfuerzos de mucha gente que participó en el desarrollo del SSE y, en particular, de Lucía Wiener, quien lo supervisa y produce análisis desde 2014. Por otro lado, Lucía Wiener, Flavio Figallo, Leandro García Calderón, Oscar Paín y César Guadalupe me han ayudado con comentarios útiles y no pocas correcciones. A todos ellos, mi agradecimiento.

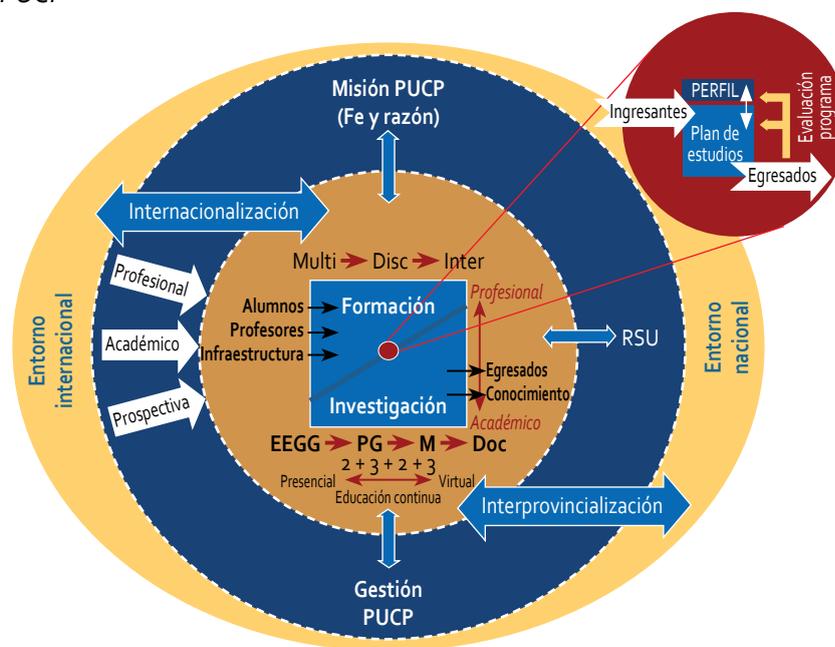
Como se verá, la presentación sigue más bien la lógica de un viaje exploratorio que la de un informe de investigación. La crónica irá presentando hitos en el recorrido, cuáles fueron los hallazgos por etapa y qué reflexiones generan a la luz de la literatura y las evidencias.

LA FORMACIÓN POR COMPETENCIAS EN EL MODELO EDUCATIVO PUCP

La formación por competencias se comenzó a incorporar formalmente en el modelo educativo PUCP alrededor del año 2010. La figura 1 está tomado de una presentación de esa época.

La figura representa, desde afuera hacia adentro, las varias influencias del entorno, el marco de nuestros instrumentos programáticos y de gestión, nuestras actividades y el núcleo del conjunto es representado por la formación y la investigación. Conviene registrar, para nuestra

FIGURA 1
Modelo PUCP



Fuente: gráfico de una presentación de 2010 del vicerrector académico (el VRAC) a las autoridades, tomado de una presentación VRAC de 2013 (Gonzales de Olarte, 2013).

ponencia, que el modelo de formación se refiere a un perfil de competencias y un plan de estudios alineados al perfil, capaz de ser reatualizado a partir de información de los egresados. Esta presentación en 2010 es una de las primeras apariciones de la formación por competencias y el seguimiento de egresados como parte del modelo educativo PUCP.

Por supuesto, este modelo formaba parte de un plan más ambicioso aún, resumido así en el discurso de orden de la apertura del año académico 2011 que versó sobre nuestro modelo académico: "Su plan de gobierno (del rectorado vigente) plantea cinco metas: consolidar la excelencia académica, impulsar la interdisciplinariedad, convertir a nuestra Casa de Estudios en una universidad de investigación y posgrado, internacionalizar nuestra universidad y modernizar nuestro modelo de gestión académica y administrativa" (Gonzales de Olarte, 2011, p. 5).

Estos temas se presentan en el círculo interno en el que se pone énfasis en la doble tarea de formación e investigación y en la articulación entre las dimensiones multidisciplinarias, disciplinarias e interdisciplinarias, asociadas a un programa de pregrado 2+3 (dos años de estudios generales y tres de estudios específicos) y de posgrado (dos años de maestría y tres de doctorado).

De esta manera, la formación por competencias en la PUCP se propone como una formación integral:

[...] la educación universitaria no es completa si únicamente se limita al desarrollo de las capacidades profesionales. Resulta imprescindible sumar a este objetivo igualmente importante: el de formar a la persona. Los egresados de la Universidad Católica son reconocidos como los mejores en sus especialidades, pero deben serlo también como buenas personas y buenos ciudadanos. [...] Optamos, en suma, por una formación integral y humanista, y la excelencia académica (Gonzales de Olarte, 2011, p. 9).

En el extremo *output* de nuestro sistema, se registra egresados de la formación y conocimientos de la investigación. En el *output* formativo, aparece una distinción que les invito a retener: tenemos egresados profesionales y egresados académicos. Es una distinción que solemos descuidar y, como veremos, tiene consecuencias para la reflexión sobre los datos de empleabilidad y niveles de satisfacción que ofrece el SSE PUCP.

UNIVERSIDAD Y DISCIPLINAS, ACADÉMICOS Y PROFESIONALES: DISTINCIONES

Para construir la distinción entre académicos y profesionales, conviene ubicar los términos dentro de un horizonte conceptual. Para ello, usaremos varias categorías de uso común en la sociología de la educación superior.

La universidad

Comenzaremos con la distinción entre la universidad y las disciplinas que introdujo Burton Clark, sociólogo que fundó el análisis de la universidad como sistema (Gumport, 2007; Marginson, 2014). Para él existen dos dimensiones principales en la institución universitaria: la dimensión empresa o institucional y la dimensión academia o disciplinar. La empresa es la organización, la unidad institucional; pero es un tipo curioso de organización: es comprehensiva, reúne varias disciplinas. La academia en cambio es especializada y se organiza según sus propios criterios. Las disciplinas dominan la vida laboral, la división del trabajo; son la unidad vital, dinámica (Clark, 1986, pp. 28-34). Este es un tópico de los estudios de administración universitaria que conciben la tarea de administrar la universidad como análoga a la de arrear gatos (Mindreau, 2014).

La idea converge con la figura de la universidad como multiversidad de Clark Kerr: la especialización disciplinar, pero también la diversidad de fines, han hecho de la vieja comunidad universitaria homogénea una cosmopolita ciudad del intelecto, una ciudad habitada por distintas comunidades disciplinares a las que la administración brinda servicios públicos, como la calefacción o los estacionamientos (Kerr, 1963). No conviene, sin embargo, subestimar la importancia de esta dimensión, “muchos de los factores que inciden en la calidad de una carrera o programa son, en verdad, de carácter institucional: las políticas de gestión de los recursos humanos, la organización presupuestaria, las políticas de selección y admisión de alumnos, los criterios de inversión en infraestructura y equipamiento entre ellas” (Lemaitre & Zenteno, 2012, p. 49). Martín Trow, por otro lado, pone énfasis en que el creciente cambio de escala institucional que acompaña el paso del sistema universitario de élite al sistema universitario de acceso universal genera una creciente importancia de la administración profesional institucional (Trow, 2006).

Y por otro lado, la multiversidad no solo implica una colección de comunidades, sino que la diversidad de colecciones viene acompañada de diversidad de misiones. La misión de formación, la misión de investigación, la misión de extensión, decíamos antes, y la novísima misión de innovación o de desarrollo de *start ups*, tienen diversos énfasis en diversas identidades institucionales.

Las disciplinas

La reflexión sobre las disciplinas constituye otro piso analítico. El sociólogo Tony Becher postula que las disciplinas son el constructo resultante de la interacción de un colectivo (una tribu) con un campo de conocimiento (su territorio) que en el trabajo genera una cultura propia (Becher, 2002).

La tabla siguiente combina el modelo sobre los campos de conocimiento que construye Becher en su capítulo introductorio, sobre la base de Kolb, Biglan y Pantin, con ideas de Snow y Kuhn.

TABLA 1

Modelo de campos del conocimiento

Snow Kolb/Biglan (Pantin)	Ciencias Abstracto/Científico	Humanidades (y CCSS) Concreto/Artístico
Reflexivo/Básico	Duro puro (*) Ciencias naturales y matemática (restringido-paradigmático) Teoría predictiva	Blando puro Humanidades y CCSS (no restringido-preparadigmático) Interpretaciones
Activo/Aplicado	Duro aplicado Ingeniería, biotecnología, medicina (método heurístico) Productos - Técnicas	Blando aplicado Profesiones sociales (juicio cualitativo) Protocolos

(*) Campo de aplicación de las reflexiones de Kuhn.

Fuente: elaboración propia sobre la base de Becher, 2002 y Snow & Leavis, 2006.

La idea principal es que existe una relación fuerte entre comunidad y territorio: el modelo adjunto describe cuatro campos (territorios tipo) en los que la comunidad universitaria puede reconocerse. La tabla indica que cada campo tiene no solo métodos de trabajo y forma de producto, propios; sino estilos cognitivos dominantes (sobre los que apoyan los test vocacionales). El esquema ayuda a dar contenido a las diferencias entre ciencias y humanidades (y ciencias sociales), así como —dentro de cada polo— a la diferencia entre el interés por conocimientos básicos respecto del interés por conocimientos aplicados.

Para Becher, los rasgos comunitarios de una disciplina son una construcción social histórica y, para representarla de modo amplio, apela a King y Brownell: “Su descripción contempla varios aspectos diferentes: una comunidad, una red de comunicaciones, una tradición, un conjunto particular de valores y creencias, un dominio, una modalidad de investigación y una estructura conceptual” (Becher, 2002, p. 38; King & Brownell, 1966).

No se requiere mucha imaginación para vincular estas ideas con el texto de Kuhn sobre la estructura de las revoluciones científicas (Kuhn, 1990). Conviene mencionar, sin embargo, que

en sentido propio las batallas que Kuhn describe entre viejo y nuevo paradigma solo caben en el cuadrante ciencia/dura. A pesar de que humanistas y científicos sociales puedan reconocerse en las prácticas de control de la disciplina que ejercen “las autoridades”, los mandarines o la *nomenklatura*, respecto del acceso a financiamiento de investigación y a canales prestigiosos de publicación, e incluso del acceso y progreso en la carrera docente, la ausencia de paradigmas imprime a las discrepancias en humanidades y ciencias sociales formas que no requieren de revoluciones y que están menos estudiadas.

Para decirlo en pocas palabras, esta reseña nos permite establecer que la dinámica sustantiva de las disciplinas es la búsqueda de conocimientos y que el tipo de conocimiento buscado genera la diversidad entre las disciplinas debido a la autonomía con que establecen sus finalidades, objetivos, prioridades, estilos y organización.

Frecuentemente, se usan de manera indistinta los términos “disciplina” y “profesión”. Hablamos de la profesión académica, las profesiones liberales, las profesiones dependientes y las profesiones de los emprendedores. En el viejo sentido, solamente las profesiones liberales eran profesiones. Veremos a continuación que todas las profesiones son disciplinas, pero no todas las disciplinas son profesiones.

Las profesiones

Las profesiones son el resultado de la división del trabajo, que implica un servicio. No nos detendremos en el tema, pero no podemos dejar de mencionar la raíz religiosa y moral del término “profesión” en la construcción social del servicio profesional (según Weber, 2012). Solo indicaremos que la idea de servicio tradicional (vocación) de las profesiones liberales estaba normalmente separada de cualquier consideración de lucro no solo por razones religiosas, sino por la antigua aversión grecoromana al trabajo y el negocio.

Son también —como las disciplinas— una institución social. Como tales, comparten las dimensiones consideradas en la descripción ya mencionada de King y Brownell. Pero tienen un rasgo propio. Un resumen muy reciente propone cuatro “rasgos de familia” de la profesión: i) conocimientos especializados, ii) acceso filtrado mediante credenciales, iii) actividades reguladas por un “*grand bargain*” y iv) compromiso con un conjunto de valores (Susskind & Susskind, 2016).

De los cuatro rasgos familiares mencionados por los Susskind, el que distingue las profesiones de las disciplinas es el tercero: “*the grand bargain*”. *Bargain* es un término interesante: puede significar convenio, regateo o una ganga. En este caso —es decir, tomado en serio—, se

trata de un convenio que permite que el ejercicio profesional sea un monopolio (Susskind & Susskind, 2016, pp. 9-45). El convenio permite un triple monopolio: de acceso, de formación y de prestación de servicios.

El juramento hipocrático es probablemente la más antigua expresión de monopolio de acceso y formación: reservar el saber experto a los miembros del colectivo y mantenerlo oculto a las personas no dignas². Las otras condiciones del juramento hacen pensar en criterios de responsabilidad, bondad y religiosidad. Pero esta es una restricción de grupo autoimpuesta; no depende de actor externo. Además, no incluye monopolio del ejercicio.

La antigua restricción adquiere nuevas formas con la mediación de la autoridad. La de los gremios medievales, por ejemplo: las diversas “*universitas*”, comunidades de heterogénea naturaleza (comerciales, artesanales, académicas) eran titulares de privilegios monopólicos públicamente otorgados por las autoridades capaces de imponerlos; eclesiales o no (desde el papa o el emperador al obispo de la localidad o la comuna municipal). El principal cambio es que la restricción autoimpuesta por la comunidad del saber experto obtiene una sanción de la autoridad que le otorga un monopolio reconocido y exigible. Es el caso, por ejemplo, de la *licentia ubique docendi* del *Studium Generale* de París: solo París podía, mientras el papa no incluyó a otras universidades, otorgar licencia para enseñar en cualquier parte de la cristiandad. La modernidad, sin embargo, ofrece dos soluciones:

- En la solución francesa, el diseño que otorga a los colegios profesionales autonomía y responsabilidad para garantizar la idoneidad de sus miembros es trastocada, ya que dejan de controlar el acceso y la formación: las universidades no solo otorgan grados académicos y licencia para enseñar, sino licencias (títulos) profesionales “a nombre de la nación”. El Estado habilita el ejercicio profesional.
- La solución anglosajona, en cambio, puede ser resumida así:

En reconocimiento de su conocimiento experto, destreza y juicio; en tanto se espera que lo apliquen a la realización de servicios accesibles, actualizados, confiables y a precio razonable; en el entendido de que mantendrán actualizados sus conocimientos y métodos,

² “Considerar a mi maestro en medicina como si fuera mi padre; compartir con él mis bienes y, si llega el caso, ayudarle en sus necesidades; tener a sus hijos por hermanos míos y enseñarles este Arte, si quieren aprenderlo, sin gratificación ni compromiso; hacer a mis hijos partícipes de los preceptos, enseñanzas y demás doctrinas, así como a los de mi maestro, y a los discípulos comprometidos y que han prestado juramento según la ley médica, pero a nadie más” (Lemarchand, 2010, p. 86).

entrenarán a sus miembros, establecerán y harán cumplir estándares de calidad para su trabajo, y que sólo admitirán individuos adecuadamente calificados en sus filas; y que siempre actuarán con honestidad y buena fe, poniendo el interés de los clientes por delante de los propios; nosotros (la sociedad) depositamos nuestra confianza en las profesiones otorgándoles monopolio exclusivo sobre una amplia gama de servicios y actividades socialmente significativas, pagándoles un salario justo, concediéndoles independencia, autonomía, derechos de autodeterminación, y reconociéndoles respeto y estatus (Susskind & Susskind, 2016, p. 21. Traducción propia del inglés)³.

En este caso la licencia para ejercer la profesión, la otorgan los colegios profesionales.

La secuencia formación-acceso cambia en la versión contemporánea con la formación otorgada por académicos universitarios que no necesariamente son miembros colegiados del cuerpo profesional. Algunos de los formadores son académicos expertos en la disciplina, dedicados a la investigación que no ejercen el servicio profesional para el que es requisito la colegiación. Esta última distinción requiere de un desarrollo que realizaremos a continuación.

La academia

Decíamos que no todas las disciplinas son profesiones. Y es que existen algunas disciplinas cuyo horizonte de ejercicio es interno a la institución universitaria. Se trata de las disciplinas académicas. El modelo de la licencia profesional adoptó el nombre del antiguo privilegio papal universitario: la licencia para enseñar concedida por la Facultad de Artes⁴. En ese sentido, hay algunas disciplinas cuyo ámbito profesional son la enseñanza y la investigación, y su entorno es el universitario, la carrera académica, el avance del conocimiento. El monopolio legal que las protege es el que restringe el acceso a la carrera docente universitaria; aunque muchas comunidades académicas hubieran preferido el respeto de las viejas autonomías.

3 *"In acknowledgement of and in return for their expertise, experience and judgement, which they are expected to apply in delivering affordable, accessible, up-to-date, reassuring, and reliable services, and on the understanding that they will curate and update their knowledge and methods, train their members, set and enforce standards for the quality of their work, and that they only admit appropriately qualified individuals into their ranks, and they will always act honestly, in good faith, putting the interest of clients ahead of their own, we (society) place our trust in the professions in granting them exclusivity over a wide range of socially significant services and activities, by paying them a fair wage, by conferring upon them independence, autonomy, rights of self-determination, and by according them respect and status".*

4 Que era la culminación del pregrado y no habilitaba para el ejercicio de la Medicina, el Derecho o Teología, que requerían maestrías o doctorados en las facultades respectivas.

En adelante, hablaremos de disciplinas académicas para referirnos a aquellas que no han considerado necesario construir un monopolio legal para su ejercicio, como Filosofía o Física. O también, son disciplinas académicas aquellas en que la licencia se corresponde con la fórmula tradicional medieval: habilita para enseñar, aunque para hacerlo siga siendo necesario ser admitido en alguno de los claustros académicos⁵.

La confusión producto de que en la tradición francesa todo pregrado conduce no solo a un grado (bachiller), sino a un título profesional (licencia), hace que se nos escapen diferencias útiles tanto para construir las trayectorias formativas y desarrollar capacidades específicas⁶ como para mirar los perfiles de egreso.

En la reflexión previa con las ideas de Becher, las disciplinas académicas se corresponden con los dos cuadrantes de investigación básica tanto para las ciencias como para las humanidades, mientras las profesiones se corresponden más bien con los cuadrantes de conocimiento aplicado.

Esto tiene consecuencias, como veremos, sobre el empleo; pero también sobre dónde se realizan las innovaciones. Mientras que en las disciplinas académicas las principales innovaciones son realizadas en la universidad, muchas de las innovaciones profesionales son resultado de acciones fuera de ella (Herrera & otros, 2012, p. 52).

SISTEMAS DE SEGUIMIENTO DE EGRESADOS Y EVALUACIÓN DE PROGRAMAS POR COMPETENCIAS (MODELOS)

Con las distinciones hechas en la sección anterior, podemos avanzar sobre el corazón de la ponencia: ¿qué pueden aportar los resultados del SSE a la evaluación de la formación por competencias?

El modelo ATOMES Leclercq-Cabrera

La pregunta la formuló Dieudonné Leclercq en su primera visita para asesorar al equipo de "Gestión curricular y apoyo académico" sobre la implementación del modelo de formación por competencias (MFC), en junio de 2015. Después de conversar sobre las actividades y productos

5 Hoy en día, sin embargo, incluso en el Perú, la licencia es insuficiente y se necesita un grado adicional: una maestría.

6 Una aplicación útil de la distinción, por ejemplo, se refiere a la distinción entre los géneros literarios académicos y profesionales (Natale, 2013; Navarro, 2012; Parodi, 2009).

del SSE PUCP, preguntó cómo los estábamos usando y sugirió que esta información podría —y debería— ser útil para la evaluación de la formación por competencias sobre la que veníamos trabajando. La sugerencia sonaba sensata. Si el MFC propone hacer explícita la conexión entre los programas formativos y las demandas del mundo laboral, el ajuste de los egresados al mercado laboral, que el SSE explora, podría ser visto como el criterio final de evaluación. El detalle es cómo lograrlo.

Al respecto, Leclercq y Cabrera tienen un texto metodológico sobre dispositivos de evaluación de aprendizajes para modelos de formación por competencias. Entonces, la manera más expeditiva para obtener pistas y responder a la pregunta fue hacerlo a partir de las propuestas de su libro (Leclercq & Cabrera, 2014, p. 26).

Los autores denominan a su modelo “ATOMES”, palabra que resulta de las siglas en francés de *Alignement à un Tableau des Objectifs, Méthodes et Évaluations dans Strata* (alineamiento en un tablero de objetivos, métodos y evaluaciones en estratos).

Los objetivos, métodos y evaluaciones constituyen los tres pilares del “programa curricular-plan de estudios” que es necesario alinear a un perfil de egreso que se organiza por competencias, según una secuencia (que los autores identifican con el acrónimo “DIME”) de cuatro niveles: las competencias dinámicas (o estratégicas), genéricas (“instrumentales” en el acrónimo), metacognitivas (o autocognitivas) y específicas. La evaluación debe atender cuatro niveles de profundidad de las evaluaciones siguiendo la pauta de Kikpatrick⁷: satisfacción del egresado, el dominio o maestría en las competencias DIME, el desempeño profesional o aplicación al terreno y los beneficios para los destinatarios del servicio (usuarios, empresa, sociedad).

De estos cuatro niveles, el SSE podría ocuparse de tres:

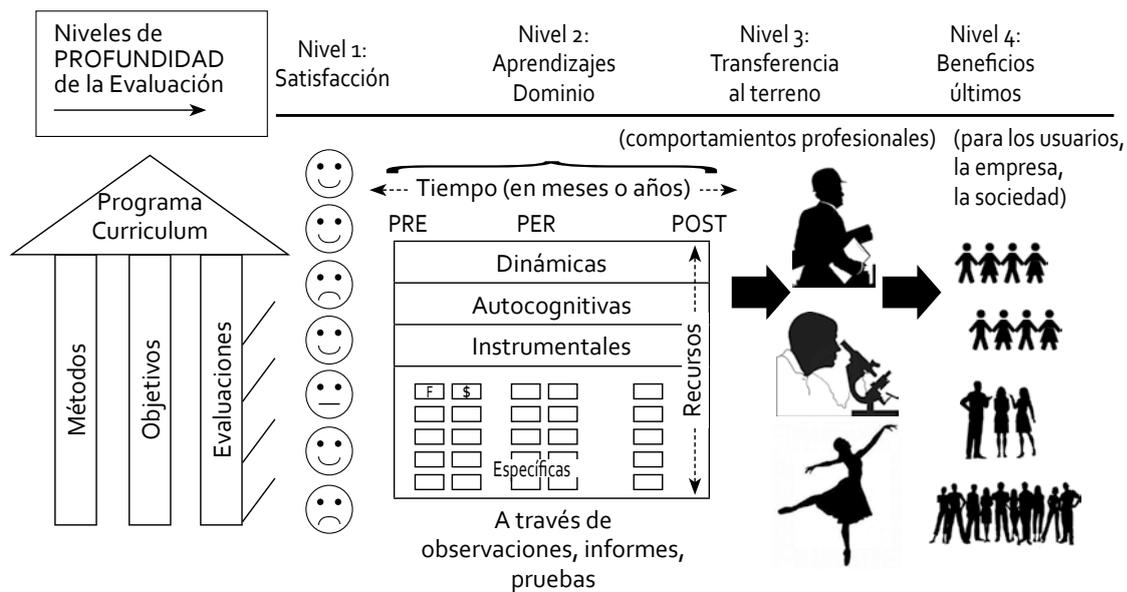
- En el nivel 1, podría explorar el nivel de satisfacción de los egresados sobre: logro de las competencias, formación que le ofreció la institución y adecuación de las competencias a las demandas de su vida luego del egreso.
- El nivel 2, referido a la maestría en el dominio de las competencias, requiere de un tipo de evaluación fuera del alcance del SSE; resulta difícil imaginar cómo aplicar las pruebas de desempeño que este nivel requiere⁸.

7 Referencia tomada de la presentación de Leclercq que figura en esta misma publicación.

8 Esto obviamente no implica que no se realice esta evaluación. La evaluación de la destreza en el dominio de las competencias, cuando se lleva a cabo, suele hacerse con los estudiantes, antes de que egresen.

FIGURA 2

Niveles y profundidad de la evaluación



Fuente: Leclercq & Cabrera, 2014, p. 26.

- El nivel 3 de aplicación en el terreno se podría aproximar más bien de manera agregada que por nivel de competencias, con indicadores de empleabilidad. La idea es que la empleabilidad debería reflejar la transformación de las competencias en comportamientos profesionales aplicados en la práctica (más bien luego de ganar experiencia que recién egresados).
- El nivel cuatro (beneficiarios últimos) podría incorporarse a los SSE a través de encuestas a los usuarios finales. El problema será identificar o aproximar a los beneficiarios finales; en ciertos casos pueden ser individuos (para profesiones liberales) y, en otros, actores institucionales (empleadores, *head hunters*, etc.).

Sin embargo, resulta interesante observar que, comenzando desde el nivel cuatro de beneficiarios finales, la reflexión puede incorporar las distinciones entre universidad y disciplina, así como entre disciplinas académicas y disciplinas profesionales.

Los profesionales liberales autoempleados tienen beneficiarios directos a los que encuestar; pero la mayoría de profesionales ejercen en condición de trabajadores dependientes y, en este caso, habría que encuestar a los empleadores o *head hunters*. El problema se complicaría de acuerdo con la escala empresarial: en una empresa de gran tamaño, ni gerentes ni responsables de personal están suficientemente al día de las competencias individuales más allá de cierto estereotipo asociado a un efecto “empresa” (prestigio institucional de la universidad). En el caso de las disciplinas académicas, los desempeños asociados a la enseñanza podrían ser verificados con entrevistas a los alumnos o a responsables de unidades académicas (lo que sería análogo a la entrevista a empleadores); los desempeños asociados a la investigación, en cambio, dependerían mucho del tipo de investigación y del tipo de organización. En cualquier caso, la dificultad mayor la representan los *free lance*.

Complejidad adicional se puede introducir especialmente en el nivel 1. Los modelos por competencias son explicados sobre todo con ejemplos de las especialidades profesionales con énfasis en la producción de productos o técnicas en ciencia aplicada y de protocolos en las humanidades y ciencias sociales aplicadas. La pertinencia del modelo DIME a las disciplinas académicas “puras” productoras de teorías predictivas o de interpretaciones convendría revisarla con mayor detalle.

Otras experiencias del SSE

No hay demasiada experiencia internacional disponible sobre el uso del SSE como herramienta para la evaluación de modelos de formación por competencias; sin embargo, hemos recogido algunas ideas.

El seguimiento de egresados es un invento de las universidades privadas norteamericanas de la *ivy league*. En la tradición anglosajona, los *Alumni*⁹ son un importante activo de la universidad. La educación de élite anglosajona se organiza con la lógica del club. Mantener viva la membresía es importante para la institución y sus miembros. Y no se debe perder de vista el no menos atípico fenómeno anglosajón de la donación al *endowment* del *Alma Mater*. Es probable que por eso sean pioneros en los estudios de egresados y resulta interesante constatar que los orígenes son fechados en los años setenta, que es la época de la masificación de la universidad en EEUU (Borden, 2003, p. 81; Cabrera, Weerts & Zulick, 2003, p. 55).

9 Término latino usado para referirse a los exalumnos.

Respecto a los estudios de egresados en América Latina, el ensayo recapitulatorio del estu-
pendo libro monográfico *Seguimiento de egresados e inserción laboral* (CINDA & GOUCh, 2012)
señala que el seguimiento de egresados forma parte actualmente del nuevo contrato social
entre el Estado y las universidades como forma de rendición de cuentas¹⁰. El mecanismo saltó
del club anglosajón a la política pública en general como política de aseguramiento de la calidad.
En este contexto:

El éxito de las universidades se mide principalmente en el resultado de sus estudiantes con res-
pecto a su situación laboral y su compromiso social. Los estudios de seguimiento de egresados,
constituyen una manera de realizar esta medición. Los estudios recopilan información sobre
el desarrollo profesional, personal y social de los egresados. La idea no es solo producir datos,
sino también analizarlos y facilitar la toma de decisiones para las autoridades universitarias
(Herrera & otros, 2012, p. 31).

Resumiendo, la empleabilidad es vista como un ajuste del plan de estudio a las demandas
laborales. No se registra la presencia de estudios asociados al cuarto nivel de profundidad en
el modelo ATOMES: retroalimentación de la formación profesional a partir de los beneficiarios
últimos; pero quizás convendría añadir que las reflexiones registradas en los textos de Leclercq
y CINDA parecieran indicar un estereotipo sesgado a las disciplinas profesionales.

Egresados y competencias en la PUCP: recorrido

Para entender la información que presentaremos conviene tener una idea de la vinculación
entre el SSE PUCP y el MFC en la PUCP. El SSE lleva ya más de una década y es anterior a la
incorporación de la formación por competencias al modelo educativo PUCP en 2010.

El proceso del SSE inició en el año 2003 con un primer estudio que daba información a nivel
institucional y que cubría un arco de doce años (tres cuatrienios). El pixelado era muy grueso.
La información institucional —como veremos— no sirve para informar cómo se comportan las
facultades y es fuertemente dependiente de las estructuras demográficas de la universidad.
Por eso, inmediatamente después (en 2004), se hizo un segundo estudio con una muestra por
facultades ponderada por especialidad. El pixelado fue un poco mejor, pero seguía siendo muy
grueso. Por ejemplo, la facultad de Humanidades alojaba Psicología, Historia, Lingüística, Geo-
grafía, etc. El promedio de la Facultad de Humanidades no interesa, sino muy eventualmente, en

10 El *New Public Management* implementado por Margaret Thatcher y continuado por Tony Blair en el Reino Unido.

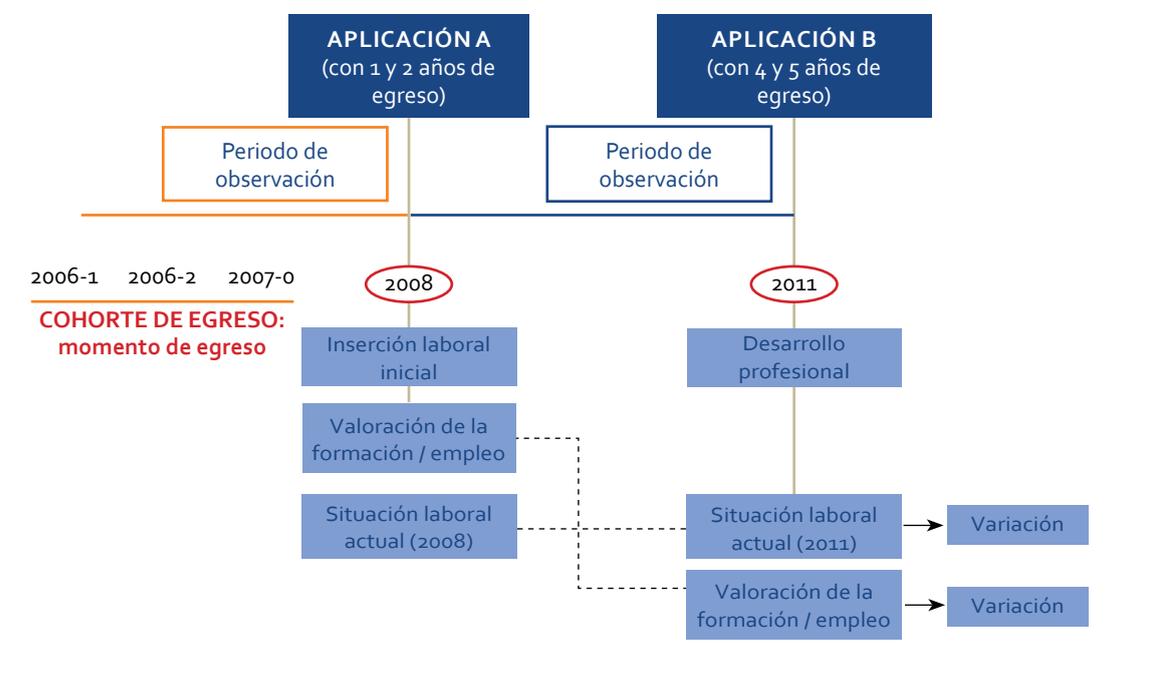
misma facultad. No hay manera de retroalimentar la actividad académica desde el promedio de la facultad, porque esta actividad es de las disciplinas, que es a quienes hay que retroalimentar como indica Clark.

Por eso, en 2007, se diseñó un nuevo modelo de sistema de seguimiento, censal y continuo, en dos etapas: una etapa luego de uno a dos años de egreso y la otra etapa a los cuatro a cinco años de egreso.

En 2008, se hizo el tercer estudio, que reportó a nivel de especialidades, descubriendo que la mayoría de estas son de N-pequeño; es decir, tienen una población (N) menor a treinta personas. Cuando egresan pocas personas, no es posible estimar el comportamiento grupal de la especialidad. Aun cuando encuestemos a los tres matemáticos egresados ese año y obtengamos el dato preciso de sus ingresos, no podemos juzgar la "representatividad" de los datos de estas tres personas respecto de la población de los matemáticos PUCP con tiempo de egreso

FIGURA 3

Esquema del sistema de seguimiento de egresados PUCP (SES PUCP): secuencia desde 2008



semejante. Un año se pueden graduar matemáticos extrovertidos, al siguiente introvertidos y luego los que hablan alemán. No hay manera de hablar de los matemáticos como grupo, ni decir si esos son los ingresos esperables para esta especialidad, hasta tener un número suficiente de matemáticos. El análisis estadístico requiere un número suficiente de casos. Por otro lado, este mismo mecanismo explica la alta variabilidad de los datos de las poblaciones que tienen en común pequeño tamaño.

Entonces, desde 2011 para la cohorte 2009, se comenzó a hacer el seguimiento anual, acumulando cinco años a la fecha. Acumular es una manera de ganar escala y, por lo tanto, representatividad. Veremos algunos datos de esta serie más adelante.

¿Cuáles son los temas de la encuesta? Sobre inserción laboral: trabajo antes de egresar, primer empleo y empleo actual. Este último es el empleo más fácil de comparar; es aquel para el que conocemos mejor el momento de formulación de la pregunta y es el que podemos indexar. Sobre satisfacción, indagamos respecto de competencias (brindadas y no brindadas, demandadas o no demandadas), el prestigio de la carrera y de la universidad y los servicios de esta. También preguntamos sobre estudios de posgrado, pero este es un tema que no se trata en esta ponencia.

EL SSE PUCP: EMPLEABILIDAD Y SATISFACCIÓN

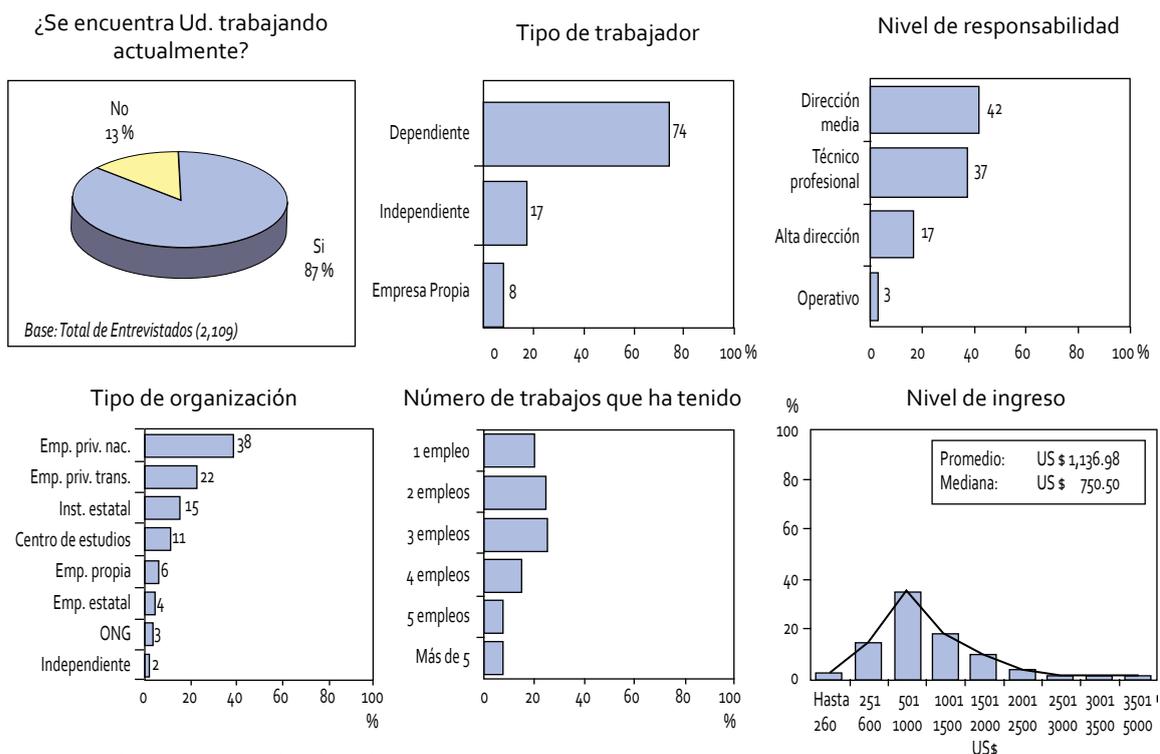
Siguiendo las pautas identificadas en el modelo ATOMES, en esta sección vamos a revisar qué nos dice el SSE PUCP sobre dos temas: la información sobre trabajo actual como proxy de nivel 3 de evaluación y algunos indicadores de autopercepción respecto al logro y utilidad de las capacidades como proxy del nivel 1 del mismo modelo. Además, introduciremos comentarios basados en las distinciones entre empresa y disciplina, así como disciplinas académicas y profesionales trabajadas en la tercera sección¹¹.

La secuencia de presentación sigue la evolución de los estudios que acabamos de desarrollar.

11 Usaremos como equivalentes disciplina, especialidad y carrera. Todas estas palabras designan al colectivo al que se incorporan los graduados luego de culminar el programa de estudios del que egresan y recibir la certificación (grado o título). Implican también el colectivo dentro del que se forman.

FIGURA 4

Vista resumen de la situación laboral actual (2003)



Base: Entrevistados que se encuentran trabajando actualmente (1,784)

Fuente: Apoyo, 2003, p. 11.

Primer estudio (2003)

Esta imagen era la reseña, en una sola lámina, de la condición laboral en el primer estudio que, conviene recordar, tenía un diseño muestral institucional para tres cuatrienios, cubriendo un total de doce años (1990-2001).

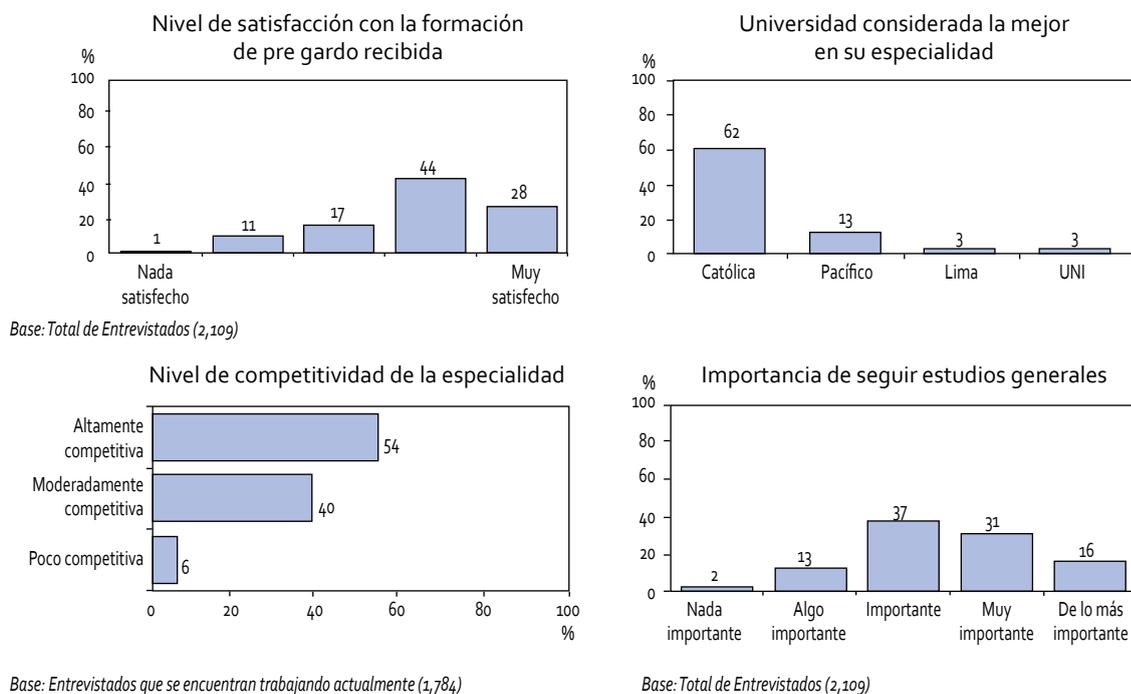
La mayoría trabaja (87%) en condición dependiente (74%), en el sector privado (60%), en nivel directivo (17% alta y 42% media), con un ingreso promedio de mil cien dólares (\$1100), una mediana de \$750 y baja rotación en el empleo.

Los resultados fueron considerados buenos, aunque quedó pendiente explorar el porcentaje de los que no trabajan, de los cuales el 58% correspondían a los del cuatrienio más reciente,

14% no declaraban el lapso de tiempo en que no trabajaron y 42% tenían un tiempo buscando trabajo menor a seis meses; sin embargo, no se obtuvo información de por qué no trabajaban actualmente.

Respecto a la satisfacción, el 72% de egresados se encontraba satisfecho con la formación de pregrado (top2); 62% declara que la PUCP es la mejor en su especialidad (Ciencias Administrativas y Economía consideran la mejor a otra reconocida universidad de Lima); 54% declara que su especialidad es altamente competitiva (es decir, que hay muchos profesionales calificados y, por tanto, mucha competencia por los puestos de trabajo); y 47% declara que pasar por los estudios generales es realmente importante (top2), con mayores proporciones entre quienes pasaron por estudios generales letras respecto de los egresados de estudios generales ciencias.

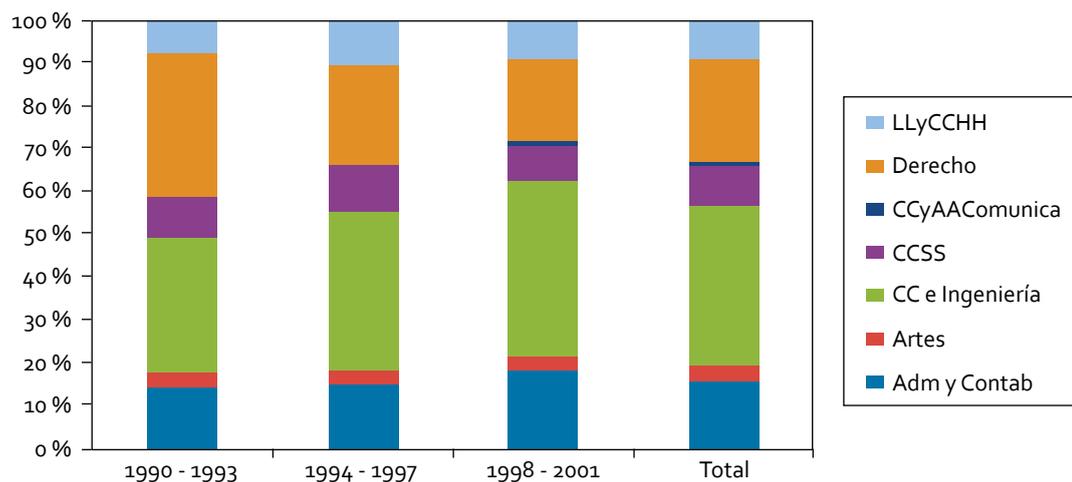
FIGURA 5
Vista resumen de la satisfacción con la PUCP (2003)



Fuente: Apoyo, 2003, p. 15.

FIGURA 6

Composición de la matrícula por facultad: estructura y evolución (2003)



Fuente: elaboración para este informe usando los datos de base.

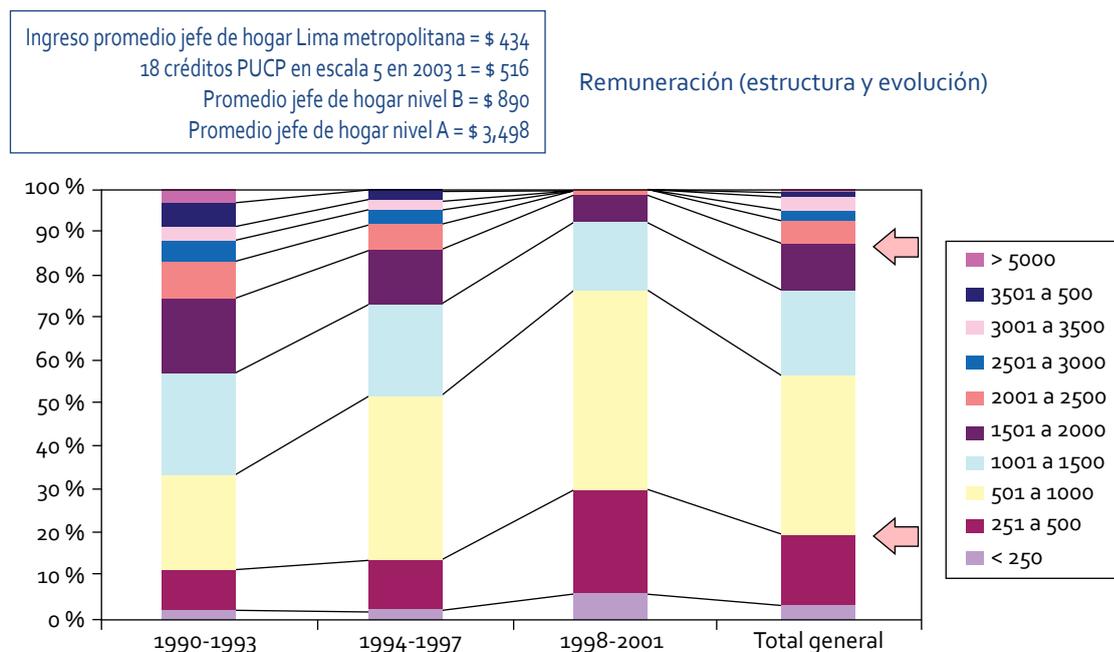
Siempre es conveniente recordar la naturaleza de los promedios, especialmente cuando se presentan series de tiempo. La figura siguiente muestra los problemas derivados de la composición por facultades según cuatrienios para la interpretación del promedio institucional.

La figura muestra que en este lapso de doce años creció la Facultad de Ciencias e Ingeniería y se redujo la de Derecho. Sin embargo, hubo un efecto sustitución, por lo que entre Derecho e Ingeniería explican dos tercios de la población. Cuando hablamos del promedio, entonces, todos los egresados de Arte, Administración y Contabilidad, Letras y Ciencias Humanas, Ciencias y Artes de la Comunicación y Ciencias Sociales comparten el otro tercio; y la mitad de ese tercio pertenece a Administración y Contabilidad, que también crece ligeramente.

La evolución por especialidad explica los cambios. Ciencias e Ingeniería crece respecto al total gracias al surgimiento de las especialidades de Electrónica e Informática, aunque se reduce la participación porcentual de Ingeniería Industrial, Mecánica, Civil y el grupo de ciencias. Entre las especialidades de letras, ya mencionamos la reducción de Derecho mientras creció Administración; Comunicaciones aparece en el periodo más reciente. Letras y Ciencias Sociales reducen su participación y Artes representa siempre un porcentaje pequeño. Dentro de la facultad de Letras y Ciencias Humanas crece Psicología.

FIGURA 7

Evolución de las remuneraciones en dólares por rango de ingreso según cuatrienios (2003)



Fuente: Vega, 2003, p. 6.

Los cambios en el tiempo son especialmente relevantes respecto a ingresos.

El perfil de remuneraciones se estimó satisfactorio, dado que el promedio y la mediana se ubicaban por encima del costo mensual promedio de la canasta del jefe de hogar en Lima metropolitana y de dieciocho créditos en la escala más alta de las pensiones en la PUCP (tarifa académica).

El análisis por cuatrienio obligó, sin embargo, a ponderar mejor los resultados promedio. La figura presenta el detalle de las remuneraciones en tramos de \$500 en el rango \$500 a \$5000 y una división de \$250 entre \$500 y menos. Además, la figura incluye dos flechas: la inferior marca el límite de los \$500 y la superior el límite de los \$3000. En ese momento \$500 era la canasta del jefe de hogar promedio en Lima metropolitana y el costo mensual estimado para un alumno regular en las escalas de pago más altas de la PUCP, de manera que \$500 nos pareció aceptable

como límite inferior. La figura muestra que, aunque en el promedio agregado el 20% estaba bajo el límite, en el grupo más viejo eran poco más del 10%, mientras en el grupo de egreso más reciente era poco más del 30%. Una buena noticia era el salto entre el cuatrienio más reciente y el segundo, con cerca de 20% de reducción en el porcentaje bajo del límite.

La flecha superior (frontera de \$3000) marca un monto cercano al ingreso promedio del jefe de hogar de nivel socioeconómico A en Lima. Prácticamente nadie superaba este límite en el grupo más joven, algo más de 10% en el segundo y más de 20% en el tercer grupo, lo cual es otro indicio de progreso en el tiempo.

El análisis de esta información suscitó más preguntas que respuestas. En particular, respecto del desempeño por facultad y especialidad. Aunque el reporte presentó algunas tablas con este detalle, mostrando una gran variabilidad y algunas pautas sorprendentes de semejanza en el comportamiento de algunas facultades, los datos eran solo indiciarios por el diseño muestral. Para tener mejor información, hubo que esperar al siguiente informe.

Segundo estudio (2004)

El segundo estudio solo reportó el quinquenio 1999-2003¹² y permitió distinguir la información por facultades. Conviene tener presente que, para comparar los agregados institucionales con el estudio anterior, habría que tomar en cuenta solo la información del último cuatrienio (1998-2001).

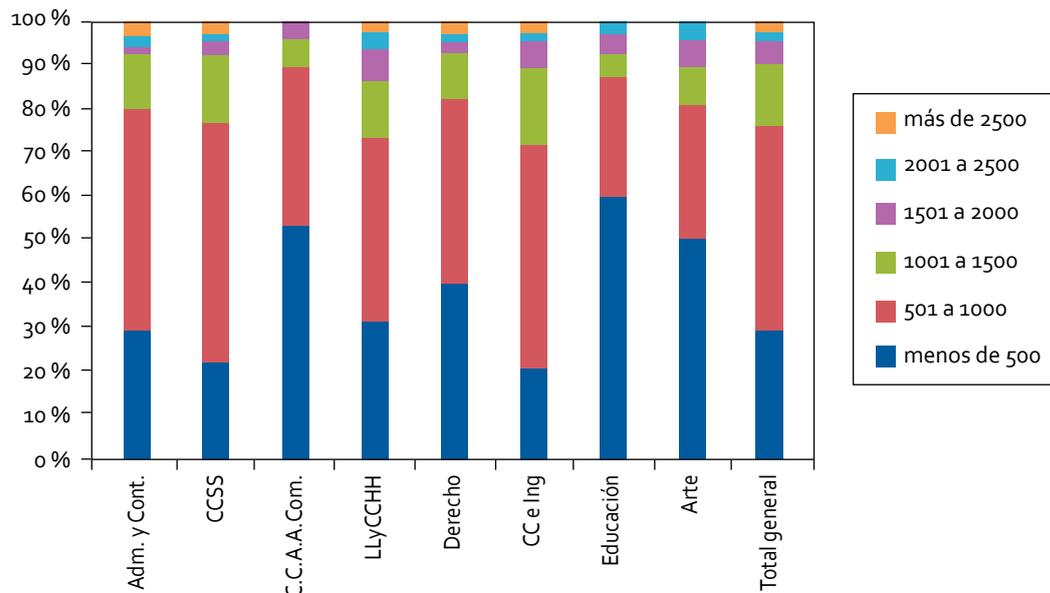
Con respecto al ingreso, el porcentaje bajo la frontera de los \$500 se había reducido en un punto en el promedio PUCP (29%) con relación al primer estudio, pero la distribución por facultades mostró que Educación (60%), Ciencias y Artes de la Comunicación (53%) Arte (49%) y Letras y Ciencias Humanas (39%) tenían más egresados por debajo del límite, mientras Administración y Derecho se ubicaban en torno al promedio y Ciencias Sociales (22%) y Ciencias e Ingeniería (20%) estaban mejor.

Otra vez conviene mirar el detalle detrás del promedio. Ciencias e Ingeniería, Derecho, así como Administración y Contabilidad siguen teniendo el mayor peso en la composición de la matrícula. Sin embargo, comprender el comportamiento de cada facultad requiere prestar atención a su composición por especialidad.

12 Otro grupo encuestado, de egresados más antiguos (1994-1998), tuvo muy baja tasa de respuesta.

FIGURA 8

Ingreso mensual promedio del empleo actual en dólares (2004)



Nota: el cálculo excluye a los que trabajan, pero no respondieron esta pregunta (21% del total).

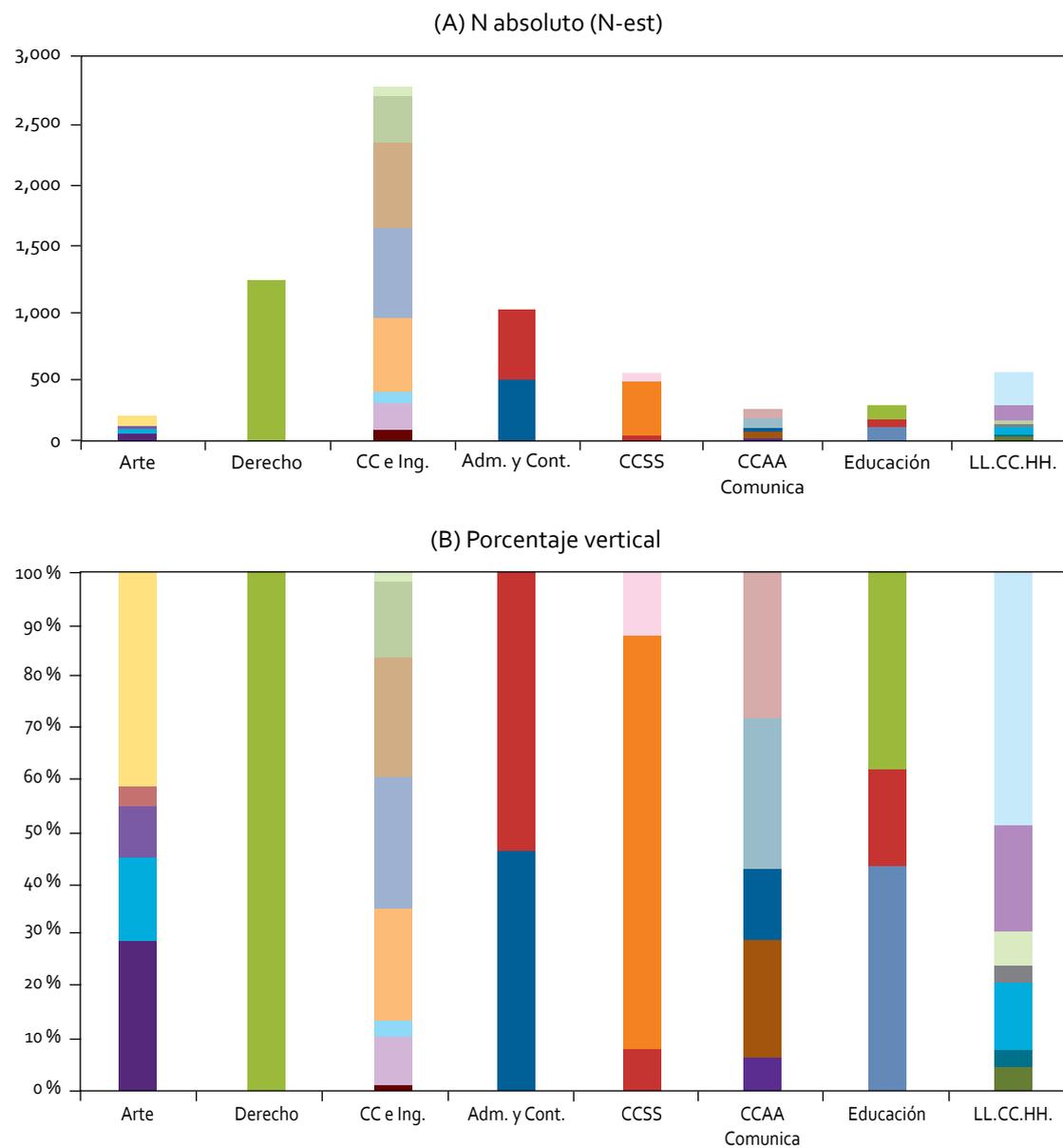
Fuente: elaboración para este informe sobre la base de N-est del segundo estudio. N-est o población estimada, es la población total resultado de la expansión de los valores muestrales obtenidos. (PUCP, 2006, p. 26).

En la estructura poblacional de Arte destacan con los valores mayores Pintura, que representa el 40% de la facultad, seguido por Diseño Gráfico (30%) y Diseño Industrial (15%); mientras que Escultura y Grabado tienen tamaño pequeño ($N < 30$). En Ciencias e Ingeniería, Física tiene tamaño pequeño y Matemática ni siquiera figura. En el caso de Ciencias Sociales, Economía domina la facultad, representando el 80% de su población. En Ciencias y Artes de la Comunicación, Artes Escénicas tiene tamaño pequeño. Y en Letras y Ciencias Humanas, Psicología representa el 48% de la facultad, mientras que Arqueología, Bibliotecología y Geografía tienen tamaño pequeño. En Educación, aunque Educación Primaria tiene el porcentaje menor de la facultad, su población es mayor a 30. Administración y Contabilidad, así como Derecho, no exigen comentario ni por distribución ni por población de tamaño pequeño.

El segundo estudio permitió apreciar la variabilidad de la satisfacción por facultades.

FIGURA 9

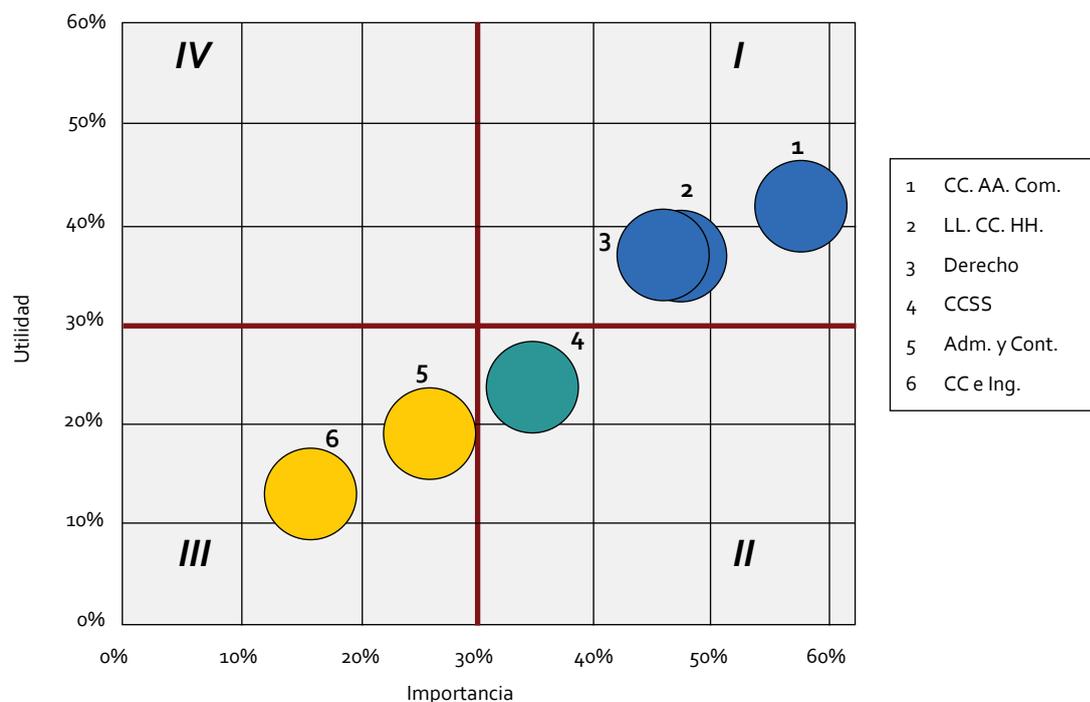
Composición de los egresados por facultad según especialidad (2004)



Fuente: elaboración para este informe sobre la base de N-est del segundo estudio (PUCP, 2006, p. 81).

FIGURA 10

Valoración de estudios generales según facultades (alta importancia y alta utilidad, 2004)



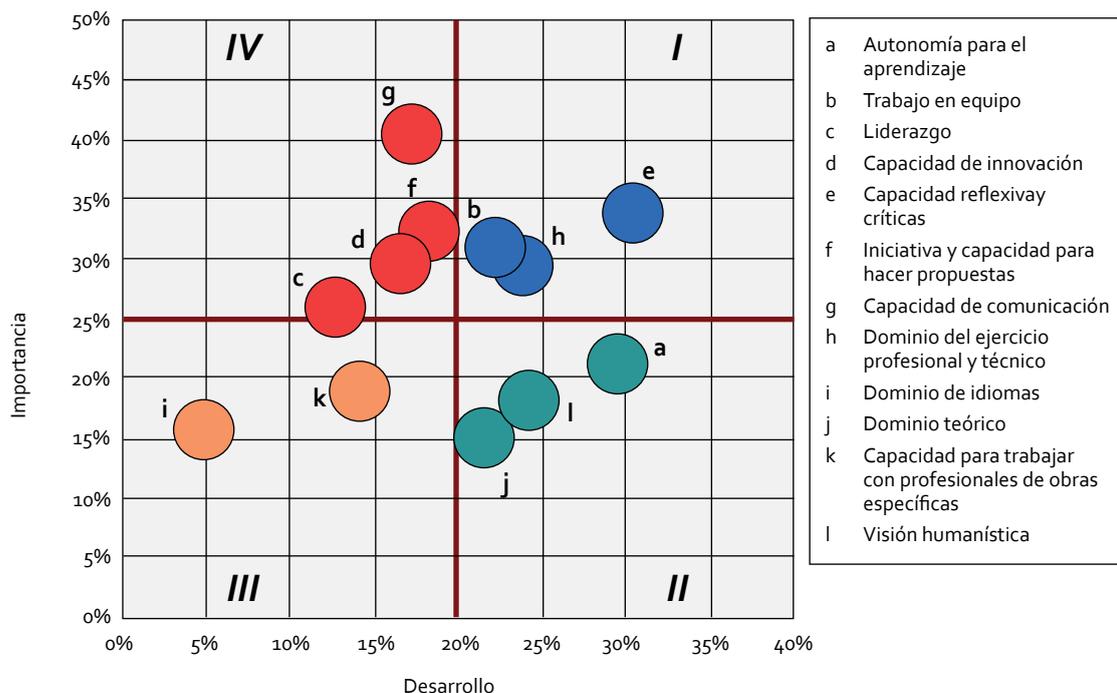
Nota: respecto a la satisfacción con los estudios generales (EEGG), se preguntó cómo evaluaban, en una escala Likert de cinco pasos, la importancia de los EEGG en la propia formación y la utilidad en el desempeño profesional. La figura registra, en el eje horizontal, el juicio sobre importancia en la formación y, en el eje vertical, el juicio sobre la utilidad; además, grafica el porcentaje de egresados que declararon que los EEGG eran “totalmente importantes” y “totalmente útiles para el desempeño profesional”. Las perpendiculares, sobre el 30%, dividen la figura en alto (> 30%) o bajo (< 30%) porcentaje de alumnos, resultando cuatro cuadrantes. El cuadrante superior derecho registra alta proporción de egresados con alta valoración en ambos juicios y, el cuadrante inferior izquierdo bajo, porcentaje en ambos juicios.

Fuente: PUCP, 2006, p. 77.

Lo sorprendente es que parece haber una correlación fuerte entre importancia y utilidad de pasar por EEGG, de modo que se alinean en una diagonal en la que las facultades de Ciencias y Artes de la Comunicación, Letras y Ciencias Humanas y Derecho se ubican en el cuadrante de altos porcentajes en ambos casos; Administración y Contabilidad y Ciencias e Ingeniería están en el cuadrante de bajos porcentajes en ambos casos; mientras que Ciencias Sociales queda cerca de estos últimos con porcentaje cercano al límite apenas por encima del 30% respecto a

FIGURA 11

Valoración de capacidades para el desempeño profesional según alta importancia y alto desarrollo (2004)



Nota: la figura representa en el eje horizontal el desarrollo y en el vertical la importancia. La escala registra en el eje vertical el porcentaje de egresados que declaró la capacidad "imprescindible" y, en el eje horizontal, el porcentaje que la declaró "totalmente desarrollada" gracias a la formación PUCP. Las perpendiculares dividen alto y bajo porcentaje, en 25% para la importancia y en 20% para el desarrollo. Así, el cuadrante superior derecho (I) registra alto porcentaje de importancia ("imprescindible") y de desarrollo ("totalmente desarrollada") y se identifica con el color azul; mientras que el cuadrante inferior izquierdo (III), bajo porcentaje para gran importancia y desarrollo, en color naranja. El cuadrante superior izquierdo (IV) es probablemente el más delicado y registra las capacidades de gran importancia, pero de bajo desarrollo, en color rojo. Fuente: PUCP, 2006, p. 68.

importancia y apenas por debajo respecto a la utilidad. Esta figura confirmó las pautas encontradas en el detalle del primer estudio sobre rasgos comunes entre Administración, Economía y las ingenierías, que no se alineaban en el vector tradicional de división en la PUCP según el canal de EEGG: Ciencias vs. Letras.

Por otro lado, se hizo en la encuesta una exploración semejante sobre una lista de doce capacidades, preguntándose respecto a cada una: si fue desarrollada por la formación y la

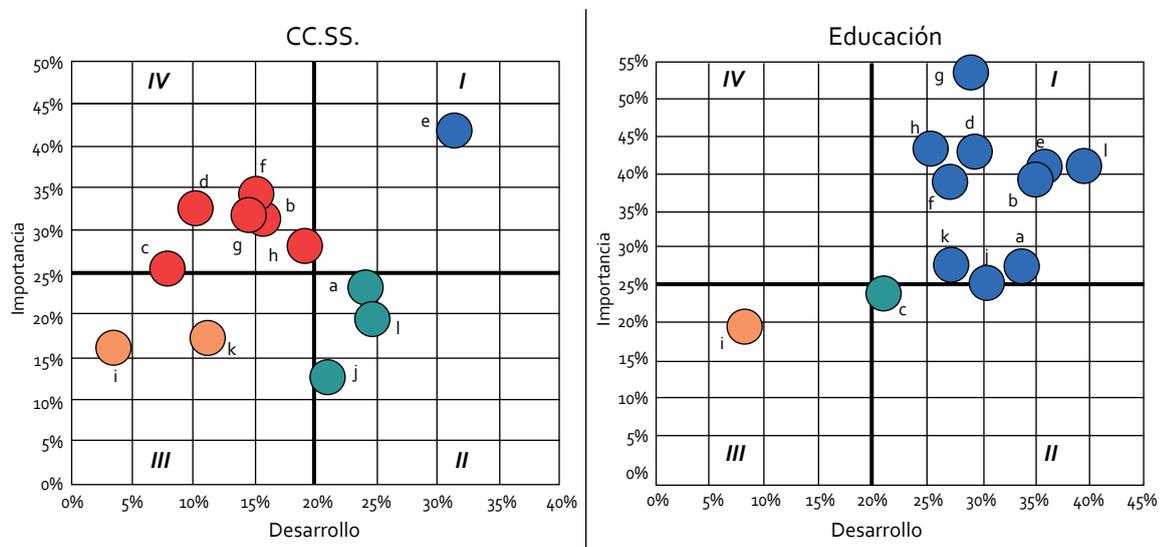
importancia para el desempeño laboral, con escalas Likert de cinco pasos respecto al desarrollo y de diez pasos sobre la importancia.

Las capacidades que a nivel institucional obtienen altos porcentajes en ambos frentes (azul) son: "capacidad reflexiva y crítica", "trabajo en equipo" y "dominio del ejercicio profesional y técnico". Obtienen bajos porcentajes en los dos frentes (naranja): "capacidad de trabajo interdisciplinario" e "idiomas". En el cuadrante más sensible (rojo) figura la "capacidad de comunicación", que tiene el más alto porcentaje de importancia (40%), pero tiene un bajo porcentaje de desarrollo (17%), acompañada de "iniciativa y capacidad de hacer propuestas", "innovación" y "liderazgo". "Dominio teórico" junto con "autonomía para el trabajo" y "visión humanista" se encuentran en el cuadrante inferior derecho (verde); es decir, se logran desarrollar, pero no se consideran tan importantes para el desempeño profesional.

Hay que recordar que estos valores institucionales expresan lo que suele suceder en el mercado laboral de la mayoría de egresados PUCP. Para entenderlo conviene revisar la composición del total de egresados por facultad y prestar atención al comportamiento de estos indicadores por facultad.

FIGURA 12

Facultades. Valoración de capacidades para el desempeño profesional según alta importancia y alto desarrollo (2004)

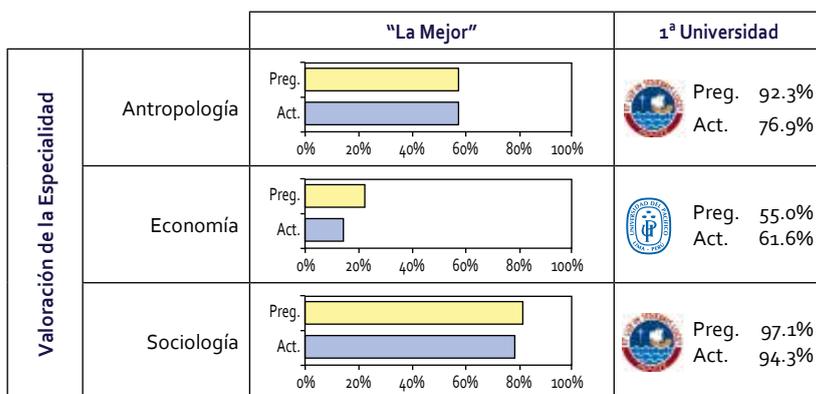
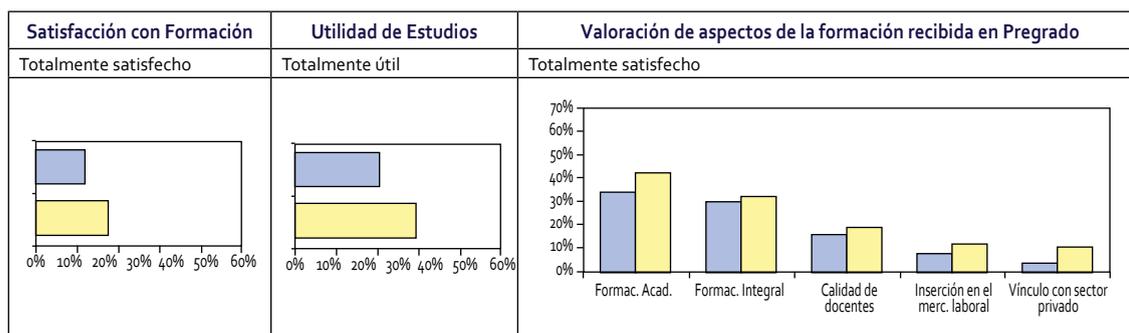


Fuente: PUCP, 2006, pp. 68-69.

Estas figuras ejemplifican cómo distintas facultades declaran muy diversas combinaciones de importancia para el trabajo/desarrollo alcanzado en la formación para este conjunto de capacidades. A modo de ejemplo —y en los extremos—, podemos observar a Ciencias Sociales que solo registra una (capacidad reflexiva y crítica) en el cuadrante I de alta importancia/desarrollo, mientras seis se acumulan en el cuadrante IV de alta importancia/bajo desarrollo. Por otro lado, en el otro extremo, Educación declara diez de doce capacidades en el cuadrante I. Una lectura lineal diría que el plan de estudio de Educación tiene el mejor ajuste al desempeño posterior de los egresados.

FIGURA 13

Ciencias Sociales: resumen de la satisfacción con la formación (2004)



Fuente: PUCP, 2006, pp. 82.

No conviene perder de vista, sin embargo, lo que señalamos respecto a la composición demográfica de las facultades. El promedio de Ciencias Sociales lo define Economía, que explica el 80% de la población y tiene los más bajos valores de satisfacción.

Tercer estudio (2008)

A partir del tercer estudio, ya censal —realizado en 2008 sobre la cohorte 2006— y con el cuidado señalado que requieren las especialidades que tienen población muy pequeña, pusimos más énfasis en el detalle de las especialidades.

FIGURA 14

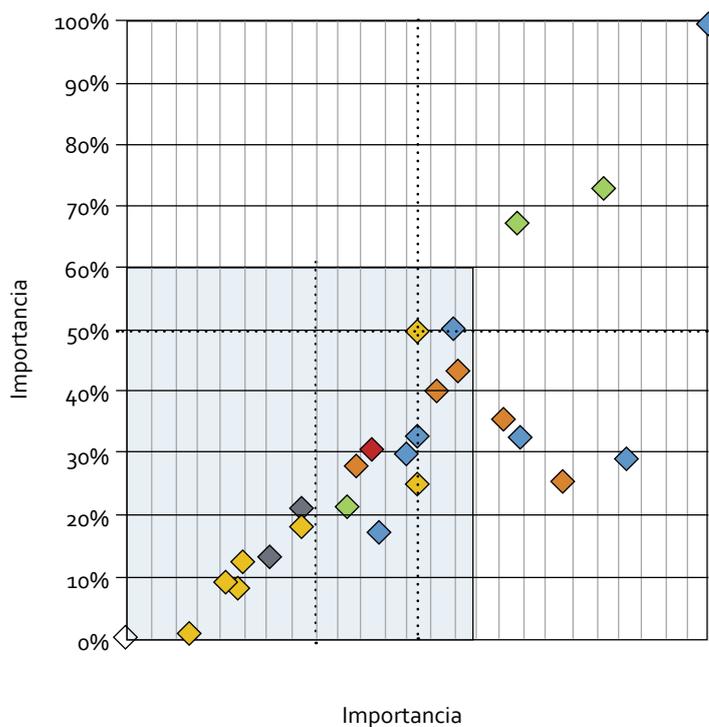
Valoración de estudios generales según facultades según alta importancia y alta utilidad (2008)

Leyenda

Los puntos representan las especialidades y el color identifica las facultades.

- Azul oscuro: Adm. y Cont.
- Mostaza: CC e Ing.
- Verde: CCSS
- Naranja: CC.AA. Com.
- Guinda: Derecho
- Azulino: LL.CC.HH.

La información no incluye Arte ni Educación, ya que no participan en los programas de EEGG. Arquitectura no se reportó por baja tasa de respuesta.



Nota: el recuadro delimita el área de la tabla equivalente de 2004.

Fuente: elaboración para la ponencia sobre la base de los datos (PUCP, 2010, pp. 28-29).

Esta figura representa el porcentaje que declara alta importancia y alta utilidad de los EEGG a nivel de especialidad. No pretende identificar cada especialidad, sino ver la distribución de las especialidades según facultad. Para identificar a cada facultad, se adoptó un código de colores.

La distribución muestra a la mayoría de especialidades de Ciencias e Ingeniería agrupadas hacia abajo con valores menores a 30% en ambos frentes, junto a las dos especialidades de Administración y Contabilidad. Las demás especialidades se ubican encima de ese límite, con Derecho con su único punto verde cerca (42,30). También es posible observar que en esta ocasión Ciencias y Artes de la Comunicación (naranja) se ha desplazado hacia el centro y ya no lidera el grupo.

Finalmente, cabe observar la dispersión por especialidad en las grandes facultades. Ciencias e Ingeniería va desde Telecomunicaciones (11,0) hasta Química (50,50). Son llamativas especialmente las dispersiones de Letras y Ciencias Humanas, Ciencias y Artes de la Comunicación, así como de Ciencias Sociales. Conviene tener presente que en Letras los valores inferiores los tienen las tres especialidades de Psicología que tienen gran importancia demográfica. El caso de Ciencias Sociales es llamativo porque Economía (66,38) se comporta como Ciencias y Administración, mientras las otras dos especialidades lideran el porcentaje de declaración de importancia y utilidad: Sociología (67,67) y Antropología (82,73). En general, las especialidades que no son de Ciencias o Administración están corridas hacia el extremo de mayor importancia (mayor de 50%); pero el techo de utilidad está en 50%, salvo las dos especialidades de Ciencias Sociales mencionadas y cuatro individuos de Lingüística Hispana (100/100).

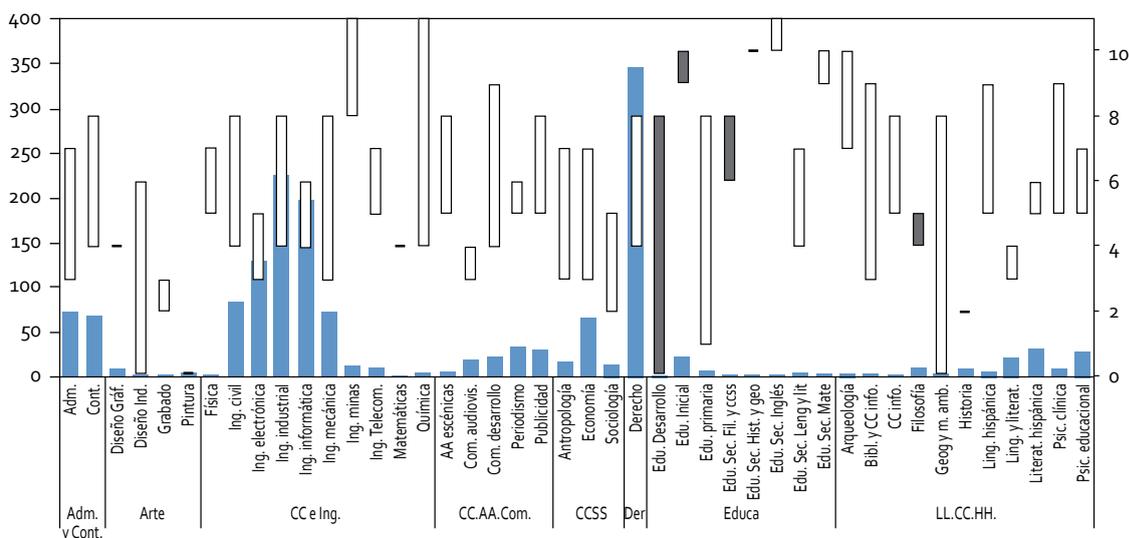
La figura 15 resume la evaluación del nivel de desarrollo en la formación PUCP y de la importancia para el primer empleo de las once capacidades para cuarenta y cuatro especialidades. Hemos resumido esta evaluación graficando el número de capacidades que obtienen valores iguales o mayores a 4; es decir, alto desarrollo y alta utilidad.

La figura muestra tres datos. En la escala vertical de la izquierda, muestra el tamaño de la población de la especialidad (en barras azules) y, en la escala vertical de la derecha, el número de capacidades con alto desarrollo o utilidad en forma de gráfico de cajas. Como señalamos, conviene siempre tener presente el peso demográfico, el cual no solo indica qué especialidades tienen población pequeña, sino cuáles pueden sesgar la información de los agregados por facultad o del total. Casi todas las especialidades (32/44) están bajo los treinta egresados.

Cabe agregar que la experiencia censal nos enseñó la necesidad de atender a problemas adicionales y nuevos que tienen que ver con las tasas de respuestas. En el censo hay que entrevistar a todos, no hay tasas de reemplazo. La idea del muestreo es que los reemplazos distribuyen los

FIGURA 15

Número de capacidades consideradas desarrolladas por la formación PUCP e importantes para del desempeño en el primer empleo, por especialidad - 2006 (2008)



Fuente: elaboración para la ponencia sobre la base de los datos (PUCP, 2010, pp. 28-29).

sesgos de autoselección y las respuestas no son afectadas. Por supuesto, altas tasas de reemplazo exigen cautela; pero la falta de respuesta en un censo podría implicar un sesgo. Esto es más importante cuanto mayor es la tasa de no respuesta. Hemos mejorado mucho en este tema.

Respecto a la figura de cajas, el color de la caja es blanco cuando el número de capacidades consideradas importantes para el primer empleo es mayor que el número de capacidades consideradas desarrolladas (este es el caso en la mayoría de especialidades). El color es negro cuando el número de especialidades consideradas desarrolladas es mayor que el número de las consideradas importantes para el empleo (esto sucede en tres especialidades de Educación y en Filosofía). En los extremos, Ingeniería de Minas, Química y Educación Secundaria Inglés consideran que las once capacidades han sido importantes para su primer empleo, mientras Pintura considera que ninguna lo fue. Respecto al alto desarrollo durante la formación, los más altos valores corresponden a Educación Inicial y Educación Secundaria Inglés con 10 capacidades y Educación Secundaria Matemática con 9; en el otro extremo, Diseño Industrial, Pintura y Geografía consideran que ninguna fue desarrollada en la formación.

En general, las capacidades que más disciplinas declaran como importantes para el trabajo son las capacidades de comunicación (39 de un total de 44: 89%) y resolución de problemas (86%). Solo en Arte, Ciencias Sociales y Educación no le dan mayor valor a la capacidad de comunicación en todas sus especialidades; mientras resolución de problemas no es declarada como altamente importante en Arte, Ciencias Sociales, Educación y Letras y Ciencias Humanas. Entre las capacidades que más egresados declaran altamente desarrolladas por la PUCP figuran análisis y síntesis (84%) y capacidad reflexiva y crítica. En el primer caso, convergen Ciencias e Ingeniería, Educación y Letras y Ciencias Humanas; mientras que en el segundo, solo Educación y Letras y Ciencias Humanas.

Los recuadros muestran el valor más alto y el más bajo respecto al número de habilidades declaradas desarrolladas e importantes dentro de las facultades, ofreciendo al mismo tiempo un indicio de la dispersión en los valores. La mayor dispersión la tiene Educación y Letras y Ciencias Humanas; mientras que la menor, Derecho y Ciencias Sociales. El límite inferior es más alto en Administración y Contabilidad, Ingeniería y Ciencias y Artes de la Comunicación. El límite superior es más alto en Ciencias e Ingeniería, Educación y Letras y Ciencias Humanas.

Informe comparativo 2006-2010 (2015)

A partir del tercer informe, como hemos indicado, se pasó a una administración censal de la encuesta, a una sola cohorte por año de manera continua, acumulando datos de diversos años y haciendo posible comparar información entre diferentes cohortes de egresados (por ejemplo, la situación laboral o las opiniones respecto a la PUCP). El primer informe de esta naturaleza comparó los resultados obtenidos para las cohortes de 2006 y 2010, entre 1 y 2 años de egreso; y, en el caso de sueldos al momento del censo, se pudo comparar con valores constantes al año 2012, en que se aplicó la encuesta a la cohorte 2010.

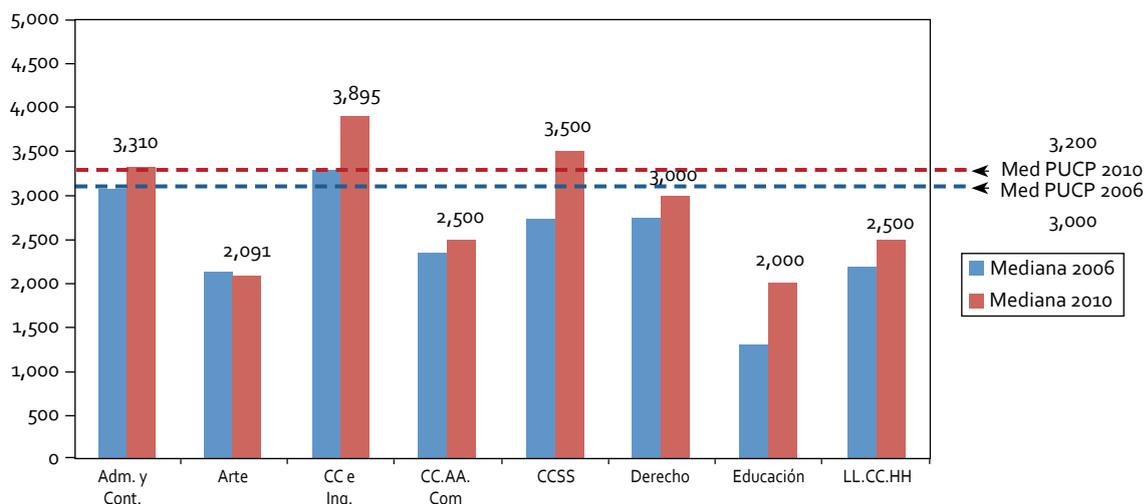
Para el reporte detallado por ingresos, se usó la mediana, que divide a la población en dos mitades ordenadas y, a diferencia de los promedios, no es afectada por valores extremos registrados en casi todos los casos.

La figura muestra que, salvo en Arte, hay un incremento en el periodo en todas las facultades y en el promedio PUCP. El incremento agregado institucional fue de 4%, mientras Educación (52%), Ciencias Sociales (27%) y Ciencias e Ingeniería (18%) tuvieron incrementos mucho mayores.

Muestra también que Ciencias e Ingeniería, Ciencias Sociales y Administración y Contabilidad tienen, en 2010, valores superiores a la mediana PUCP. En el otro extremo figuran Educación y Arte.

FIGURA 16

Mediana de ingreso mensual de empleo actual por facultad, para las cohortes 2006 y 2010, en soles contantes de 2012



Fuente: elaboración para este informe a partir del reporte (PUCP, 2015, pp. 102-103).

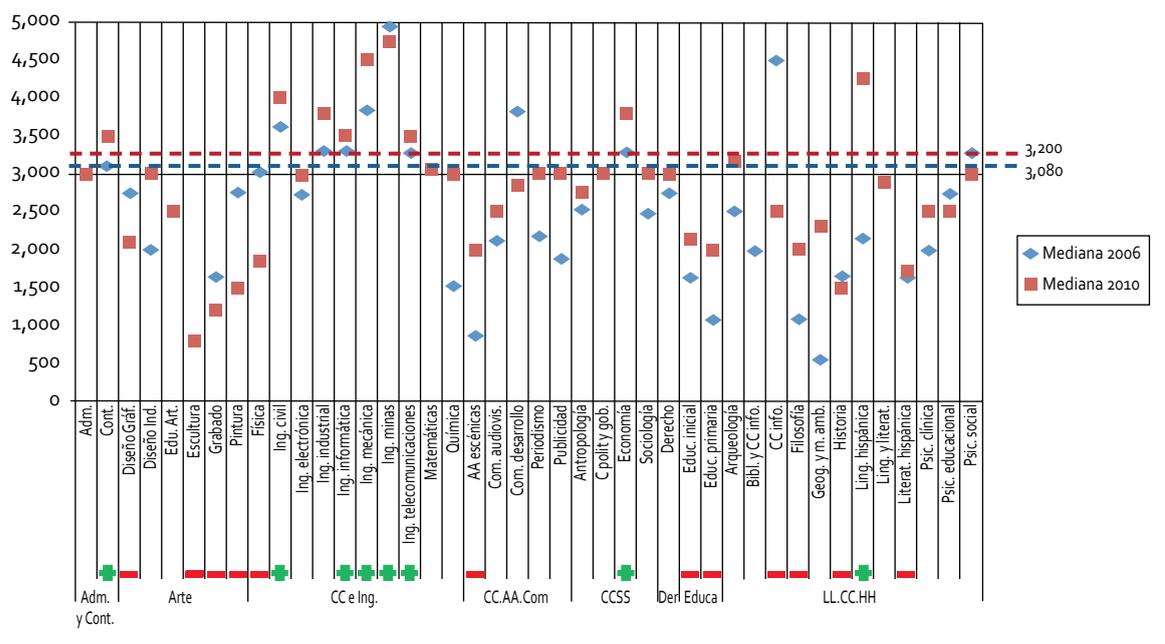
Las reflexiones previas sobre la distribución de los egresados según facultades invitan a descomponer esta información por disciplina. Y lo que se muestra es una variabilidad mayor a la registrada entre facultades, verificando las cautelas mencionadas sobre el impacto del peso demográfico en los promedios y otras medidas de tendencia central, incluso la mediana.

La figura 17 muestra cómo las medianas de ingreso por disciplina tienen una distribución mucho más dispersa. Lo más destacable es la concentración de valores superiores en la facultad de Ciencias e Ingeniería y las disciplinas de Contabilidad, Economía y, contra todo lo esperado, en Lingüística Hispánica (tres egresados en 2010). Por otro lado, las facultades de Letras y Ciencias Humanas, Educación y Arte concentran valores inferiores (rojos). Y, finalmente, Arte, Ciencias y Artes de la Comunicación, Derecho y Educación solo tienen valores debajo de la mediana.

También se puede comentar las distancias entre los valores de 2006 y 2010. En la mayoría de las especialidades, la mediana es mayor en 2010 y, en algunos casos, bastante mayor: Geografía (4,2 veces), Artes Escénicas (2,3 veces), Lingüística Hispánica (2 veces), Química y

FIGURA 17

Mediana de ingreso mensual de empleo actual por disciplina, para las cohortes 2006 y 2010, en soles contantes del 2012



Nota: las marcas en el área de títulos del eje horizontal señalan las especialidades que se ubican por encima y por debajo de la mediana PUCP en 2010. Las marcas verdes indican valores superiores; las marcas rojas indican valores al menos 1/3 por debajo de la mediana —es decir, las de menor ingreso—; las que no tienen marca están por debajo, pero a menor distancia.

Fuente: elaboración para este informe a partir del reporte (PUCP, 2015, pp. 102-103).

Filosofía (1,9 veces) y Educación Primaria (1,8 veces). Pero conviene tener presente dos hechos: se trata de las especialidades con ingresos bajos en 2006 y son especialidades de población pequeña. Por otro lado, entre las que las que tienen mayor reducción están Pintura, Ciencias de la Información, Física y Grabado (en torno a la mitad), que son también de tamaño pequeño y algunas están entre las de más bajos ingresos en 2010 que se concentran en la facultad de Artes.

Como se puede observar en las figuras por especialidad, comienza a verificarse ciertos patrones, pero también una gran dispersión que hace difícil el análisis. Para enfrentar este hecho, aplicamos un método de *Data Mining* para construir una tipología sobre la base de rasgos comunes emergentes de los datos acumulados.

Tipología de especialidades (2016)

Para el ejercicio se trabajó la información de un lustro (2009-2013) y se hizo un análisis de clústeres. La virtud de este procedimiento es que permite identificar estructuras presentes en la información sin partir de ideas previas. La unidad de observación fue la especialidad. A pesar del lustro, algunas especialidades siguieron teniendo valores muy pequeños por lo que se las agrupó reduciéndose el número de casos de 51 a 35 especialidades que incluyen siete grupos que representan 25 especialidades —Diseño, 2 especialidades; Arte, 4; Ciencias, 3; Educación, 5; Gestión Institucional y Pública, 2; Humanidades, 5; y Lingüística y Literatura, 3—. Es importante destacar que para este análisis, al asumir a la especialidad como unidad de observación, el tema del tamaño de la especialidad deja de tener relevancia.

Se estudió una base de 76 variables del SSE, por especialidad, sin incluir la información sobre satisfacción, y algunos datos administrativos (escala de pagos, porcentaje de créditos aprobados y porcentaje de graduados a tiempo). El análisis de correlación permitió reducir las a 29, organizadas en cinco ejes: características del egresado (dos variables), características de la especialidad (2), inserción laboral (11), situación laboral al momento del censo (12) y estudios de posgrado (2). Se realizaron varios análisis con distintos algoritmos y distinto número de clústeres. El mejor ajuste lo obtuvo el procedimiento K-medias para cuatro clústeres. La siguiente tabla presenta los resultados:

El clúster 1, con un mayor predominio de docencia, agrega a especialidades con mayor proporción de egresados de las escalas más altas, con menor número de horas de trabajo, menor ingreso promedio y mayor número de trabajadores independientes. También tienen mayor proporción de trabajo en la universidad y estudios de maestría, pero más énfasis en formación práctica. En general, se trata de las carreras de Artes, Letras y Ciencias Humanas (salvo las psicologías educacional y social) y Comunicaciones (salvo Comunicación para el Desarrollo y Publicidad). Este grupo tiene también una mayor proporción de egresados que se considera infraeducado para realizar las actividades del primer empleo luego de egresar.

El clúster 2, de especialidades con predominio en investigación, tiene una mayor frecuencia en el empleo en ONG y menor proporción de desempleados, comparado con los demás clústeres. No extraña que declaren una mayor importancia para el primer empleo de las capacidades de abstracción, análisis y síntesis y tengan una mayor proporción de egresados que culminaron sus estudios en tiempo ideal. Figuran en este clúster las carreras de Arquitectura, el conjunto agrupado de las carreras de ciencias (Física, Matemática y Química), dos de las cuatro de ciencias sociales (Antropología y Sociología), y las carreras de Letras y Ciencias Humanas y

TABLA 2

Clústeres: visión de conjunto

	Especialidades	Características
1	Diseño, Arte, Artes Escénicas, Comunicación Audiovisual, Periodismo, Educación, Humanidades, Filosofía, Lingüística y Literatura, Psicología Clínica.	<ul style="list-style-type: none"> • Promedio del porcentaje de créditos en la unidad. • Escala 4 y 5. • Maestría. • Docencia. • Independientes. • Universidad. • Promedio de horas de trabajo. • Ingreso promedio. • Mecanismo de inserción: referencia de empleadores. • Mecanismo de inserción: amigos, conocidos/referencias de egresados, profesores.
2	Arquitectura y Urbanismo, Ciencias, Antropología, Sociología, Comunicación para el Desarrollo, publicidad, Psicología Educativa, Psicología Social.	<ul style="list-style-type: none"> • Culminaron estudios en tiempo ideal. • No buscan trabajo. • Investigación. • ONG. • Mecanismo de inserción: amigos, conocidos/referencias de egresados, profesores. • Importante para el primer empleo: capacidades como abstracción, análisis y síntesis. • Infraeducados.
3	Administración, Contabilidad, Ingeniería de las Telecomunicaciones, Electrónica, Ingeniería Industrial, Informática, Ciencias Políticas y Gobierno, Economía, Derecho, Gestión Empresarial, Institucional, Pública y Social.	<ul style="list-style-type: none"> • Hombres. • Escala 4 y 5. • Sector privado. • Sector público. • Mecanismo de inserción: bolsa de ttrabajo PUCP. • Aprecio de servicio PUCP de inserción laboral.
4	Ingeniería Civil, Ingeniería de Minas, Ingeniería Mecánica-Mecatrónica.	<ul style="list-style-type: none"> • Hombres. • Diplomado. • Sector privado. • Promedio de horas de trabajo. • Ingreso promedio. • Especialización de su trabajo. • Importante para el primer empleo: conocimientos técnicos y específicos. • Experiencia laborar antes de egresar.

Fuente: Wiener & Vega, 2016.

Comunicaciones que no figuran en el clúster 1. Comparte con el clúster 1 que una proporción mayor se considera infraeducados respecto de los clústeres 3 y 4.

El clúster 3 incluye a las carreras de Administración y Contabilidad, Gestión, Derecho, la mayoría de las de ingeniería (salvo las tres del clúster 4) y las dos restantes de ciencias sociales (Economía y Ciencias Políticas). Las caracteriza el predominio masculino, la menor presencia de escalas 4 y 5, el trabajo dependiente en instituciones privadas o públicas y una mayor importancia de la bolsa de trabajo PUCP como mecanismo de inserción, así como un mayor aprecio del servicio PUCP para la inserción laboral. Como veremos, este es el clúster mayoritario en términos poblacionales.

Finalmente, el clúster 4, constituido por tres ingenierías (Civil, Minas y Mecánica-Mecatrónica), se distingue por la mayor importancia de los conocimientos técnicos y específicos, la especialización de su trabajo y un mayor promedio de horas de trabajo e ingresos. Su trabajo se realiza en empresas privadas, cuentan con una mayor proporción de diplomados y comparte con el clúster 3 el predominio masculino y mayores ingresos promedio.

Análisis de indicadores de satisfacción por clúster

Como indicamos previamente, la tipología no incluyó la sección sobre valoración de la PUCP del SSE. Así, es posible observar qué tipo de relación existe entre las condiciones materiales más objetivas y los aspectos valorativos más subjetivos. Aunque conviene tener presente que la valoración incluye juicios sobre utilidad en el mundo laboral.

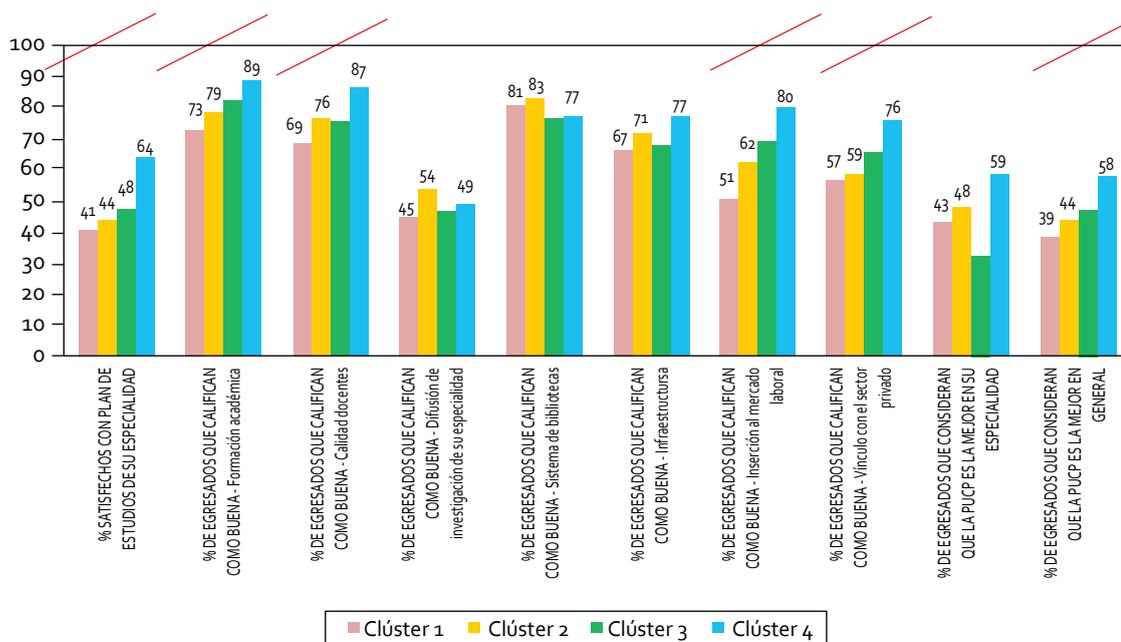
La figura nos proporciona una visión de conjunto sobre la valoración por clúster de distintas dimensiones de la PUCP. Las barras indican los porcentajes, pero también las variables en que están ordenados de menor a mayor según clúster.

Así, resulta que las especialidades más académicas (1 y 2) están menos satisfechas que las profesionales respecto del plan de estudios de su especialidad, la formación académica, la calidad de los docentes, la inserción en el mercado laboral, los vínculos con el sector privado y el juicio de que la PUCP es la mejor universidad.

En dos indicadores, el clúster identificado con investigación rompe la serie: difusión de investigaciones de su especialidad y biblioteca. Respecto de bibliotecas, los indicadores son altos en todos los clústeres, con una ligera ventaja para los académicos (1 y 2). Finalmente, respecto al juicio de la propia especialidad como la mejor, llama la atención que el valor del clúster 3 sea el más bajo (32%) en el ítem, pero también en toda la tabla.

FIGURA 18

Valoración de las capacidades según clúster



Fuente: Wiener & Vega, 2016.

Salvo en el caso del juicio sobre difusión de investigación y evaluación de las bibliotecas, el clúster 4 es el más satisfecho, liderando todos los indicadores. El clúster 1 muestra los valores menores en todos los indicadores, salvo respecto del porcentaje que considera que la PUCP es la mejor en su especialidad en que el valor del clúster 3 —profesional (y mayoritario)— es menor.

Un rasgo general que conviene tener presente es que los porcentajes son más altos respecto del juicio sobre formación académica, sistema de bibliotecas y calidad de docentes, y menores respecto de difusión de la investigación, satisfacción con el plan de estudios, la excelencia de la PUCP en su especialidad.

En la figura 19 figuran los resultados de la evaluación de la importancia para el primer empleo y del desarrollo en la PUCP de las capacidades estratégicas.

En general, los porcentajes respecto al nivel de desarrollo en la PUCP son menores a los de la importancia en el trabajo, salvo la capacidad de investigación, que en todos los clústeres

tienen un mayor porcentaje en desarrollo PUCP que en demanda laboral. Por otro lado, en ambos casos los valores menores corresponden a dominio de idiomas.

El análisis por clúster muestra un menor número de porcentajes alineados de menor a mayor (línea roja continua), aunque es posible distinguir también un comportamiento diferenciado entre los clústeres académicos —1 y 2— y los profesionales —3 y 4— (línea punteada azul).

“Liderazgo” y “resolución de problemas” están alineados, con una gradiente de menor en clúster 1 a mayor en clúster 4, respecto al desarrollo en la formación y, respecto de la importancia para el empleo, la resolución de problemas se comporta igual, mientras el liderazgo muestra un comportamiento diferenciado entre clústeres académicos y profesionales.

“Dominio de idiomas” y “dominio de tecnologías” muestran comportamiento diferenciado entre académicos (menor porcentaje) y profesionales (mayor), tanto en la formación como en la importancia para el primer empleo.

“Autonomía para el aprendizaje” y “trabajo en equipo” tienen un comportamiento creciente del clúster 1 al 4 respecto a la importancia en el primer empleo; pero, respecto al desarrollo en la PUCP, su comportamiento es bastante plano con el clúster 1 por debajo.

El clúster 1 destaca en “creatividad”, probablemente asociado a la presencia de artes en el grupo, con el grupo 4 en el segundo puesto, tanto respecto a desarrollo como a importancia (compartido con clúster 2). El perfil tiene forma de “U”. El valor menor lo tiene el clúster 3.

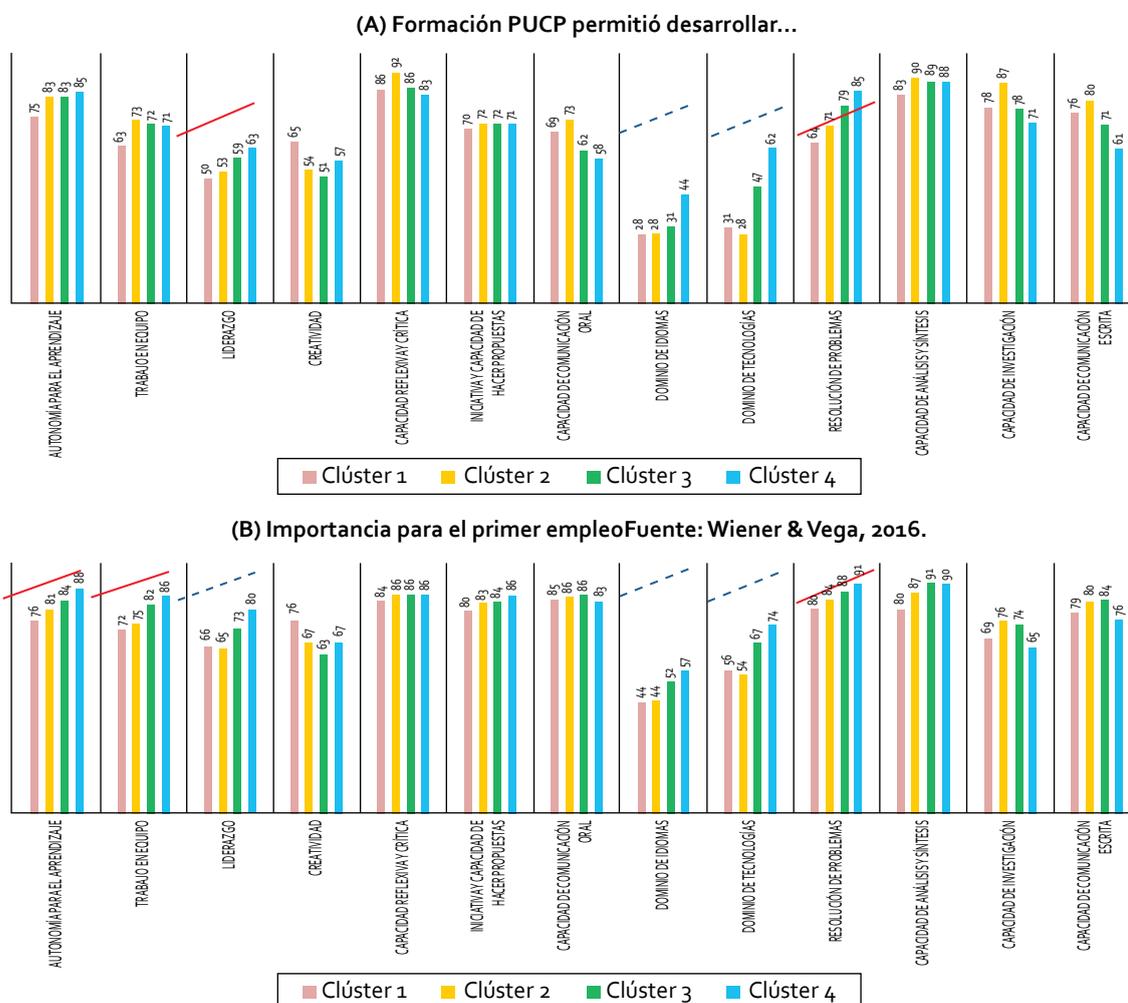
“Capacidad reflexiva y crítica”, “iniciativa y capacidad de propuesta” y “capacidad de comunicación oral” tienen un perfil bastante plano en términos de la importancia para el primer empleo, con un comportamiento igualmente plano en el desarrollo solo para “iniciativa y capacidad de propuesta”. El clúster 2 obtiene mejores valores en la evaluación del desarrollo en la PUCP de la “capacidad reflexiva y crítica”, de “comunicación oral”, de “investigación” y de “comunicación escrita”; la distribución de valores en estas capacidades tiene forma de “U” invertida con el pico en el clúster 2 y valores menores en los clústeres 3 y 4 en el desarrollo; y el clúster 1 y, el clúster 6, en la importancia para el empleo.

Respecto al desarrollo PUCP de capacidades, el clúster 1 obtiene los valores menores de “autonomía para aprendizaje”, “trabajo en equipo”, “liderazgo”, “iniciativa y propuesta”, “dominio de idiomas” (junto al clúster 2), “resolución de problemas” y “capacidad de análisis y síntesis”. El clúster 2 obtiene el menor valor en “dominio de tecnologías”; el clúster 3, en “creatividad”; y el clúster 6 obtiene el menor valor en capacidad de “comunicación oral”, “investigación” y “comunicación escrita”.

Respecto a la importancia para el primer empleo, el clúster 1 obtiene los valores más bajos en “autonomía para el aprendizaje”, “trabajo en equipo”, “capacidad reflexiva y crítica”, “dominio de idiomas” (junto al clúster 2), “resolución de problemas” y “capacidad de análisis y síntesis”. El clúster 2 obtiene el menor valor en “liderazgo” y “dominio de tecnologías”; el clúster 3, en “creatividad”; y el clúster 6 obtiene el menor valor en “comunicación oral”, “investigación” y “comunicación escrita”.

FIGURA 19

Desarrollo de las capacidades e importancia para el empleo según clúster



Fuente: Wiener & Vega, 2016.

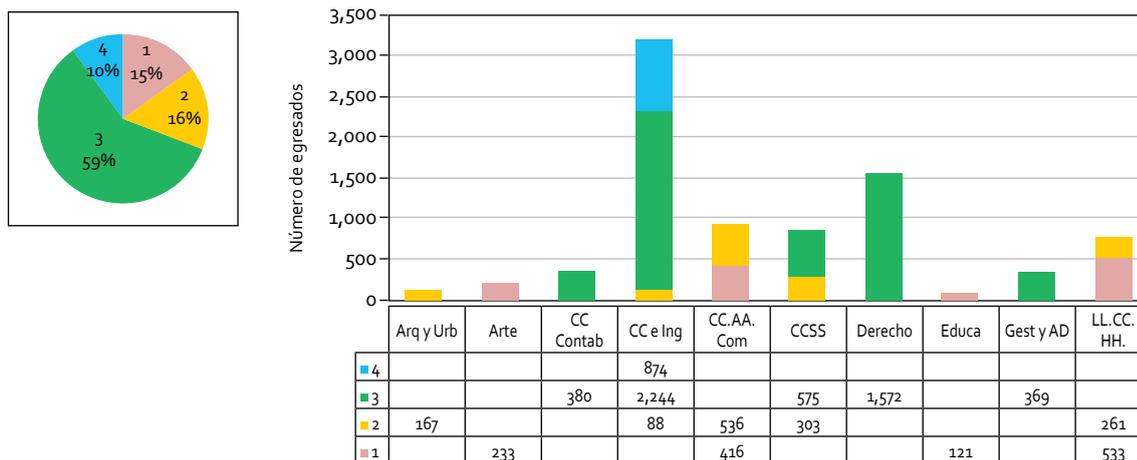
Los clústeres y la demografía PUCP

Aunque para el análisis de clúster no importa el tamaño porque la idea es identificar comportamientos semejantes con independencia de la escala, una vez identificados los clústeres resulta interesante comprobar que la PUCP, según el SSE, forma mayoritariamente para los clústeres de perfil más profesional (69% en los clústeres 3 y 4).

Respecto a la composición de las facultades, es interesante distinguir entre las de composición mixta y las homogéneas. La más variada es Ciencias e Ingeniería. Pero conviene tener presente que al menos en Derecho la homogeneidad resulta de la inexistencia de especialidades en el pregrado y que, en algunos otros casos, debido al bajo número de casos, se construyó una categoría agregada.

FIGURA 20

Peso poblacional y presencia de los clústeres según facultad



Fuente: realizado para la ponencia sobre la base de los datos del informe.

A MODO DE CONCLUSIÓN

No tenemos una respuesta clara a las preguntas de Leclercq sobre el aporte del SSE a la evaluación de la formación por competencias; más bien tenemos nuevas preguntas.

Yendo de las condiciones materiales (empleabilidad, niveles 3 y 4) a las subjetivas (satisfacción, nivel 1), la imagen de que la empleabilidad podría ser la prueba ácida del éxito de todo

programa de formación requiere de muchos matices. La empleabilidad no depende solo, ni necesariamente de manera predominante, de las competencias desarrolladas en la universidad. Depende del comportamiento de los mercados y los mercados laborales tienen diferencias según el tiempo (ciclos económicos) y el espacio (jerarquía territorial/escalas en el espacio). Pero, además, los filtros de acceso en los mercados no dependen tampoco única ni principalmente del *know how*; sino que son influidos fuertemente por el *know who*¹³, en términos de capital social o redes sociales, derivados de nivel socioeconómico, género o cultura del egresado, pero también de los filtros del demandante. A este paquete pertenece el efecto institucional (empresa) que hace que egresados más competentes de universidades de bajo prestigio tengan empleabilidad menor que egresados menos competentes de universidades de alto prestigio.

La variabilidad en la satisfacción, por otro lado, no siempre sigue la dinámica de la empleabilidad. A este respecto encontramos distintos índices de satisfacción desacoplados de ingresos y asociados más bien al tipo de especialidad.

Quizás el hallazgo principal sea que la evidencia muestra variabilidad asociada a la naturaleza de las especialidades.

Así, el viaje nos deja algunas lecciones y pistas:

1. El hallazgo principal es la necesidad de distinguir, en la variabilidad observada entre especialidades, el aporte endógeno del modelo de formación por competencias de los aspectos exógenos propios de la composición por nivel socioeconómico, de los mercados y de las diferencias entre especialidades, tarea a la que recién nos asomamos.
2. La idea de un modelo institucional de formación por competencias no debe perder de vista el hecho de que la formación está referida a especialidades, las cuales son las unidades más vitales en la dinámica de las instituciones universitarias; pero que además son heterogéneas, tanto en su especialización temática como en sus fines, y su diseño y evaluación debería prestar más atención a esta diversidad.
3. Por otro lado, cada especialidad, para serlo, ha construido un nicho en el que podemos distinguir un componente disciplinar cuyo ritmo es marcado por dinámicas académicas internacionales y un componente laboral predominantemente referido a un mercado local. Con frecuencia formulamos esta distinción como el componente académico y el

13 Esta frase viene del comentario de César Guadalupe al texto extenso.

componente profesional. Convendría incorporar en el SSE información sobre la dimensión internacional disciplinar.

4. Sin embargo, el peso de la dimensión disciplinar internacional será diferente en las disciplinas orientadas al servicio que cuentan con sistemas regulatorios que les otorga monopolio en el ejercicio, que denominamos con otra acepción de lo profesional, que en las disciplinas más orientadas al desarrollo del conocimiento (la enseñanza superior y la investigación) a las que denominamos con otra acepción de lo académico. También estas distinciones generan consecuencias respecto a la inserción laboral y los motivos de satisfacción que el SSE pone en evidencia y requiere una capacidad de análisis más fina.
5. En la experiencia disponible, la información sobre empleabilidad y satisfacción de los egresados, producida por el SSE, es requerida por las especialidades en el proceso de revisión de sus planes de estudio y en los procesos de acreditación. Contar con datos alimenta la discusión sobre fortalezas y debilidades que ayuda a modular las nuevas propuestas curriculares y los planes de mejora. Por otro lado, hemos encontrado útil brindar no solo la información de cada disciplina, sino los datos de la media institucional y de las otras disciplinas de su facultad. La tipología podría sugerir incluir otras disciplinas del clúster al que pertenece, como insumo en esta reflexión.
6. Un elemento que encuentra sitio en este contexto y que debe tener impacto creciente en las evaluaciones de las especialidades es el que resulta del observatorio de educación y empleo “Ponte en carrera”¹⁴. Disponer de indicadores de la variabilidad de las remuneraciones por especialidad y de quienes son las universidades que lideran los ingresos máximos por especialidad retroalimentará las percepciones colectivas y de los propios egresados. Habrá que estar atento a la evolución de las opiniones ahora que los juicios podrían formularse sobre la base de estas evidencias.
7. “Ponte en carrera”, sin embargo, contiene información (no exhaustiva) de un fragmento del mercado laboral: el de las instituciones formales, públicas y privadas que reportan empleo dependiente contratado con todos los derechos. En un país con gran parte del empleo fuera del mercado dependiente y con parte importante del empleo dependiente fuera de las reglas formales de empleo, descuidar estos hechos puede generar graves sesgos en el análisis (Novedades Académicas, 2015a, 2016). Para hacer comparable la

14 El observatorio es una iniciativa conjunta del Ministerio de Trabajo y Promoción del Empleo, el Ministerio de Educación, e IPAE, organización privada empresarial (ver <http://www.ponteencarrera.pe/>).

información del SSE con la del observatorio “Ponte en carrera”, convendrá precisar qué parte del empleo de nuestros egresados podría pasar desapercibido en el observatorio.

8. “Ponte en carrera” exige considerar los otros niveles, las facultades y las universidades. Al respecto, conviene recordar la idea de Burton Clark y Clark Kerr de que la dimensión institucional en la universidad tiene muchas restricciones para funcionar como un sistema orgánico; la dimensión “empresa” ofrece una articulación débil explicada por el peso de las especialidades. Esto tiene una consecuencia estadística ingenua tal que los valores institucionales serán promedios explicados por la composición de las especialidades en cada aspecto: demografía de alumnos y profesores, de pregrado y posgrado; pero también infraestructura de investigación, o producción académica de libros o de artículos en revistas indexadas.
9. Otra dimensión relacionada con la esfera institucional es que conviene también incluir en el análisis de los logros de los egresados PUCP los aportes respecto a ingresos de determinantes distintos a la calidad/prestigio institucional identificados en los estudios internacionales. Un estudio británico reciente indica que los determinantes serían, según orden de importancia: especialidad, institución académica, nivel socioeconómico familiar y género. En general, los ingresos promedio de quienes estudian Medicina serán mayores a los que estudien Arte sin importar los otros determinantes; pero existe un premio para quienes proceden de familias de nivel socioeconómico más alto sobre el bono de la institución y, en la base, una discriminación de género (Britton, Dearden, Shephard & Vignoles, 2016). La gradiente entre especialidades está bien documentada en estudios anglosajones y españoles sobre ingresos (Carnevale, Cheah, & Hanson, 2015; Novedades Académicas, 2015b).
10. Dicho esto, sin embargo, convendrá tener en cuenta también la creciente diversidad de misiones institucionales en el diseño y evaluación de los SSE. Una universidad de investigación (de posgrado con la mayor parte del financiamiento vinculada a proyectos en los que los profesores investigan con sus posgraduandos) debería generar no solo un perfil diferente dentro de un SSE compartido con otros tipos de universidad, sino un SSE distinto que el de una universidad de formación (de pregrado y la mayor parte del financiamiento vinculada a enseñanza). La diversidad incluye diferencias de escala; de modalidad presencial, mixta o a distancia; de mercado laboral de referencia (centro/periferia); de énfasis académico o profesional; y, por supuesto, de portafolio disciplinario (un área de conocimiento, varias, todas, etc.). Para misiones diferentes, criterios diferentes. Por ejemplo, en universidades con vocación de formación de actores en el espacio público, habría que incorporar en el SSE este tema; mientras que en

universidades con vocación de innovación habría que incluir el seguimiento de *start ups* y patentes, por ejemplo.

Debemos concluir agradeciendo la invitación a responder a las preguntas de Dieudonné Leclercq en este III Encuentro Internacional Universitario: “La Evaluación de Competencias en la Educación Superior” y por habernos empujado a precisar esta reflexión.

Bibliografía

- Apoyo (2003). *I estudio institucional sobre egresados*. Lima: PUCP.
- Becher, T. (2002). *Tribus y territorios académicos. La indagación intelectual y la cultura de las disciplinas*. Barcelona: Gedisa.
- Borden, V. M. (2003). Las encuestas a egresados universitarios como medio para la mejora de las universidades: lecciones desde Estados Unidos. En Vidal García, J. (ed.), *Métodos de análisis de la inserción laboral de los universitarios* (pp. 81-91). Salamanca: Universidad de León.
- Britton, J., Dearden, L., Shephard, N. & Vignoles, A. (2016). How English domiciled graduate earnings vary with gender, institution attended, subject and socio-economic background. IFS Working Paper W16/o6, Institute for Fiscal Studies. Disponible en: <<http://www.nuffieldfoundation.org/sites/default/files/files/wp16o6.pdf>>.
- Cabrera, A., Weerts, D. & Zulick, B. (2003). Encuestas a egresados: tres fundamentos conceptuales en el seguimiento de egresados universitarios. En Javier Vidal García (ed.), *Métodos de análisis de la inserción laboral de los universitarios* (pp. 55-80). Salamanca: Universidad de León.
- Carnevale, A., Cheah, B. & Hanson, A. (2015). The economic value of college majors. Washington, DC: Georgetown University, Center on Education and the Workforce. Disponible en: <<https://cew.georgetown.edu/wp-content/uploads/The-Economic-Value-of-College-Majors-Full-Report-web-FINAL.pdf>>.
- Centro Interuniversitario de Desarrollo (CINDA) & Grupo Operativo de Universidades Chilenas (GOUCh) (eds.) (2012). *Seguimiento de egresados e inserción laboral: experiencias universitarias* (primera edición). Santiago de Chile:

- CINDA. Disponible en: <[http://www.cinda.cl/download/libros/01%20Cinda%20Seguimiento%20indd%20\(2\).pdf](http://www.cinda.cl/download/libros/01%20Cinda%20Seguimiento%20indd%20(2).pdf)>.
- Clark, B. (1986). *The Higher Education System: Academic Organization in Cross-national Perspective*. Berkeley: University of California Press.
- Gonzales de Olarte, E. (2011). Nuestro modelo educativo. Discurso de orden (14 de marzo). Disponible en: <<http://textos.pucp.edu.pe/pdf/450.pdf>>.
- Gonzales de Olarte, E. (2013). Modelo académico PUCP. Presentado en curso IGLU (noviembre). Lima, Perú.
- Gumport, P. (2007). Sociology of Higher Education: An Evolving Field. En Patricia Gumport (ed.), *Sociology of Higher Education: Contributions and Their Contexts* (pp. 17-50). Baltimore: Johns Hopkins University Press.
- Herrera, R. & otros (2012). Aspectos referenciales sobre la inserción laboral y seguimiento de egresados. En Centro Interuniversitario de Desarrollo (CINDA) & Grupo Operativo de Universidades Chilenas (GOUCH) (eds.), *Seguimiento de egresados e inserción laboral: experiencias universitarias* (primera edición, pp. 19-56). Santiago de Chile: CINDA. Disponible en: <[http://www.cinda.cl/download/libros/01%20Cinda%20Seguimiento%20indd%20\(2\).pdf](http://www.cinda.cl/download/libros/01%20Cinda%20Seguimiento%20indd%20(2).pdf)>.
- Kerr, C. (1963). *The uses of the university* (presentado en The Godkin Lectures). Cambridge, Mass.: Harvard University Press. Disponible en: <<http://raley.english.ucsb.edu/wp-content/uploads/Reading/Kerr.pdf>>.
- King, A. & Brownell, J. A. (1966). *The curriculum and the disciplines of knowledge: A theory of curriculum practice*. New York: John Wiley.
- Kuhn, T. (1990). *La estructura de las revoluciones científicas*. México: FCE.
- Leclercq, D. & Cabrera, A. (2014). *Ideas e innovaciones : dispositivos de evaluación de los aprendizajes en la educación* (primera edición). Santiago de Chile: Universitaria.
- Lemaitre, M. J. & Zenteno, M. E. (eds.) (2012). *Aseguramiento de la calidad en Iberoamérica. Educación superior. Informe 2012* (primera edición). Santiago de Chile: CINDA. Disponible en: <<http://www.cinda.cl/wp-content/uploads/2014/02/Aseguramiento-de-la-calidad-en-Iberoam%C3%A9rica-2012.pdf>>.
- Lemarchand, G. (2010). Ciencia para la paz y el beneficio de la humanidad: El concepto del Juramento Hipocrático para Científicos. En Lemarchand, G.

- (ed.), *Ciencia para la paz y el desarrollo: el caso del Juramento Hipocrático para científicos* (primera edición, pp. 39-111). Montevideo: Oficina Regional de Ciencia para América Latina y el Caribe. Disponible en: <unesdoc.unesco.org/images/0018/001884/188400s.pdf>.
- Marginson, S. (2014). Clark Kerr and the Californian Model of Higher Education (presentado en 2014 Clark Kerr Lecture). Berkeley, California: Center for Studies in Higher Education (CSHE). Disponible en: <<http://www.cshe.berkeley.edu/publications/clark-kerr-and-californian-model-higher-education>>.
- Mindreau, E. (2014). *La universidad desde adentro: características, comportamiento y gestión* (primera edición). Lima: Universidad del Pacífico. Disponible en: <<http://www.up.edu.pe/investigacion-centros/fondo-editorial/catalogo/la-universidad-desde-adentro>>.
- Natale, L. (2013). Integración de enfoques en un programa institucional para el desarrollo de la escritura académica y profesional. *Revista Mexicana de Investigación Educativa*, 18(58), 685-707.
- Navarro, F. (2012). ¿Qué son los géneros profesionales? Apuntes teórico-metodológicos para el estudio del discurso profesional. En Actas del IV Congreso Internacional de Letras "Transformaciones Culturales. Debates de la teoría, la crítica y la lingüística en el Bicentenario" (volumen 22, pp. 1294-1303). Buenos Aires: Universidad de Buenos Aires. Disponible en: <<http://2010.cil.filo.uba.ar/sites/2010.cil.filo.uba.ar/files/189.Navarro.pdf>>.
- Novedades Académicas (2015a). "Ponte en Carrera": información resaltante y algunas precisiones a tener en cuenta. Lima: PUCP. Disponible en: <<http://vicerrectorado.pucp.edu.pe/academico/noticia/ponte-en-carrera-informacion-resaltante-y-algunas-precisiones-a-tener-en-cuenta/>>.
- Novedades Académicas (2015b). Ranking: empleabilidad y competencias profesionales que dan las universidades españolas. Disponible en: <<http://vicerrectorado.pucp.edu.pe/academico/noticia/ranking-empleabilidad-y-competencias-profesionales-que-dan-las-universidades-espanolas/>>.
- Novedades Académicas (2016). Ponte en carrera. Problemas de cocina. Disponible en: <<http://vicerrectorado.pucp.edu.pe/academico/noticia/ponte-en-carrera-problemas-de-cocina/>>.

- Parodi, G. (2009). El Corpus Académico y Profesional del Español PUCV-2006: semejanzas y diferencias entre los géneros académicos y profesionales. *Estudios Filológicos*, 44, 123-147.
- PUCP (2006). *II Estudio institucional sobre egresados PUCP. Primer reporte: Resultados a nivel institucional y de facultades 1999-2003* (primera edición). Lima: PUCP DAPE.
- PUCP (2010). III Estudio institucional sobre egresados PUCP. Reporte por especialidades. En *III Estudio institucional sobre egresados PUCP: egresados 2006-1, 2006-2 y 2007-0* (p. 36). Lima: PUCP DAPE.
- PUCP (2015). *Análisis comparativo entre egresados: promociones 2006 y 2010 (entre 1 y 2 años de egreso)* (primera edición). Lima: PUCP DAA. Disponible en: <http://vicerrectorado.pucp.edu.pe/academico/wp-content/uploads/2015/03/comparativo_egresados_2006_2010.pdf>.
- Snow, C. P. (2000). *Las dos culturas* (H. Pons, trad.; primera edición). Argentina: Nueva Visión.
- Susskind, R. & Susskind, D. (2016). *The Future of the Professions. How Technology Will Transform the Work of Human Experts* (primera edición). Oxford, United Kingdom: Oxford University Press.
- Trow, M. (2006). Reflections on the Transition from Elite to Mass to Universal Access: Forms and Phases of Higher Education in Modern Societies since WWII. En James Forest & Philip Altbach (eds.), *International Handbook of Higher Education* (pp. 243-280). Netherlands: Springer. Disponible en: <http://link.springer.com/chapter/10.1007/978-1-4020-4012-2_13>.
- Vega, J. F. (2003). *I Estudio institucional sobre egresados PUCP. Análisis detallado*. Lima: PUCP.
- Weber, M. (2012). *La ética protestante y el espíritu del capitalismo* (J. Abellán, editor) (primera edición). México: Alianza.
- Wiener, L. & Vega, J. F. (2016). *Tipología de especialidades PUCP en base a la serie 2009 2013 del SSE*. Lima, Perú: PUCP.

El reto de la evaluación por competencias en las asignaturas de Humanidades

 *Gemma Puigvert Planagumà*
Josep Antoni Iglesias-Fonseca¹

El trabajo que presentamos pretende ofrecer un marco de reflexión ante el proceso de evaluación de los aprendizajes por competencias en el ámbito de las humanidades que se impuso a raíz del proceso de adaptación de las titulaciones al Espacio Europeo de Educación Superior (EEES) llevado a cabo en la Universitat Autònoma de Barcelona (UAB) entre los años 2004 y 2006.

En esa línea, en primer lugar, se describirá la evolución que siguió este proceso en la UAB, así como la adopción de un nuevo modelo educativo por parte de la titulación de Humanidades, cuyo eje central fue un cambio de paradigma docente que consistió en pasar de una docencia centrada en la enseñanza a una docencia centrada en el aprendizaje. Se verá también la importancia de reflexionar sobre el perfil de competencias de un estudiante de Humanidades y la necesidad de diseñar actividades de evaluación de aprendizajes acorde con las competencias fijadas a priori.

Finalmente, se propondrán ejemplos de actividades con la explicitación de las competencias implicadas en cada una de ellas, los conocimientos y habilidades implícitos, los contextos en que se trabajarán, así como los resultados de aprendizaje esperables y observables.

¹ Gemma Puigvert Planagumà es profesora titular de Filología Latina y Josep Antoni Iglesias-Fonseca, de Ciencias y Técnicas Historiográficas (Paleografía, Codicología y Diplomática) del Departament de Ciències de l'Antiguitat i de l'Edat Mitjana de la UAB.

LA EVALUACIÓN DE LOS APRENDIZAJES POR COMPETENCIAS EN EL CONTEXTO DEL ESPACIO EUROPEO DE EDUCACIÓN SUPERIOR

La Unión Europea (UE) tuvo en sus inicios una orientación fundamentalmente económica y centrada en conseguir la apertura de fronteras y la creación de un espacio económico común. Sin embargo, poco a poco, propició la convergencia en otros ámbitos como los jurídicos, sociales y educativos.

El esfuerzo en educación universitaria se dirigió a promover un EEES que permitiera el reconocimiento de titulaciones y asegurara una formación óptima de los estudiantes y su integración en un mercado laboral unificado y sin fronteras.

Los primeros antecedentes se remontan a la Carta Magna de las universidades europeas, firmada por los rectores asistentes en Bolonia (Italia), el 18 de setiembre de 1988, con motivo del IX centenario de su universidad. En ella se alienta la movilidad de profesores y estudiantes y se considera que una política general de equivalencia en materia de estatus, títulos, exámenes (aun manteniendo los diplomas nacionales) y de concesión de becas constituirá un instrumento esencial para garantizar el ejercicio de la misión actual de la universidad.

Posteriormente, han sido varias las declaraciones que han ahondado en este sentido (cuya síntesis puede verse en García & Morant, 2005)².

En ese contexto, se imponía dos conceptos clave: el de crédito europeo y el de competencia.

El ECTS (*European Credit Transfer System*) suponía una nueva forma de conceptualizar la docencia y el aprendizaje en la universidad. A partir del Real Decreto de setiembre de 2003³, el crédito europeo fue la nueva medida académica que representaba el total del trabajo realizado por el estudiante para cumplir los objetivos del programa de estudios marcado por la titulación que cursaba. En esta unidad, quedaban integradas las enseñanzas teóricas y prácticas, así como otras actividades académicas dirigidas, incluidas las horas de estudio y trabajo autónomo que el estudiante debía realizar para conseguir los objetivos formativos propios de cada una de las materias del plan de estudios correspondiente. El número medio de horas por crédito quedó fijado en 25 y el número máximo en 30.

2 Declaración de la Sorbona (mayo, 1998), Declaración de Bolonia (junio, 1999), I Convención de Instituciones de Educación Superior (Salamanca, marzo, 2001), Conferencia de Ministros (Praga, mayo, 2001), II Convención de Instituciones de Educación Superior (Graz, mayo, 2003), Conferencia de Ministros (Berlín, setiembre, 2003), Conferencia de Bergen (Noruega, mayo, 2005), Conferencia de Londres (mayo, 2007).

3 Real Decreto 1125/2003, del 5 de setiembre (BOE núm. 224, 18 de setiembre de 2003: artículo 5).

La adopción de los créditos ECTS en todas las titulaciones de la UAB y su asignación estaba directamente relacionada con uno de los puntos básicos de la convergencia europea: la estructura de las titulaciones teniendo en cuenta los perfiles profesionales que determinan sus contenidos específicos.

Rué y Martínez (2005) planteaban la necesidad de realizar un acercamiento progresivo a tres niveles (nivel 1: especificar el perfil de la formación; nivel 2: definir la formación en ECTS; nivel 3: configurar las materias y asignaturas en ECTS).

La determinación del perfil profesional de una titulación, como descripción detallada y pormenorizada que delimita las características propias de un grupo profesional, define su identidad y nos orienta sobre la naturaleza de los aprendizajes que deben ser priorizados para que el profesional sea competente en su actividad. El perfil formativo enlaza así con las competencias que la formación debe desarrollar (Sánchez Delgado & Gairín Sallán, 2008).

Detrás del término “competencia” hay toda una teoría psicopedagógica que enfatiza en el desarrollo de capacidades personales por encima de la simple asimilación de contenidos. Vincular las competencias con las capacidades supone destacar el carácter aplicativo de los aprendizajes, a la vez que reconocer una perspectiva global en la que intervienen tanto elementos cognitivos, como actitudes y habilidades. El término “competencia” debe, pues, integrar tres dimensiones: ser, saber hacer y saber estar; tres dimensiones o componentes que, para garantizar el éxito, se tienen que movilizar y articular conjuntamente en cualquier ejecución académica o profesional (Allen, 2000; Lévy-Leboyer, 2003; Carreras & Perrenoud, 2006; Yáñez & Villardón, 2006; Puigvert & otros, 2009).

Cada titulación fijó dos amplios grupos de competencias: las específicas (propias de la titulación, orientadas a la consecución de un perfil específico del graduado o graduada) y las genéricas o transversales (comunes a la mayoría de las titulaciones, aunque con una incidencia diferente: nos referimos a competencias personales como la gestión del tiempo y la responsabilidad del propio aprendizaje; competencias interpersonales, cómo comunicarse, trabajar en equipo, liderar o negociar; competencias relacionadas con la gestión de la información, los idiomas, el dominio de las TIC; etc.).

El proyecto Tuning (2003)⁴ ya señalaba tres grandes factores que explicaban el interés de desarrollar las competencias en programas educativos:

4 El proyecto Tuning es el programa de más impacto que lanzaron las universidades europeas para responder a la Declaración de Bolonia y al Comunicado de Praga. Consistió en elaborar propuestas de calidad que pudieran convertirse en una referencia para el conjunto de titulaciones universitarias. El informe final del proyecto se puede consultar en: *Tuning Educational Structures in Europe* (Universidad de Deusto, 2003).

1. La necesidad de mejorar la “ocupabilidad” de los graduados en la nueva sociedad del conocimiento.
2. La creación del EEES —necesidad de establecer referentes comunes para las titulaciones (descriptores de Dublín para *bachelor* y máster).
3. Un nuevo panorama educativo centrado en el aprendizaje de los estudiantes y que pone mayor énfasis en los resultados y objetivos de la enseñanza.

LA EVALUACIÓN DE LOS APRENDIZAJES POR COMPETENCIAS EN LA UAB

En la UAB, fue la Unitat d’Innovació Docent en Educació Superior (IDES) quien lideró el proceso de adaptación de las titulaciones al nuevo marco europeo (Martínez & Añaños, 2008). Se pidió a las titulaciones que iniciaran un proceso de reflexión sobre cuál era el perfil de estudiante que se quería conseguir después de su período de formación en la UAB, pensando en que cuando finalizase sus estudios se tendría que insertar en un mercado laboral y en una sociedad con unas necesidades cambiantes. Se nos pidió también que pensáramos en el perfil de competencias de los estudiantes, en cómo se iban a trabajar dentro de la estructura de estudios de la titulación y en cómo se evaluarían.

Cada titulación creó una comisión de titulación presidida por su coordinador e integrada por una selección de profesores y estudiantes que se encargó de llevar a la práctica todos los elementos que debían ser tenidos en cuenta en el proceso de adaptación y en el modelo establecido. También se pidió la colaboración de agentes externos (profesionales, empleadores, centros de prácticas, exalumnos, etc.).

Estas reflexiones condujeron al diseño de los planes de estudio actuales donde cada asignatura tiene una guía docente que incluye los apartados siguientes: 1. Datos de la asignatura (nombre de la asignatura, código, número de créditos ECTS, curso y período en el que se imparte, horario, lugar donde se imparte, lengua en que se imparte —castellano o catalán—); 2. Equipo docente (nombre del profesor, departamento, número de despacho, teléfono, correo electrónico, horario de atención docente); 3. Prerrequisitos para cursar la asignatura; 4. Contextualización y objetivos formativos; 5. Competencias y resultados de aprendizaje; 6. Contenidos; 7. Metodología docente y actividades formativas (dirigidas, autónomas y evaluación); 8. Evaluación; 9. Bibliografía y enlaces web; 10. Programación de la asignatura, donde se explicitan las actividades formativas y las entregas correspondientes.

LA EVALUACIÓN DE LOS APRENDIZAJES POR COMPETENCIAS EN LA TITULACIÓN DE HUMANIDADES DE LA UAB

La titulación de Humanidades, una de las 13 titulaciones que se imparten en la Facultat de Filosofia i Lletres de la UAB⁵, inició su andadura en el año 1992⁶, cuando se vio la necesidad de formar a estudiantes versátiles, capaces de adaptarse a retos intelectuales y a propuestas profesionales cambiantes, diversas o, incluso, inéditas. Los estudios de humanidades se fundamentan sobre las mismas disciplinas que integraban los *studia humanitatis* del siglo XV (ética, poesía, historia, retórica y gramática), aunque bajo los nombres de filología, historia, filosofía y ciencias sociales. La transversalidad de conocimientos que constituye la singularidad principal de esta titulación hace que los estudiantes que la cursan estén a la vanguardia de un contexto cada vez más global⁷.

La titulación de Humanidades fue una de las 11 titulaciones piloto de la UAB en el plan de adaptación al nuevo marco impuesto por el EEES (curso 2004-05) y su diseño curricular se fijó, inicialmente, en tres años (180 créditos), tal como marcaban las directrices europeas, si bien más tarde las directrices del Ministerio fijaron todos los grados de Humanidades de España en 4 años⁸.

En el año 2007, la Agència per a la Qualitat del Sistema Universitari de Catalunya (AQU) nos concedió un proyecto para que pudiéramos liderar este proceso⁹ y diseñar una guía exhaustiva donde quedaran reflejadas las competencias específicas fundamentales que se asocian a un

5 Las demás titulaciones impartidas son: Arqueologia, Estudis Anglesos, Estudis Clàssics, Estudis Francesos, Filosofia, Història, Història de l'Art, Llengua i Literatura Catalanes, Llengua i Literatura Espanyoles, Musicologia, Antropologia Social i Cultural, Geografia i Ordenació del Territori.

6 Real Decreto 913/1992, del 17 de julio (BOE núm. 206, del 27 de agosto).

7 Para entender el espíritu interdisciplinar con el que fue concebida la titulación de Humanidades en la UAB, véase De-la-Mota & Puigvert, 2009.

8 Real Decreto 1393/2007, del 29 de octubre (BOE núm. 260, del 30 de octubre de 2007). En el año 2015, el gobierno español introdujo una modificación sustantiva: el Real Decreto 43/2015, del 2 de febrero de 2015, modificaba el RD 1393/2007. Este cambio, según se dice en el texto del RD, promovía la internacionalización de los estudios homogeneizando la duración de los grados impartidos en universidades españolas con los que se ofrecen en la gran mayoría de universidades europeas. La principal novedad era la que contemplaba la posibilidad de diseñar títulos de 180 ECTS (tres años). En estos momentos se está en proceso de revisión de los títulos de 240 ECTS y el debate se centra en la posible coexistencia de grados de 240 y 180 ECTS a los que seguiría una estructura de máster de 60 ECTS y 120 ECTS, respectivamente.

9 Resultado de este proyecto fue la publicación de la *Guia per a l'avaluació de les competències en l'àmbit de les Humanitats* (Barcelona, 2009). Tanto la versión en catalán como la versión en castellano se pueden consultar en: <http://www.aqu.cat/doc/doc_95455311_1.pdf> (consultado el 15 de octubre de 2016).

estudiante de Humanidades (no solo de la titulación que lleva ese nombre, sino de las demás titulaciones que podemos englobar bajo este ámbito: filología, historia, filosofía, arte). Se trataba de la primera guía, y hasta ahora la única, que aborda la evaluación por competencias en el ámbito de las humanidades.

Para la elaboración de esta guía, contamos con la experiencia del profesorado que integró el equipo de investigación, ya que todos ellos eran profesores de la titulación de Humanidades y cubrían el espectro de las principales titulaciones implicadas en este grado: Filología clásica, Filología catalana, Filología hispánica, Filosofía, Historia y Arte. También se incluyó el ámbito de la Traducción, aunque esta disciplina se imparte en la Facultad de Traducción e Interpretación. Al equipo se incorporaron profesionales de la pedagogía con experiencia dilatada en evaluación por competencias.

La implantación de un nuevo modelo basado en competencias hacía variar la naturaleza, la cantidad y la calidad de las sesiones de formación en las aulas. Las clases magistrales (entendidas como mecanismo de transmisión de conocimientos, esencialmente unidireccional, del profesor al estudiante), que habían sido la tipología dominante en el ámbito de las Humanidades, debían complementarse y enriquecerse con otras modalidades de enseñanza y aprendizaje. Esto conllevaba un cambio sustancial en el sistema de aprendizaje, ya que el estudiante debía implicarse más activamente en la comprensión y la construcción de su propia formación; debía estar dispuesto, asimismo, a seguir una evaluación continuada, que debía permitirle ajustes continuos de su propio proceso. Para ello, el perfil del alumno es, *a priori*, el de un estudiante motivado, a tiempo completo y responsable activo de su aprendizaje.

En este proceso es clave el diseño de actividades de aprendizaje encaminadas a la consecución de una determinada competencia y la elección de los mejores instrumentos para la evaluación.

La pirámide de Miller (1990) distingue dos grandes tipos de pruebas, que podríamos clasificar en evaluación tradicional y evaluación de ejecuciones:

1. La evaluación tradicional engloba las típicas “pruebas de papel y lápiz”, en las que se pone más énfasis en los objetivos de conocimientos y de saber. En la evaluación tradicional hay pruebas que enfatizan habilidades de bajo orden (recuerdo, comprensión), mientras que hay otras que enfatizan el pensamiento de alto orden (aplicación, síntesis, evaluación).
2. La evaluación de ejecuciones es muy variada y permite abrazar un rango mucho más amplio de competencias, sea de habilidades disciplinarias (así, por ejemplo, este tipo

de pruebas evalúa si el estudiante sabe programar una unidad didáctica) o de competencias transversales (comunicación oral, pensamiento crítico, etc.)¹⁰.

Sobre la base de esta pirámide, se han diseñado las pruebas de evaluación que pueden consultarse en nuestra guía y que abrazan la totalidad de disciplinas que se incluyen dentro del Grado en Humanidades de la UAB.

LAS COMPETENCIAS GENÉRICAS O TRANSVERSALES EN EL ÁMBITO DE LAS HUMANIDADES

Para las competencias genéricas o transversales que pueden identificar el conjunto de titulaciones que integran el área de Humanidades, seguimos la clasificación establecida por Villa y Poblete (2007).

1. Capacidad de razonamiento crítico.
2. Capacidad de análisis y síntesis de la información.
3. Capacidad de reflexión y de argumentación instrumental.
4. Capacidad de resolución de problemas.
5. Capacidad de toma de decisiones.
6. Capacidad de expresión oral.
7. Capacidad de expresión escrita.
8. Capacidad de organización y planificación del trabajo individual.
9. Capacidad de dominio de la informática y capacidad de gestión de la información.
10. Capacidad de trabajo en equipo.
11. Compromiso ético.

A la hora de evaluar las competencias genéricas, se puede optar por estrategias diversas (algunas de las cuales se recogen en el anexo en relación con las actividades propuestas). Entre otras, tests objetivos, preguntas breves, pruebas orales, exámenes o prácticas de laboratorio con material de consulta, trabajos de investigación, informes, artículos, ensayos, *dossiers* o carpetas de aprendizaje, trabajos en cooperación, etc.¹¹.

10 Para los dos tipos de evaluación, véase Mateo Andrés & Martínez Olmo, 2005, pp. 16-17. Para ver el potencial de cada prueba con relación a la evaluación de competencias y consideraciones diversas sobre la fiabilidad y validez de cada una, recomendamos ver Puigvert 2009, pp. 21-25.

11 Para las características de cada prueba y para conocer su fiabilidad y validez, véase Puigvert, 2009, pp. 20-25.

LAS COMPETENCIAS ESPECÍFICAS PROPIAS DEL ÁREA DE LAS HUMANIDADES

Aquí detallaremos solo aquellas que cubren las actividades de aprendizaje y evaluación presentadas en el anexo siguiente¹²:

1. Analizar y aplicar los conceptos gramaticales, fonéticos, morfológicos, sintácticos y semánticos de las lenguas clásicas.
2. Traducir e interpretar textos latinos.
3. Saber valorar la información específica sobre el mundo clásico y generar opiniones propias justificadas.
4. Transcribir, anotar e interpretar un documento histórico original manuscrito.

ANEXO

Ejemplos de actividades¹³

1. Traducción e interpretación de un texto latino¹⁴.
2. Cinefórum. Comentario de una película/documental inspirada en el mundo clásico o medieval.
3. Transcripción, anotación e interpretación de un documento original manuscrito.
4. Comentario de un texto histórico.

I. Traducción e interpretación de un texto latino

ASIGNATURA: Poesía lírica latina (asignatura obligatoria).

NIVEL ACADÉMICO: tercer curso (grado de Estudios Clásicos).

DESCRIPCIÓN BREVE DE LA ACTIVIDAD

La actividad de aprendizaje y evaluación consiste en la traducción e interpretación de un texto latino. En este caso, proponemos un poema de uno de los autores líricos que más ha influido en la tradición literaria posterior: Horacio.

¹² Para el resto de competencias específicas del área de Humanidades, véase Puigvert, 2009, pp. 30-33.

¹³ Cada una de las actividades presentadas seguirá una misma estructura, que es la que se propone en Puigvert, 2009, pp. 35-38. Es la que figura también en Puigvert, 2010.

¹⁴ Para la traducción de un texto latino de dificultad mediana (primer curso de carrera), véase Puigvert, 2009, pp. 103-109.

La actividad propuesta forma parte de Poesía lírica latina, asignatura obligatoria de tercer curso del grado de Estudios Clásicos e incluida dentro de la materia Lengua y Literatura Clásicas.

COMPETENCIAS QUE SE PUEDEN DESARROLLAR A PARTIR DE LA ACTIVIDAD¹⁵

Mediante esta actividad, buscamos evaluar algunas de las competencias específicas de la titulación de Estudios Clásicos, en concreto las que hacen referencia a:

1. Analizar y aplicar los conocimientos gramaticales adquiridos al análisis y comprensión de los textos latinos.
2. Interpretar textos latinos, escritos en verso, aplicando el método filológico¹⁶.

OBJETIVOS¹⁷

Mediante esta actividad evaluativa, se busca que el estudiante:

1. Aplique, a partir de lo que se ha trabajado en clase, las técnicas que permiten la comprensión rápida del texto latino.
2. Traduzca de 15 a 20 versos en 1 h 30 min.
3. Haga un comentario gramatical y métrico de un texto de poesía lírica a partir de cuestiones propuestas por el profesor¹⁸.

RESULTADOS DE APRENDIZAJE¹⁹

El estudiante alcanza las capacidades siguientes:

1. Aplica las técnicas que permiten la comprensión rápida de un texto latino en poesía.
2. Traduce y realiza un comentario gramatical y métrico de un texto de poesía lírica.

15 Implican el uso integrado de conocimientos, habilidades y actitudes en la acción.

16 Si se propusiera un texto trabajado previamente en clase se pediría al estudiante que ofreciera, además de la traducción del texto, un comentario de *realia*. En este tipo de pruebas, además de la pericia en la traducción, se examinan otras competencias, entre las que destaca la adquisición de vocabulario.

17 Están escritos desde el punto de vista del profesor. Son afirmaciones relativas a la docencia.

18 En esta asignatura, como en todas las de traducción, es fundamental el trabajo diario que se hace en casa: *nulla dies sine linea*. La traducción y comentario (trabajo autónomo) es revisado y examinado en clase, lo cual nos asegura que se cumplen dos de las competencias transversales incluidas en esta asignatura: el estudiante es capaz de organizar y planificar su trabajo; y el estudiante es capaz de desarrollar habilidades de aprendizaje autónomo.

19 Son afirmaciones sobre lo que se espera que un estudiante pueda conocer, comprender y ser capaz de demostrar después de haber completado un proceso de aprendizaje. Se centran en aquello que puede demostrar el estudiante al finalizar la actividad de aprendizaje.

DESCRIPCIÓN DE LA ACTIVIDAD

La actividad evaluativa que se propone consiste en: traducir al catalán o español los versos que se proponen; realizar el comentario morfosintáctico de algunos versos destacados en negrita; y realizar el comentario métrico de un verso.

Los versos propuestos no han sido trabajados previamente, lo cual supone que, si el estudiante no se ha ejercitado en los trabajos diarios de traducción, es difícil que supere la prueba, teniendo en cuenta la dificultad que presenta el autor, superior a la que presenta el poeta que han visto en el curso precedente: Virgilio, de cuya *Eneida* conocen bien la historia. El contexto de Horacio cambia en cada poema, lo cual aumenta la complejidad.

La evaluación de una traducción siempre es difícil (no tanto puntuar los comentarios gramaticales y los métricos). Es importante disponer de una pauta de corrección y de puntuación²⁰ que facilite, en la medida de lo posible, la objetividad. Así pues, partimos de una traducción ideal y la puntuamos teniendo en cuenta los diferentes grados de dificultad que presentan algunas estructuras. Consideramos importante que el corrector puntúe los aciertos parciales, ya que también son orientadores del nivel de conocimiento alcanzado por el estudiante.

CRITERIOS DE EVALUACIÓN

A la hora de evaluar esta actividad se tienen en cuenta los ítems siguientes:

1. La traducción correcta de los versos con adecuación de la lengua catalana o española al registro lingüístico original²¹.
2. Los aciertos parciales en la traducción.
3. La coherencia global en la traducción.
4. La precisión en el comentario de las estructuras destacadas en negrita.
5. La corrección en el comentario métrico-estilístico.

20 De vez en cuando se facilita a los estudiantes una prueba modelo con la pauta de corrección incluida para que ellos mismos se la puntúen y sean conscientes del avance en sus aprendizajes.

21 La traducción correcta en cualquiera de las dos lenguas nos sirve también para evaluar otra competencia transversal incluida en la guía docente de la asignatura: comunicar por escrito con corrección, precisión y claridad los conocimientos adquiridos.

EJEMPLO PROPUESTO

Prueba de evaluación: lírica latina

I. Lee y traduce los versos siguientes:

Vidi egomet nigra succintam uadere palla
Canidiam, pedibus nudis passoque capillo,
cum Sagana maiore ululatem: pallor utrasque
fecerat horrendas adspectu. Scalpere terram
unguibus et pullam diuellere mordicus agnam
coeperunt; cruor in fossam confusus, ut inde
manis elicerent animas responsa daturas.
Lanea et effigies erat altera cerea: maior
lanea, quae poenis conpesceret inferiorem;
cerea suppliciter stabat, seruilibus ut quae
iam peritura modis. Hecaten uocat altera, saeuam
altera Tisiphonen: serpentes atque uideres
infernus errare canes lunamque rubentem,
ne foret his testis post magna latere sepulcra.
Mentior at siquid, merdis caput inquiner albis
coruorum, atque in me ueniat mictum atque cacatum
Iulius et fragilis Pediatia furque Voranus.

II. Realiza el comentario métrico del primer verso.

PLANTILLA DE CORRECCIÓN

[Vidi egomet nigra **succintam uadere** palla
Canidiam, pedibus nudis passoque capillo,
cum Sagana maiore **ululatem: 1 punto: la traducción tiene que contemplar la construcción de infinitivo y los dos predicativos dependientes de Canidiam**]:
[pallor utrasque
fecerat horrendas adspectu: **0,5 punto**]. [Scalpere terram
unguibus et pullam diuellere mordicus agnam
coeperunt: **1 punto**]; [cruor in fossam confusus: **0,5 punto**], [ut inde

manis elicerent animas responsa daturas: 1 punto: en la traducción deberá haber interpretado correctamente el valor de la conjunción *ut* y la proposición infinitiva].

[Lanea et effigies erat altera cerea: **0,5 punto**]: [maior

lanea, quae poenis conpesceret inferiorem: **1 punto:** la traducción debe reflejar el valor del relativo + subjuntivo];

[cerea suppliciter stabat, seruilibus ut quae

iam peritura modis: **1 punto:** el estudiante debe haber traducido de forma correcta la proposición encabezada por la conj. *ut*]. [Hecaten uocat altera, saeuam

altera Tisiphonen: **0,5 punto**]: [serpentes atque uideres

infernas errare canes lunamque rubentem,

ne foret his testis post magna latere sepulcra: **1 punto:** la traducción debe reflejar la correcta interpretación de las proposiciones de infinitivo y el valor de la conj. *ne*].

[Mentior at siquid, merdis caput inquiner albis

coruorum, atque in me ueniat mictum atque cacatum

Iulius et fragilis Pediatia furque Voranus]. **1 punto para llegar a la excelencia.**

La medida correcta del primer verso vale **1 punto**

Nota: cada error se penaliza descontando **0,5**.

La corrección se presenta a través de un informe que se entrega al estudiante.

II. Cinefórum. Comentario de una película

ASIGNATURA: Cultura clásica I

NIVEL ACADÉMICO: segundo curso (grado en Humanidades)

DESCRIPCIÓN DE LA ACTIVIDAD

Visionado de la película *El nombre de la rosa*.

Se trata de ver (previa lectura de la obra) la película *El nombre de la rosa* como complemento a uno de los apartados de la guía docente de la asignatura Cultura clásica I que se imparte en el segundo curso de la titulación de Humanidades. Se trata del apartado: "La transmisión de la cultura clásica".

En muchos casos, ya han visto la película en sus años de formación preuniversitaria, pero ahora se les pide que vayan un paso más allá y que reflexionen sobre la importancia de la transmisión de los clásicos grecolatinos. Como material de soporte al trabajo, tienen la lectura obligatoria de la obra *Copistas y folólogos* (1986), de L. D. Reynolds y N. G. Wilson (traducida al español por M. Sánchez Mariana; Madrid, Gredos).

A partir de lo explicado y comentado en clase y de la lectura obligatoria, se les pasa un cuestionario que tendrán que responder en grupos de tres personas, una vez vista la película.

COMPETENCIAS QUE SE PUEDEN DESARROLLAR A PARTIR DE LA ACTIVIDAD

1. Ser capaces de valorar y explicar el proceso de transmisión de los textos clásicos hasta el mundo contemporáneo.
2. Desarrollar un pensamiento y un razonamiento críticos y saberlos comunicar de manera efectiva en la lengua propia. Se trata que sepan distinguir cuáles son los elementos fundamentales y cuáles los accesorios o circunstanciales, valorar la relevancia de estos elementos y presentar sucintamente las opiniones que se generen. Se trata también que sepan redactar textos adecuados al registro concreto, gramatical y ortográficamente correctos, utilizando la terminología específica de cada ámbito de estudio.
3. Trabajar en equipo: capacidad de saber trabajar en colaboración con otros (confiar) y ser capaz de asumir las responsabilidades que correspondan, teniendo siempre presente que el interés del equipo queda por delante de los intereses personales de los individuos que lo componen.

OBJETIVOS

Los objetivos que se persiguen con la actividad que se propone son que el estudiante:

1. Conozca y comprenda los procesos que se derivan desde la copia de un texto en la antigüedad hasta que un lector contemporáneo lee el original griego o latino o una traducción.
2. Organice el escrito de manera adecuada, eficaz y lingüísticamente correcta.
3. Valore en un proyecto común las aportaciones y los puntos de vista de los demás.

RESULTADOS DE APRENDIZAJE

Se espera que, con esta actividad, los estudiantes sean capaces de:

1. Identificar y valorar críticamente hechos referentes a la tradición manuscrita y a la recepción de los clásicos grecolatinos con la máxima competencia lingüística.

2. Aplicar las estrategias comunicativas y de comportamiento (asertividad) necesarias para trabajar en grupo.

CRITERIOS DE EVALUACIÓN

La evaluación, a través de un informe escrito que se facilita a cada estudiante, tiene en cuenta no solo la calidad de las respuestas (80%), sino también la propia valoración de cada miembro del grupo respecto a sus capacidades de trabajo en equipo. Cada estudiante debe emitir una valoración sobre el trabajo realizado por él mismo y por los demás componentes del grupo (20%). Es el profesor quien, después de haber leído las argumentaciones de los tres integrantes, decide cuál es la puntuación final.

EJEMPLO PROPUESTO

Cuestionario-comentario sobre la película *El nombre de la rosa* (dirigida por Jean-Jacques Annaud, 1986)

1. Describe el contexto histórico en que se sitúa la película y explica la principal controversia religiosa de la época. (2 puntos: 1 + 1).
2. Explica a qué orden monástico pertenece Guillermo de Baskerville y qué papel había ejercido él en otro tiempo. (1 punto: 0.5 + 0.5).
3. Explica el motivo por el que Guillermo de Baskerville, acompañado por el novicio Adso de Melk, llega a la abadía benedictina. (1 punto).
4. Detalla las cinco muertes acaecidas en el monasterio: los nombres, el papel que desempeñan dentro de la comunidad, cómo mueren. (5 puntos).
5. ¿Quién es el autor de unas letras misteriosas escritas sobre un pergamino? (1 punto).
6. Comenta brevemente la frase: "Con una mano sobre el ídolo pulsa sobre el primero y el séptimo de cuatro". (1 punto).
7. ¿Cuál es el móvil de los crímenes? ¿Quién es el asesino? (2 puntos).
8. ¿Por qué, en un momento determinado de la película, cuando Adso se pierde en el laberinto de la biblioteca, Guillermo de Baskerville le hace saber que está orgulloso de sus conocimientos sobre la cultura clásica? (1 punto).
9. ¿De qué se le acusa a Salvatore? ¿Y a Remigio? ¿Y a la campesina? ¿Cuál es el destino final de los tres? (2 puntos).
10. Cita el nombre de uno de los códices ilustrados que aparecen en la biblioteca y que Guillermo de Baskerville hace observar detenidamente a Adso y di qué representa. (1 punto).

11. ¿Cuál es el soporte de escritura que has identificado en los códices custodiados en la biblioteca? (1 punto).
12. Valoración crítica del trabajo en grupo. (20% del total del ejercicio).

III. Transcripción, anotación e interpretación de un documento original manuscrito²²

ASIGNATURA: Paleografía (asignatura obligatoria).

NIVEL ACADÉMICO: segundo curso (grado en Historia).

DESCRIPCIÓN DE LA ACTIVIDAD

Transcripción, anotación e interpretación de un documento histórico original manuscrito.

La actividad está orientada a estudiantes a partir del segundo curso porque disponen a partir de ese momento de nociones elementales de análisis paleográfico y diplomático que les permiten iniciarse, aunque sea de forma tímida y sobre la base de documentos de escasa complejidad, a una de las tareas primordiales de todo historiador: la investigación a partir de documentación original.

Para ello, el profesor indicará el tipo documental original manuscrito objeto de la práctica (o bien dará a los alumnos las indicaciones para que sean ellos mismos los que localicen el documento que trabajarán); en general, serán documentos de escasa complejidad, muy formularios, de fácil lectura y transcripción, con numerosos ejemplos editados que pueden servir de modelo (testamentos, inventarios, subastas, capítulos matrimoniales, etc.). Antes de la práctica individual, en clase se verá un ejemplo sobre el cual se trabajarán, parcialmente, los objetivos de transcripción, anotación e interpretación propuestos. Una vez el alumno dispone del documento a analizar, se inicia un periodo de trabajo individual por parte del alumno que puede contar con alguna asesoría por parte del profesor.

COMPETENCIAS QUE SE PUEDEN DESARROLLAR A PARTIR DE LA ACTIVIDAD

En esencia, interpretar correctamente un documento histórico original manuscrito.

OBJETIVOS

Entre los objetivos didácticos que se pretende conseguir con esta actividad, se identifica los siguientes:

²² Tanto este anexo como el siguiente siguen, en lo fundamental, las indicaciones recogidas, en los apartados correspondientes, en Puigvert, 2009, pp. 39-44 y 48-50, cuyo autor es J. Antón Pelayo.

1. La lectura de documentos manuscritos originales del pasado.
2. La aplicación de criterios de transcripción adaptados a la actualidad.
3. La anotación de aquellos aspectos necesarios (históricos, institucionales, jurídicos, etc.) para aclarar el contenido de los textos analizados.
4. La identificación de la tipología documental objeto de estudio.
5. La interpretación y contextualización del contenido del documento en el marco histórico e institucional que lo ha generado.

RESULTADOS DE APRENDIZAJE

El estudiante aplicaría sus conocimientos en el ámbito de la paleografía y la diplomática para poder:

1. Identificar la escritura de la época, así como las prácticas diplomáticas que pudiesen reconocerse (formularios, abreviaturas, prácticas de registro, etc.).
2. Puntuar y acentuar el texto de acuerdo con los criterios actuales.
3. Transcribir el texto ofreciendo una presentación moderna del mismo (como si de una edición moderna se tratase).
4. Utilizar los diccionarios etimológicos, de términos históricos e institucionales, de abreviaturas, etc.
5. Hacer uso de la bibliografía específica que describe esta fuente documental para encontrar elementos parangonables con el documento en estudio.
6. Hacer uso de la bibliografía específica que analiza la utilidad de esta fuente documental para identificar aplicaciones comparables.

CRITERIOS DE EVALUACIÓN

La extensión del trabajo vendrá determinada por la transcripción, anotación e interpretación del documento, que se recogerá en una presentación que agrupa tanto el análisis de la tipología documental como la interpretación de su contenido (este comentario debería ocupar entre tres y cinco páginas).

El profesor asignará un porcentaje concreto a cada uno de los valores que se evaluarán a través de esta práctica (transcripción, anotación e interpretación del texto). Además, en el conjunto de la evaluación de la asignatura, esta práctica tendrá un peso específico, a determinar también por el profesor, de acuerdo con la dificultad de otras prácticas programadas (comentario de un texto histórico, debates, reseña(s) de lecturas, transcripción de documentos originales, etc.).

Valoración	Alto (Excelente-MH)	Medio (Notable)	Bajo (Aprobado)	Insuficiente (Suspenso)
Transcripción del texto	El estudiante aplica correctamente las normas de transcripción, puntuación, acentuación y solución de las abreviaturas en el 90% o más del texto.	El estudiante aplica correctamente las normas de transcripción, puntuación, acentuación y solución de las abreviaturas entre el 70% y el 89% del texto.	El estudiante aplica correctamente las normas de transcripción, puntuación, acentuación y solución de las abreviaturas entre el 50% y el 69% del texto.	El estudiante aplica correctamente las normas de transcripción, puntuación, acentuación y solución de las abreviaturas en el 49% o menos del texto.
Anotación del texto	El estudiante ha anotado correctamente un porcentaje muy importante de las palabras o conceptos merecedores de una nota a pie de página.	El estudiante ha anotado correctamente un porcentaje importante de las palabras o conceptos merecedores de una nota a pie de página.	El estudiante ha anotado correctamente un porcentaje suficiente de las palabras o conceptos merecedores de una nota a pie de página (aproximadamente la mitad de las que el profesor considera que tendrían que ser objeto de anotación).	El estudiante ha anotado correctamente un porcentaje insuficiente de las palabras o conceptos merecedores de una nota a pie de página (menos de la mitad de las que el profesor considera que tendrían que ser objeto de anotación).
Interpretación del texto (capacidad de análisis y de síntesis)	<p>Contenido: el estudiante interpreta correctamente el contenido del texto y es capaz de extrapolarlo a una tipología documental. La bibliografía usada es variada y representativa.</p> <p>Presentación: la estructura y la redacción del comentario que acompaña la transcripción son claras, diáfanos, inteligibles y de fácil lectura.</p>	<p>Contenido: el estudiante interpreta bastante bien el contenido del texto, pero no establece ningún elemento relacional. La bibliografía consultada es limitada y poco representativa.</p> <p>Presentación: la estructura y la redacción del comentario que acompaña la transcripción son correctas.</p>	<p>Contenido: el estudiante comete errores de interpretación del texto y no establece ningún elemento relacional. La bibliografía usada es escasa y, en el peor de los casos, nada representativa.</p> <p>Presentación: la estructura es desordenada y la redacción del comentario que acompaña la transcripción es básica.</p>	<p>Contenido: el estudiante interpreta incorrectamente el texto y no establece ningún elemento relacional. La bibliografía usada es nula.</p> <p>Presentación: la estructura es desordenada, incoherente, poco trabajada y la redacción del comentario que acompaña la transcripción es claramente insuficiente para un estudiante universitario.</p>

IV. Comentario de un texto histórico

ASIGNATURA: Codicología (asignatura optativa).

NIVEL ACADÉMICO: tercer/cuarto curso (grado en Historia).

DESCRIPCIÓN DE LA ACTIVIDAD

Comentario de un texto histórico, de naturaleza diversa (en este caso, una carta) para profundizar en el conocimiento histórico a través de un testimonio directo del momento.

La actividad está orientada a estudiantes a partir de segundo curso. Se pretende que el estudiante, a través del comentario de texto y las lecturas necesarias para contextualizarlo, se implique de manera activa y participativa en un momento clave de la historia. Para ello, el profesor, previamente, habrá facilitado algún ejemplo de comentario de texto a los estudiantes y estos tendrán que hacer lo propio en una redacción ordenada, clara y precisa con una extensión que no debería superar las cinco páginas.

COMPETENCIAS QUE SE PUEDEN DESARROLLAR A PARTIR DE LA ACTIVIDAD

1. Capacidad de análisis y de síntesis.
2. Razonamiento crítico.
3. Comunicación escrita y oral (en caso de tener que defenderlo públicamente).

OBJETIVOS

Los objetivos didácticos que se pretenden conseguir son que el estudiante:

1. Comprenda el contexto histórico del texto y su trascendencia posterior.
2. Redacte un comentario de texto (y, eventualmente, sea capaz de defenderlo).

RESULTADOS DE APRENDIZAJE

El estudiante alcanzará las capacidades siguientes:

1. Hacer una lectura atenta y reflexiva del texto, establecer el significado correcto de las palabras y términos que aparecen en el mismo (haciendo uso, para ello, de cuantos diccionarios, enciclopedias y manuales sean necesarios).
2. Comprender y resumir las ideas esenciales, acciones y prácticas contenidas en el texto sin caer en la paráfrasis o reiteración.
3. De manera más específica, sobre la autoría, individual o colectiva, recoger la biografía intelectual del autor o autores para valorar la importancia y la intencionalidad del texto

dentro de su(s) trayectoria(s); para ello, podrá recurrirse también a enciclopedias biográficas relacionadas con la nacionalidad del autor o autores.

4. Introducir un comentario sobre la naturaleza del texto (sobre el género concreto: crónica, epistolografía, historiografía, etc.); si es posible, identificar el formato estilístico y los elementos narrativos que sirven para fundamentar la argumentación.
5. Informarse del contexto y la intencionalidad del texto y valorarlos: fecha y lugar de creación, campo de divulgación, finalidad, resultados concretos, etc. Para ello, será necesario recurrir a material bibliográfico específico.
6. Incluir una valoración global del texto: antecedentes, consecuencias, trascendencia histórica, comparación, si es posible, con textos similares en el tiempo y en el espacio (u otras áreas geográficas próximas al estudiante), etc.
7. Redactar los apartados del comentario de texto de una manera armónica y clara y utilizando la terminología apropiada.

CRITERIOS DE EVALUACIÓN

El profesor asignará un porcentaje concreto a cada uno de los valores que se evaluarán a través de esta práctica (capacidad de análisis y de síntesis, razonamiento crítico y comunicación escrita). Además, en el conjunto de la evaluación de la asignatura, esta práctica tendrá un peso específico, a determinar también por el profesor, de acuerdo con la dificultad de otras prácticas programadas (debates, reseña(s) de lecturas, transcripción de documentos originales, etc.).

Valoración	Alto (Excelente-MH)	Medio (Notable)	Bajo (Aprobado)	Insuficiente (Suspenso)
Capacidad de análisis y de síntesis (lectura comprensiva del texto)	El estudiante categoriza la información correctamente.	El estudiante establece una jerarquía de los contenidos buena, pero no acaba de interpretar correctamente aspectos clave.	El estudiante presenta bastantes dificultades a la hora de establecer la categorización de los contenidos	El estudiante no categoriza la información de manera correcta. Manifiesta dificultades en la lectura comprensiva, sin separar lo anecdótico de lo fundamental.
Razonamiento crítico	El estudiante plantea juicios argumentados de la obra, fruto de un trabajo de investigación, un análisis comparativo y una reflexión personal.	El estudiante plantea bastantes juicios argumentados, aunque no todos los posibles, siendo algunos cuestionables o formulados incorrectamente.	El estudiante propone pocos juicios justificados o simples opiniones no acreditadas.	El estudiante no emite ningún juicio del texto (o no argumenta el o los ofrecidos). No se aprecia trabajo personal de investigación, comparación o reflexión personal.

Valoración	Alto (Excelente-MH)	Medio (Notable)	Bajo (Aprobado)	Insuficiente (Suspenso)
Comunicación escrita	El estudiante demuestra conocer la técnica del comentario de texto, elabora una narración bien estructurada, con frases correctamente construidas y una terminología adecuada. Se deja ver la impronta del redactor y resulta una lectura agradable para el lector.	Siendo un comentario estructurado, no está suficientemente bien expuesto. La terminología usada es imprecisa y dubitativa en ocasiones. Es un comentario impersonal, sin ritmo, poco atractivo para el lector.	Es un comentario básico, sin orden en la exposición y con un nivel gramatical y de léxico elemental, básico. Puede ofrecer alguna dificultad de lectura la comprensión de determinados pasajes del comentario o del texto original.	El estudiante ofrece un comentario con muchas deficiencias ortográficas, gramaticales, sintácticas y de coherencia. Se hace difícil evaluar el contenido último del comentario.

1416, junio-julio: Poggio Bracciolini anuncia a sus amigos florentinos el descubrimiento de códices de clásicos en el monasterio de Sankt Gallen/San Galo (Suiza). Forma parte de un episodio significativo de la recuperación de la antigüedad grecorromana por parte de los humanistas italianos del Renacimiento. Se trata de un testimonio significativo de este período cultural en referencia a los textos clásicos y también, indirectamente, a episodios de la historia de la Iglesia (como el papel de los grandes centros eclesiásticos como depositarios de la herencia clásica o el importantísimo concilio de Constanza, al que asistía Bracciolini, que acabó con el Gran Cisma de Occidente):

La naturaleza, madre de todas las cosas, ha dado al género humano el intelecto y la razón, que constituyen óptimos guías para vivir bien y felizmente; tales son, que no se puede pensar en nada superior. Pero no sé si no son verdaderamente aún más excelentes, entre todos los bienes que nos ha concedido, la capacidad y el orden del decir, sin lo cual la misma razón y el intelecto nada podrían valer. De hecho, solo por medio del discurso conseguimos expresar la virtud del ánimo, distinguiéndonos así de los demás animales. Hemos de estar, pues, sumamente agradecidos, ora a los inventores de las demás artes liberales, ora, y sobre todo, a los que, con sus investigaciones y sus cuidados, nos transmitieron los preceptos del decir y una norma para expresarse con perfección. Lo hicieron, en efecto, de tal modo que, incluso en aquello en lo que los hombres sobrepujan a los demás seres animados, no fuésemos en cambio capaces de ir más allá de nuestros propios límites humanos. Y, habiendo sido muchos los autores latinos, como sabéis, que fueron eminentes en el arte de perfeccionar y de adornar el discurso, entre todos ilustre y excelente fue Marco Fabio Quintiliano, quien

tan claramente y cumplidamente, y con suma diligencia, expuso las cualidades que necesariamente deben adornar a un orador perfecto; de modo que, siguiéndolo, no parece que falte cosa alguna, en mi opinión, para poder alcanzar una doctrina suprema o una singular elocuencia. Y, aunque solo quedase su obra, aun cuando faltase el padre de la elocuencia, Cicerón, conseguiríamos con él una ciencia perfecta en el arte del decir. Pero él, entre nosotros los italianos, había sido tan destrozado, tan mutilado por culpa, creo, de los tiempos, que no se reconocía ya en él ningún aspecto ni aderezo humano [...] Y, en verdad, ¡por Hércules!, si no se le hubiese prestado ayuda, estaba ya para siempre cercano al día de la muerte. Pues no cabe la menor duda de que aquel hombre espléndido, esmerado, elegante, lleno de calidad, lleno de argucia, no hubiese podido soportar más aquella torpe cárcel, ni la desolación del lugar ni la crueldad de los carceleros. Estaba, en efecto, triste y sucio, como solían estarlo los condenados a muerte, con la barba escuálida y los cabellos llenos de polvo, de modo que con su mismo aspecto y con su vestimenta mostraba haber sido destinado a una injusta condena. Parecía tender las manos, implorar la fe de los quirites, que lo protegiesen de un juicio injusto; e, indignamente, aquel que una vez, con su auxilio, con su elocuencia, había salvado a tantos, sufría ahora por no hallar ni un solo defensor que se apiadase de su desventura, que se tomase el trabajo de ocuparse en su salvación, que impidiese que fuese arrastrado a un injusto suplicio. Pero, como dice nuestro Terencio, ¡qué inesperadamente ocurren las cosas que uno no se atrevía a esperar! Una casualidad, afortunada para él y sobre todo para nosotros, quiso que, mientras estaba ocioso en Constanza, me viniese el deseo de visitar el lugar donde había estado recluido.

Pues está junto a esa ciudad el monasterio de San Galo, a una distancia de unas veinte millas. Así que me dirigí allí para distraerme y a la vez para ver los libros, de los que se decía había en gran número. Allí, en medio de una gran masa de libros, que sería demasiado largo enumerar, he hallado a Quintiliano aún a salvo e incólume, aunque todo lleno de moho y polvo. Aquellos libros, efectivamente, estaban en la biblioteca, y no como lo exigía su dignidad, sino como en una tristísima y oscura cárcel, en el fondo de una torre en la cual no se habría encarcelado ni siquiera a los condenados a muerte.

Encontré entre otros los tres primeros libros y la mitad del cuarto de las Argonáuticas de Cayo Valerio Flaco y los comentarios a ocho discursos de Cicerón de Quinto Asconio Pediano, hombre elocuentísimo, obra recordada por el propio Quintiliano. Estos libros los he copiado yo mismo, y a toda prisa, para mandárselos a Leonardo Bruni y a Niccolò Niccoli, los cuales, habiendo sabido que yo había descubierto ese tesoro, insistentemente me solicitaban por carta que les mandase lo más rápidamente posible a Quintiliano (Poggio Bracciolini, en *Rerum Italicarum Scriptores*, XX, pp. 160-161; Sabbadini, *Le scoperte dei codici latini e greci...*, pp. 77-79; trad. española en E. Garin, *El renacimiento italiano*, Barcelona: Ariel, 1986, pp. 41-42).

CONCLUSIONES

Las cuatro actividades de evaluación propuestas son ejemplos que obligan necesariamente al profesor universitario a reflexionar sobre su propia docencia: que enseñamos, cómo enseñamos, qué evaluamos, cómo evaluamos.

El EEES reclamó un nuevo rol del profesor universitario, pues todo lo indicado anteriormente requiere, en última instancia, un cambio de cultura de los docentes y un esfuerzo por inventar nuevos modos de facilitar el aprendizaje. Se propugna un papel más activo del profesorado en el diseño, desarrollo y evaluación de actividades (Imbernon, 1994).

Este proceso no estuvo exento de algunos problemas en su aplicación; por ejemplo, si bien es cierto que algunos profesores ofrecieron cierta resistencia al nuevo modelo, no lo es menos que ante la actitud positiva de la mayor parte del colectivo docente aquellos acabaron por aceptar la propuesta. Un ejemplo claro son las guías docentes que han pasado de ser simples guías orientativas, con una relación escueta del programa y un listado de bibliografía, convertirse en una suerte de contrato de derechos y obligaciones tanto para el docente (viera o no con buenos ojos el nuevo modelo) como para el discente. Las resistencias se han ido venciendo porque incluso los más escépticos han visto cómo este modelo reflejaba mejor el aprendizaje del propio estudiante, cómo se evaluaba de manera más objetiva y cómo mejoraban considerablemente sus resultados. La experiencia de estos años de aplicación permite constatar que ha descendido de manera considerable el número de alumnos no presentados y que los que lo hacen consiguen superar las asignaturas en mayor proporción respecto al sistema anterior. Este dato también ha ido a favor de una mayor aceptación de este nuevo paradigma docente. Por ejemplo, ahora sería difícil volver a prácticas que ya se consideran superadas, como evaluar con un único examen el contenido de una asignatura o basar la docencia exclusivamente en la clase magistral.

La reflexión sobre las competencias que deben adquirir nuestros estudiantes obliga también a plantearse cuáles son las competencias que debería poseer el profesor universitario (como la necesidad de adaptación a unos nuevos entornos para los que no fue preparado).

Otra cuestión que ha emergido de todo este proceso es cómo evalúa la docencia del profesor la propia institución universitaria. En la actualidad, las universidades españolas y catalanas han iniciado políticas de evaluación de la docencia. Si bien hasta hace poco la evaluación del profesorado se limitaba a la valoración de su actividad investigadora, desde la LOU la docencia pasa a tener mayor peso. Cada universidad, previa aceptación por parte de la agencia de calidad correspondiente, diseña su propio modelo de evaluación del profesorado, donde se valora el autoinforme que realiza el profesor, la dirección de trabajos finales de grado y máster,

la participación en proyectos de innovación docente, en seminarios y cursos formativos sobre diseño de herramientas para la calidad en la enseñanza universitaria, además de las encuestas institucionales de satisfacción de los estudiantes.

Para este último colectivo, tampoco fue fácil la adaptación al nuevo sistema, pues suponía romper con una tradición bastante extendida de concentrar sus esfuerzos en las semanas previas a la preparación de un único examen. Con el nuevo sistema de evaluación continua, el alumno se ve obligado a realizar pruebas distintas durante todo el curso académico que, si bien favorecen la conciencia sobre el propio aprendizaje, suponen un sobreesfuerzo si ha de compatibilizar sus estudios con un trabajo. Este sistema requiere un estudiante a tiempo completo si pretende finalizar sus estudios en el plazo previsto y obtener una nota superior a la media. Una vez más, la guía docente es el instrumento más útil para recordar al estudiante sus derechos y obligaciones.

En consecuencia, después de años de aplicación del nuevo modelo y después de haber sido sometido a evaluación por parte de la agencia de calidad del sistema universitario de Catalunya (AQU-Catalunya), contamos con evidencias más que suficientes para afirmar que son muchas más las luces que las sombras en todo el proceso hasta ahora descrito.

Bibliografía

- Allen, D. (ed.) (2000). La evaluación de los aprendizajes de los estudiantes. Una herramienta para el desarrollo profesional de los docentes. Barcelona: Paidós.
- Carreras, J. & Perrenoud, P. (2006). El debat sobre les competències a l'ensenyament universitari. Barcelona: ICE-UB.
- De-la-Mota, C. & Puigvert, G. (2009). La investigación en Humanidades. Madrid: Biblioteca Nueva.
- García, Felix & Morant, J. L. (2005). Declaración de Bolonia. El Espacio Europeo de Educación Superior. Madrid: Acta.
- Garin, E. (1986). El renacimiento italiano. Barcelona: Ariel.
- Iglesias-Fonseca, J. A. (2009). Tot descornat e desfullat, que una carta no.s tenia ab altre: de cómo fragmentos y membra disiecta de manuscritos son también testimonio de la actividad de la librería en la Barcelona tardomedie-

- val. En Carme De-La-Mota & Gemma Puigvert (2009), *La investigación en Humanidades* (pp. 73-89). Madrid: Biblioteca Nueva.
- Imbernón, F. (1994). *La formación y el desarrollo profesional del profesorado. Hacia una nueva cultura profesional*. Barcelona: Graó.
- Lévy-Leboyer, F. (2003). *Gestión de las competencias: cómo analizarlas, cómo evaluarlas, cómo desarrollarlas*. Barcelona: Gestión.
- Martínez, M. & Añaños, E. (coords.) (2008). *Experiències docents innovadores de la UAB en Ciències Socials i en Ciències Humanes / Experiències docents innovadores de la UAB en Ciències Experimentals i Tecnologies i en Ciències de la Salut*. Barcelona: UAB.
- Mateo Andrés, J. & Martínez Olmos, F. (2005). *L'avaluació alternativa dels aprenentatges*. Barcelona: Universitat de Barcelona.
- Miller, G. (1990). The assessment of Clinical Skills/Competence/Performance. *Invited Reviews*, 65, 63-67.
- Puigvert, G. (coord.) (2009). *Guies per a l'avaluació de les competències dels estudiants*. Barcelona: AQU, Agència per a la Qualitat del Sistema Universitari a Catalunya.
- Puigvert, G. & otros (2010). Ejemplos de evaluación de competencias en Filología Clásica. *Revista de Estudios Latinos (RELat)*, 10, 189-226.
- Rué, J. & Martínez, M. (2005). *Les titulacions UAB en l'Espai Europeu d'Educació Superior*. Madrid: Narcea.
- Sabbadini, R. (1967). *Le scoperte dei codici latini e greci ne'secoli XIV e XV*. Florencia: G.C. Sansoni Editore.
- Sánchez Delgado, P. & Gairín Sallán, J. (2008). *Planificar la formación en el Espacio Europeo de Educación Superior*. Madrid: Universidad Complutense de Madrid.
- Villa, A. & Poblete, M. (2007). *Aprendizaje basado en competencias. Una propuesta para la evaluación de las competencias genéricas*. Bilbao: Universidad de Deusto.
- Yáñez, C. & Villardón, L. (2006). *Planificar desde competencias para promover el aprendizaje*. Bilbao: Universidad de Deusto.

Conversatorios

Primer conversatorio: Marco general de la evaluación de competencias y la evaluación en el aula

■ LEANDRO GARCÍA-CALDERÓN

Quiero darles la más cordial bienvenida a la segunda jornada de nuestro encuentro internacional universitario “La evaluación de competencias en la educación superior”.

Hoy continuamos con la reflexión en torno a la evaluación por competencias en los distintos niveles curriculares en un espacio más distendido que el de ayer. Tendremos un conversatorio entre los expositores y también con los expositores y el público. Este conversatorio tendrá dos partes. Durante la primera, discutiremos acerca de la evaluación en el espacio microcurricular y, en la segunda, vamos a dialogar en torno a la evaluación en los niveles intermedio y macrocurricular.

Para esta primera parte del conversatorio, “La evaluación por competencias en el aula”, nos acompañan el profesor Dieudonné Leclercq de la Universidad de Lieja, el profesor Álvaro

Cabrera del Teacher’s College de Columbia University, la profesora Patricia Escobar del Departamento de Educación de la Pontificia Universidad Católica del Perú y la profesora María Teresa Moreno que tiene a su cargo la coordinación del Área de Gestión curricular de la DAA.

Dejo con ustedes a la profesora María Teresa Moreno, quien va a conducir este primer bloque del conversatorio de esta mañana.

■ MARÍA TERESA MORENO

Muchas gracias, Leandro. Muchísimas gracias a todos por su presencia esta mañana. Yo creo que sería importante, Dieudonné, según queramos ir comentando estos temas, que nos cuenten un poco más por qué es tan importante tener en cuenta la “triple concordancia” como punto de partida. En otras palabras, ¿por qué es importante este proceso de calibración dentro de la evaluación por competencias en el aula?

■ DIEUDONNÉ LECLERCO

Me parece evidente que el profesor y el grupo, el equipo de profesores, deben haber pensado de antemano para qué sirve cada intervención, para qué sirve cada evaluación, para poder decir a los estudiantes a dónde queremos ir, para que sepan cuál es el reto último. Porque a un caminante que no sabe dónde va no le es posible caminar antes que el profesor, está condenado a seguirlo. Es decir que no puede tomar iniciativas porque puede ser que estas resulten inútiles; pero, si sabe a dónde vamos, se pueden preparar, puede tomar iniciativas en cuanto a la dirección que todo el mundo quiere lograr. Y también los profesores deben poder explicar para qué sirven todos los métodos y evaluaciones, para qué sirven ambos. En una lámina, he enseñado esta visión, que he llamado “ver el bosque” y no que el bosque quede escondido por los árboles, para que el estudiante pueda vincular cada trozo del “currículo”, cada curso, cada intervención con el reto, con el objetivo último, con las competencias y, lo más difícil de las competencias, las situaciones. Estas últimas pueden ser intermedias porque son competencias que son tan complejas que usualmente solo al final de una carrera pueden ser evidenciadas por los estudiantes, pero pueden existir competencias o niveles intermedios de estas.

■ MARÍA TERESA MORENO

Sin embargo, en algún momento hemos hablado que esta es una evolución de un tér-

mino que ya John Biggs usaba y se refería a él como “alineamiento curricular”; pero tú has preferido hablar de “triple concordancia”. ¿Cuál sientes que es una diferencia en esta precisión de una “triple concordancia”?

■ ÁLVARO CABRERA

Mientras Dieudonné piensa la respuesta, ¿puedo agregar un punto a la primera pregunta? Quiero ir un poco más atrás, algo tal vez incluso más básico, pero lo experimenté trabajando en Chile con profesores. Cuando uno les plantea esta idea de la triple concordancia, el primer desafío es tener total claridad de qué es lo que se quiere enseñar, qué es lo que efectivamente quiero que mis estudiantes logren en el curso. Tú le haces esa pregunta al profesor y es una pregunta en que él queda con una duda y te dicen: “Bueno, mi curso se trata de estos contenidos, de esta teoría” y uno piensa: “Ok, pero qué es lo que usted quiere que sus estudiantes hagan, qué es lo que va a demostrar... ¿que comprendan, que analicen, que apliquen estas teorías?”. Entonces esa primera reflexión ya es fundamental y es lo primero que necesitamos para después poder llegar a establecer la triple coherencia. Y cuando lo vemos desde el punto de vista de las evaluaciones, algo que también me daba cuenta muy seguido trabajando con profesores, es que la triple concordancia nos permite establecer que la evaluación que hacemos es totalmente pertinente y representa aquellos aprendizajes que son claves. Y no generamos

instrumentos de evaluación solamente con el contenido más difícil o, como nos pasaba en algunos momentos, profesores que generaban una evaluación en donde decían: “He ideado un problema donde sé que la mitad de los estudiantes va a fracasar... perfecto, tengo lista la prueba”. Y ahí le decimos: “Si usted considera que hay un aprendizaje que es clave, es fundamental en su curso y el cien por ciento de sus estudiantes lo logra y lo va a tener bien, eso habla bien de su docencia. Es una pregunta que usted tiene que mantener”. Algunos profesores decían: “Si tengo una pregunta que todo el mundo la va a tener bien, esa pregunta no la hago, la saco de mi prueba, ¿me entienden?”. En cambio ahora, con la triple concordancia, uno puede decir: “Voy a preguntar todo lo que sé que me demuestra a mí que el estudiante ha dominado todos los aprendizajes que yo considero claves en mi curso, independientemente que lo dominen todos o que sean contenidos más difíciles que los dominen algunos”, pero para nosotros era por esta idea que el profesor tenía que plantearse seriamente la pregunta: ¿qué es efectivamente lo que quiero que logren mis estudiantes a partir de mi curso?

■ PATRICIA ESCOBAR

Coincido totalmente. No es solo un problema en Chile, creo que aquí también en Perú tenemos muchos profesores, en general, de los distintos niveles, que tienen un problema para identificar qué aspectos o qué es lo que

los alumnos deben lograr. Pero una pregunta más: ¿cómo mi curso, además, contribuye al logro del perfil de egreso? Esta mirada no es solamente desde el curso, sino también cómo es que está contribuyendo a unas competencias que forman parte del perfil de egreso de esa carrera. Entonces, muchas veces esa mirada hacia qué estoy contribuyendo no se hace porque estamos mirando solamente a nuestro curso. Me preocupo de mi curso y no me preocupo de cómo colaboro con el perfil de egreso. Y esa coherencia que debe haber es una mirada de los procesos de enseñanza, de los procesos de aprendizaje y de la evaluación. Es una coherencia, una mirada de esos tres elementos que son fundamentales. Porque a veces los profesores nos preocupamos mucho por las metodologías de enseñanza, queremos innovar la metodología, queremos no utilizar ya la técnica expositiva sino el aprendizaje colaborativo, es decir una serie de técnicas innovadoras; pero nuestra evaluación sigue siendo una evaluación tradicional, alejada de los procesos de aprendizaje y enseñanza, una evaluación que simplemente no toma en cuenta aquello que debemos evaluar.

■ DIEUDONNÉ LECLERCO

Volviendo a la pregunta sobre el concepto de alineamiento (*alignment* en inglés), regresa hasta Taylor, en 1950. La idea es que primero son los objetivos, después los métodos y luego la evaluación. La ventaja de presentar la triple

concordancia es que, cuando se construyen las ideas, no vienen todas de los objetivos. Hemos definido los objetivos o las competencias una vez totalmente, después concebimos los métodos. Hay un camino en el que uno influencia al otro y es permanente. Por ejemplo, si sigo dando cursos cada año, cambio algo, pero lo que cambio puede ser un año una evaluación, aunque si cambio una evaluación tengo que pensar si va a servir a un objetivo o no. O esta evaluación puede revelarme que he olvidado un objetivo, que he olvidado una competencia, entonces modifico mi currículo, pero la idea al principio de la evaluación a menudo viene de la metodología. Hemos encontrado en un libro una visita, un nuevo método que nos parece interesante; pero si es tan interesante, tiene que estar al servicio de un objetivo y debe ser evaluado, porque los estudiantes no aprenden esencialmente lo que se evalúa y, si es válido, si es interesante, se debe evaluar. Es decir que la triple concordancia ilustra el hecho de que las buenas ideas vienen de todas partes, pero que siempre se asegura la triple concordancia.

■ MARÍA TERESA MORENO

Ayer se habló mucho de evaluación basada en evidencias. De pronto, el resultado de aprendizaje ya puede insinuar alguna evidencia y se puede pensar en la evaluación, no al final como tradicionalmente se hacía, sino un poco antes. Yo quisiera preguntarles: cuando hemos trabajado el tema de la triple concordancia,

muchas veces con los profesores, hay hasta duda en ellos sobre cómo la concebimos como una oportunidad de aprendizaje para los estudiantes, que es un poco lo que Álvaro Cabrera remarcaba, sin correr el riesgo de esta percepción de que estamos bajando el nivel de exigencia, por ejemplo, en la evaluación. ¿Qué opinan?

■ ÁLVARO CABRERA

No me queda claro de dónde podría venir esta idea de bajar el nivel de exigencia.

■ MARÍA TERESA MORENO

Te lo comento. En el ámbito universitario, a diferencia del ámbito de la educación básica regular, es probable que se planteen situaciones de evaluación o exámenes, con preguntas que no se han desarrollado muy profundamente en el aula. Una es la demanda cognitiva —es decir, el nivel de exigencia o de profundidad con el que se trabajaban algunos temas en el aula— y otra es la exigencia o el nivel de profundidad que se plantea en la evaluación. Entonces, hay esta idea de que estamos en la universidad, no estamos en la escuela. Sé que no es una pregunta sencilla.

■ ÁLVARO CABRERA

Yo lo abordaría haciendo otra pregunta: ¿para qué finalmente estamos evaluando? ¿Estamos evaluando para fomentar los aprendizajes de

los estudiantes o —muy entre comillas, entre varias comillas— “para separar los buenos de los malos”, para generar un ranking “estos sirven, estos no”? Evaluamos y concebimos la evaluación como una parte fundamental del proceso de aprendizaje y en ella se trata de recolectar evidencia de esos aprendizajes, no solamente para poner una nota, sino también, o incluso en primer lugar, para retroalimentar al estudiante acerca de dónde está, cuánto ha avanzado, qué le falta, qué más tiene que hacer. Por la tanto, me parecería incluso una cuestión de justicia, o de injusticia en ese caso, trabajar durante el semestre ciertos tipos de problemas, ciertos tipos de preguntas con el estudiante, y luego, durante la evaluación, hacer algo totalmente distinto o con un nivel de dificultad mucho mayor. Eso me parecería incluso un asunto de injusticia con esos estudiantes. Obviamente, no se trata de hacerles siempre el mismo examen o repetirles siempre la misma pregunta, pero sí preguntas equivalentes en el nivel de dificultad, temas que de alguna manera se han trabajado durante el semestre con ellos. Porque eso es lo que queremos: que finalmente aprendan, que la mayoría apruebe los cursos. Suponemos que hacia allá nos estamos moviendo y por tanto la coherencia es un elemento fundamental si pensamos que estamos evaluando para fomentar el aprendizaje y no para distinguir aquellos estudiantes que aprenderían aunque no les enseñemos, ¿verdad? Todos sabemos que tenemos estudiantes brillantes, que tú los

metes en un armario y en cuatro años van a salir sabiendo más. La evaluación no se trata a lo mejor de detectar solamente a esos estudiantes brillantes, sino de que la evaluación sea un proceso orgánico con el resto, con la enseñanza, con las demás acciones que hacemos para fomentar el aprendizaje.

■ PATRICIA ESCOBAR

Quizá complementar ahí señalando que, al contrario, no es correr el riesgo de bajar el nivel de exigencia, sino asegurar más bien una exigencia del propio estudiante. Si el estudiante tiene claro por qué y para qué se le evalúa, creo que se compromete con el proceso de enseñanza y aprendizaje, con ser partícipe de la evaluación. Y quizá también es importante señalar esta intención formativa que debe tener la evaluación. Alguien decía que la evaluación debe tener intención formativa, ¿no es cierto?; aunque a veces esta no se logra, pero es la mirada, la mirada de la evaluación. Y quizás ahí también cabe mencionar el tema que indicabas de las evidencias de aprendizaje. Muchas veces los profesores no nos preocupamos mucho de seleccionar cuál evidencia de aprendizaje es la que finalmente va a dar cuenta del resultado de aprendizaje que hemos planteado en nuestro curso. Cuando nosotros diseñamos los cursos y tenemos en cuenta nuestras competencias, desempeños y resultados de aprendizaje, yo tengo que apuntar a estos resultados. ¿Qué evidencia de aprendizaje me va a permitir

evidenciar ese resultado de aprendizaje en mi curso? Entonces, la evaluación tiene que ser planificada. Es intencional, hay una intención de evaluar, por lo tanto tenemos que planificar y diseñar muy bien qué es lo que queremos evaluar y seleccionar muy bien cuáles son esas evidencias de aprendizaje.

■ ÁLVARO CABRERA

Algo muy breve. Justamente con lo de la exigencia. Yo creo que estamos todos de acuerdo en que las exigencias sean altas, pero eso no puede estar únicamente expresado en la evaluación, ¿verdad? Nuestros resultados de aprendizaje tendrán que expresar también esa alta exigencia. Las metodologías que utilizamos tendrán que ser demandantes también con los estudiantes y, en coherencia, podremos tener alta exigencia en la evaluación también; pero no puede de verdad ser la alta exigencia solamente en la evaluación, sin que eso esté expresado en los otros dos componentes.

■ DIEUDONNÉ LECLERCO

Sí. Y pienso que hay ambigüedades e ilusiones en el término “exigencias”. Hay un ejemplo que lo ilustra bien. En los años setenta, se inició un movimiento de renovación de los estudios de medicina que comenzó en Hamilton, Ontario, y sus seguidores, que ahora son los que están más comprometidos con eso y los más expertos con quienes trabajamos en Europa,

son los de la universidad de Maastricht, de la formación de medicina (seis años), aprendizaje basado en problemas, en casos, porque para los médicos generales, los médicos de familia, los problemas son casos individuales, personas que están enfermas. Y desde 1960 hasta ahora, han habido tantas experiencias y medidas comparativas de estudiantes que habían sido formados en el sistema tradicional basado en la transmisión de conocimiento y más o menos la mitad de las universidades del mundo ahora han adoptado el sistema, de modo que hay mucha oportunidad de comparar. Pero comparándoles, ¿con qué instrumentos?, ¿mirando qué para compararles? Primero, conocimiento de memoria, conocimiento en general, y existen pruebas internacionales o norteamericanas (si hay médicos, las conocerán): NBME, National Board of Medical Exams, que son preguntas de selección múltiple. Hay tantas experiencias que hicieron metaanálisis en 1993, en 2003 y en 2013, cada diez años. El resultado de eso es que los formados con los métodos tradicionales de transmisión de conocimiento superan a los formados por problema de casos, por aprendizaje basado en casos, basado en problemas. Pero también lo han medido en sus capacidades clínicas con pacientes y ahí son los otros que superan. ¿Para qué sirve el conocimiento que no se puede aplicar con pacientes? Han hecho comparaciones muy importantes de tipo: la tasa de consultas que no son solamente físicas, sino también psicológicas; también el costo al estado, porque son

análisis complementarios que alguien debe pagar, a menudo es el Estado el que paga el sistema de seguridad social, todo eso. Hay comparaciones muy detalladas y eso plantea el problema de las exigencias. ¿Qué exigencias? ¿Que conocían mucho o que lo saben aplicar? Hay un problema interesante, la cuestión que se plantea y que revela la ambigüedad de la palabra “exigencias”. Por ejemplo, a menudo en Europa, los de mi edad, dicen “el nivel baja”. Hace dos mil años ya también decían que “el nivel baja”. Y el nivel ha bajado tanto durante tanto tiempo que ahora solo pueden ir a la Luna, ¿ves? Tanto ha bajado el nivel. Es tonto.

■ MARÍA TERESA MORENO

¿Qué es lo que yo quiero evaluar? Y esto es interesante. Ayer, creo, Álvaro traía esto a colación. No es un tema simple, sino complejo, que supone tomar decisiones pensadas. Me gustó eso que decías ayer: hay que aproximarse a este asunto como cuando nos aproximamos a nuestro quehacer de investigación como docentes. Mencionabas este otro concepto, que trae Dieudonné y tocas tú en el libro *Ideas e Innovaciones*, que es el dispositivo de evaluación del aprendizaje (DEA). Y lo definieron ayer como un conjunto de evaluaciones, un grupo coherente de evaluaciones, que permiten evaluar ya sea un curso o un plan de estudios. ¿Podemos hablar un poquito, ya que ustedes son coautores de este libro, de esta idea del dispositivo de evaluación de aprendi-

zaje? ¿Cómo lo entienden? ¿Como un conjunto coherente de evaluaciones en esta lógica de tomar decisiones?

■ DIEUDONNÉ LECLERCO

En una de las primeras láminas, cuando se trataba de reanimar en la calle, se enseñaba que hay que medir los saberes, las destrezas aisladas en papel, y también—es lo que faltaba antes— en situaciones más complejas. Las dos. Cuando has dicho eso, he pensado en algo que dijo ayer Álvaro, refiriéndose a que nosotros los docentes tenemos la posibilidad de hacer investigación de clase, porque es una cosa que nos gusta mucho, somos universitarios, hemos renunciado a la investigación para enseñar, tenemos el corazón partido en dos. Nosotros, los profesores, los docentes, tenemos una ventaja que los investigadores de laboratorio no tienen. Ellos pueden hacer algo que nosotros, los docentes, no podemos hacer, que es anticipar grupos de control, grupo testimonio, un grupo experimental, dos grupos, eso no podemos hacer porque no se puede controlar la vida. En el laboratorio, con ratones, se puede; pero nosotros tenemos una ventaja: tenemos la misma cohorte, al menos durante un año, y podemos hacer investigaciones de evolución—lo que se llama “longitudinal”—, ver el desarrollo. Y el desarrollo es una palabra clave en términos de competencias: desarrollar las competencias, ver las evoluciones y encargar a los estudiantes ser su propio observador de

su desarrollo, algo en lo que un investigador de fuera tiene dificultades porque no está en esta clase. Y se reemplaza la fuerza de los investigadores de laboratorio que pueden decir: hemos testeado dos mil estudiantes, pero una vez (lo que se llama en inglés *one-shot*). No saben nada de antes o después. Nosotros tenemos una ventaja: la de poder hacer investigaciones que se publican, que son más clínicas, pero que están basadas en la evolución, llamadas “longitudinales”.

■ ÁLVARO CABRERA

Yo contestaría con dos ideas. Una es que, debido al tipo de aprendizaje que ahora queremos lograr, o que siempre hemos querido lograr, que son múltiples aprendizajes en nuestro curso, no existe un instrumento de evaluación que nos permita evaluar todo y por eso tenemos que aproximarnos a este tema pensando en una variedad de instrumentos de evaluación. Ninguno, ni el portafolio ni la simulación, nos va a permitir a lo mejor evaluar todo lo que queremos. Y la otra parte de ese mensaje es decir que todas las formas de evaluar son válidas también. Lo mencionábamos ayer y Dieudonné trabaja mucho con las preguntas de selección múltiple. Tampoco se trata, a lo mejor, de decir que las preguntas de selección múltiple hoy día ya no sirven o están obsoletas. Sobre todo si utilizamos preguntas de selección múltiple que exigen procesos cognitivos superiores, que no

son solamente reconocimiento o evocación de un conocimiento, si utilizamos los grados de certeza, etc. Lo importante es que, si las utilizamos, sepamos que las estamos usando efectivamente para lo que sirven y tenemos conciencia que otro tipo de desempeños no pueden ser evaluados de esta manera y, por lo tanto, tenemos que combinar. Entonces, para mí, la fuerza de este dispositivo es decir: “necesitamos una combinación de instrumentos” y adquiere incluso más fuerza cuando eso lo podemos concebir en una línea de cursos, a través de varios semestres, varios años, y los profesores pueden juntarse y decir: “bueno, lo que estos cuatro, cinco, seis cursos en conjunto quieren desarrollar es esto y vamos a concebir un dispositivo donde en cada semestre apliquemos esta variedad e incluso a nivel del programa completo. Entonces, no hay un instrumento mágico que nos permita evaluar todo, pero a la vez cada instrumento tiene su función y tenemos que ser conscientes de cuál es esa función para poder combinarlos de buena forma.

■ DIEUDONNÉ LECLERCO

Y como hoy, así como ayer, se trata de evaluación, tengo una buena noticia. Hattie es un profesor de Australia que ha dedicado treinta años de su vida a estudiar los métodos que son más eficaces en los aprendizajes y ha publicado un libro grande que se llama *Visible Learning* en el año 2009. Su medición

es amplitud del efecto, llamada *effect size* en inglés. En términos de amplitud del efecto, del beneficio sobre los aprendizajes, una de las cosas que es la más eficaz de todas es la evaluación con intención formativa, pero hay condiciones. ¿Qué condiciones? No basta con evaluar, hay que retroalimentar. Hay que forzar a los estudiantes a que piensen no solamente en dónde se han equivocado, sino por qué se han equivocado. Y el porqué es un problema de diagnóstico. El dónde es un problema de juicio, de grados de certeza: está bien o mal, pienso no estoy seguro, juicio. Pero el por qué no estás seguro, eso es un problema de diagnóstico. Como cuando vamos al médico, le decimos que nos sentimos mal y este nos dice: “Sí, es correcto, estás mal”; pero no vamos al médico para que nos diga que estamos mal, sino para que encuentre el porqué, el diagnóstico. La evaluación con intención formativa, para ser eficaz, tiene condiciones: que el estudiante tenga una reflexión sobre el porqué, que llamamos el diagnóstico, y después el remedio; pero este último es, en medicina, la prescripción, lo que hay que hacer, y no hay que prescribir si no hemos tenido un diagnóstico antes. A menudo, lo que pasa en las universidades —trabajo en varias universidades— es que hay una prueba y hay una carta que dice: “Los que no han tenido un diez sobre veinte se les espera el miércoles a las diez menos veinte para la ‘remediación’”. Entonces van y hay una persona que hace la “remediación” al

por mayor, la misma para todo el mundo. Un poco como si en medicina hubiese un cartel que dijese: “Todos los enfermos, el próximo martes van a recibir la píldora, la misma que va a servir a todos”. Si no hay diagnóstico y, mejor aun, autodiagnóstico —doy cursos en aulas de trescientos, cuatrocientos estudiantes, que tienen dieciocho o veinte años; no puedo hacer el diagnóstico—, pueden hacerlo, deben hacerlo, porque son profesionales. Debemos ser exigentes con ellos. Deben hacerlo y yo les pido que me envíen —son cuatrocientos— una metacognición retrospectiva sobre las cuatro evaluaciones que han hecho: una en octubre, una en noviembre, una en diciembre y en enero deben enviarme el análisis metacognitivo de los diagnósticos que han hecho sobre sus errores. ¿Por qué se hace eso? La buena noticia es que uno de los métodos más eficaces de formar es evaluar. En inglés se dice: *the best way to teach is to test*. Pero no solamente testear, el resto también es muy importante. Les recomiendo que lean a Hattie¹.

■ MARÍA TERESA MORENO

Gracias, Dieudonné. Yo quisiera más adelante entrar un poquito más al tema de la metacognición y a profundidad, pero antes me gustaría

¹ Nota del editor: el autor hace referencia, particularmente, al libro *Visible learning* (London: Routledge, 2009).

preguntarle a Patricia: ¿qué pasa cuando un profesor se toma en serio el tema de implementar una evaluación por competencias en sus cursos? Te preguntaría en esa línea, ¿cuáles han sido las cosas con las que estás satisfecha, cuáles son los retos que has enfrentado, cómo les has salido al frente? Cuéntanos un poquito de tu experiencia en esa línea.

■ PATRICIA ESCOBAR

Creo que, en general, uno de los aspectos —por lo menos en mi caso, en la Facultad de Educación, que ya desde 2013 venimos implementando un modelo curricular por competencias—, es que ha habido primero un compromiso de todos los profesores por asegurar y comprometernos con un perfil de egreso por competencias donde están las competencias y los desempeños. Luego, la tarea de cada profesor ha sido elaborar los resultados de aprendizaje para cada uno de sus cursos, articulados a los desempeños y a las competencias, de tal manera que nosotros podemos decir: en mi curso, yo puedo contribuir al logro de esa competencia. Porque no es que yo voy a asegurar la competencia en mi curso; ya que la competencia, como hemos visto, es un proceso, entonces eso es lo primero. Pero un reto sumamente importante es el de planificar la evaluación, sentarnos a decir qué es lo que yo quiero si tengo estos resultados de aprendizaje; muy bien, cómo voy a trabajar durante este semestre con mis estudiantes para ayudarlos a lograr esos resultados de aprendizaje. Tengo

que diseñar mis evidencias de aprendizaje, mis actividades de evaluación, que me permitan decidir de qué manera estos alumnos van a participartambién en la evaluación. Porque creo que ese es otro elemento importante y un reto también para los profesores porque pensamos que somos los únicos que tenemos que evaluar. Si el proceso de enseñanza-aprendizaje es del estudiante, este tiene que participar en la evaluación. Tenemos que hacerlo partícipe de alguna u otra manera, sea autoevaluando su propio proceso de aprendizaje o coevaluando a sus compañeros. Entonces, creo que algo que a mí particularmente me ha resultado es comenzar con la coevaluación, en trabajos en equipo: ayudarlos a que, con una ficha de evaluación donde están los criterios y los indicadores claramente trabajados con ellos de manera consensuada, puedan evaluar el trabajo de equipo de sus compañeros. Esto permite no solamente que puedan evaluar el trabajo de otro equipo, sino también aprender porque, finalmente, creo que el objetivo último de la evaluación es aprender de ella, de los aciertos y de los errores. Creo que otro elemento también importante, que lo mencionó Dieudonné hace un momento, es el tema de la retroalimentación. ¿Qué sucede? Cuando los profesores planificamos nuestras actividades de evaluación y los dispositivos hacen que tengamos que diversificar los instrumentos —o sea, no solamente tener un examen parcial y un examen final, sino qué otras actividades tenemos—, tengo que darme el tiempo y el espacio para hacer una

retroalimentación efectiva; de tal manera que, si el trabajo que presentó fue un ensayo, un artículo o una monografía, el alumno reciba efectivamente una retroalimentación que le permita aprender. Pero muchas veces simplemente los profesores, en algunos casos por tiempo, nos fijamos solo en los aspectos quizá negativos o le ponemos una nota y punto. Entonces el alumno recibe su trabajo con un doce sobre veinte y luego dice: “Bueno, ¿qué significa el doce? ¿Qué hice bien y qué hice mal?”. Nosotros tenemos la responsabilidad como profesores de ayudar a que el estudiante pueda identificar esos aciertos y esos errores en sus trabajos. Y obviamente un instrumento que ha ayudado a eso es la ‘rúbrica’, pero esta no es nada fácil. O sea, el que diga que hacer una rúbrica en realidad es sencillo, miente. En realidad, es muy laborioso hacer una de tal manera que te permita ver estos descriptores en sus distintos niveles y tú puedas, como profesor, evaluar el trabajo; pero la rúbrica tiene la función también de que el estudiante pueda evaluar su propio trabajo mirando en qué nivel, de estos cuatro o cinco —dependiendo de cómo esté elaborada la rúbrica—, él pueda también decir: “Ah, perfecto, voy a entregar mi trabajo pero antes me autoevalúo. Tengo esta rúbrica que me permite decir: ‘Yo creo que en este criterio tengo esto; en este criterio estoy así, y he cumplido o no he cumplido con las exigencias del trabajo’”. Entonces, creo que el implementar una evaluación por competencias es todo un reto; en realidad, no es sencillo y se trata de un trabajo en equipo. Es un trabajo

colaborativo, pero también muy de cerca con los estudiantes. Si nosotros no tenemos una relación cercana con los estudiantes, donde los ayudamos a que sientan que son partícipes y responsables de su aprendizaje, no vamos a lograr el cambio. Muchas veces el profesor es solamente el que se capacita, el profesor es el que cambia, pero los alumnos también lo hacen. Los alumnos vienen muchas veces y tienen realidades diferentes. Proviene de escuelas diferentes donde su proceso de enseñanza-aprendizaje ha sido probablemente distinto al que van a encontrar en la universidad y también tienen que aprender a cambiar sus formas de aprendizaje, a introducirse en el mundo universitario donde nadie les va a dictar un curso, sino donde tienen que tomar apuntes y eso ya es un aprendizaje, eso ya es un reto para ellos. Nosotros tenemos que comprender que en las aulas universitarias tenemos una heterogeneidad de alumnos, sobre todo en el caso de Educación, donde la heterogeneidad es muy grande, y supongo que en las otras carreras también.

■ DIEUDONNÉ LECLERQ

Quisiera añadir algo a lo que Patricia acaba de decir. Una pregunta muy interesante para plantear a los estudiantes cuando hay trabajos recurrentes —por ejemplo, presentar y redactar un informe—: comparado a mi anterior informe, este ha mejorado en los siguientes criterios. El estudiante se compromete a mejorar y a decir dónde ha mejorado. Demasiadas

veces aceptamos: “Te he dicho que no hay que hacer los párrafos así, que hay que estructurar, poner títulos, enumerar las páginas”. La primera, no lo hice; la segunda, no lo hice; y la tercera, no lo hice. No tienen progreso. Deben comprometerse y debemos evaluar la progresión. Y tienen la responsabilidad de evaluarse, de decir: “Aquí están las evidencias de que he mejorado”, porque no sirve de nada que les demos retroalimentación si no hacen lo que les decimos.

■ MARÍA TERESA MORENO

Eso me lleva, Dieudonné, a la siguiente pregunta que tenía para todos nosotros. A menudo se dice que uno realmente aprende cuando enseña. Y, en esa línea, me preguntaba: ¿cómo se puede trasladar a los estudiantes la responsabilidad por su aprendizaje a través de la evaluación? Dieudonné lo ha puesto sobre la mesa y me gustaría recoger sus ideas en esa línea.

■ ÁLVARO CABRERA

Primero, solo colgándome un poco de la respuesta anterior, encuentro mucha razón en lo que dice Patricia acerca de cuáles son los desafíos, las dificultades de implementar esto. Es superar todo tipo de inercia, no solamente nosotros como profesores, sino también la inercia que traen los estudiantes, muy acostumbrados a un tipo de evaluación desde las escuelas. Ahí el mensaje es no perder la espe-

ranza porque uno empieza a hacer estas innovaciones y, el primer semestre, funciona más o menos; el segundo semestre, un poquito mejor; y a veces uno tiene ganas de decir: “Bueno, esto no funcionó, vuelvo a lo que hacía”. Y en realidad hay que perseverar, porque es un cambio cultural no solamente de los profesores y de la institución, sino también de los estudiantes. Parte de este cambio cultural es hacerlos que asuman un mayor protagonismo. Para mí, la clave es una palabra que ya hemos repetido muchas veces: la retroalimentación, y el desafío para nosotros como profesores es asumirla no como una pérdida de tiempo o una carga, algo que va a ocupar el tiempo que deberíamos estar enseñando, ¿verdad? A veces lo sentimos así. ¿Cómo voy a dedicar la mitad de la clase después del examen de mitad del semestre a retroalimentar si tengo que seguir pasando la materia? Por eso, el desafío es entender que la retroalimentación es un momento clave del proceso de aprendizaje y tenemos que incluirlo y diseñarlo de esa manera. Y —es fácil decirlo aquí en un escenario— el desafío tal vez es, del contenido que nosotros queremos enseñar, poder detectar con mucha claridad cuáles son esas ideas fundamentales (en inglés lo llaman las *core ideas*): esa especie de bloques fundamentales sobre los cuales el resto del conocimiento de lo que queremos explicar o queremos lograr en los estudiantes se va construyendo. Y en estas *core ideas*, en estas ideas fundamentales, acompañar a los estudiantes a aproximarse, a analizar; luego evaluar, retroalimentar. Acompañarlos en el proceso de cómo

un profesional, un académico, se aproxima a esta idea, y tomarnos el tiempo de hacer eso. Y otros contenidos, simplemente decir: “Bueno, yo voy a acompañarlos a ustedes en el proceso en torno a esta idea. Pero acá hay otra idea que la vez anterior yo solía enseñar, en mi cátedra. Esta idea ahora ustedes la van a explorar de forma autónoma, pero yo ya he modelado la manera en que uno debe aproximarse a esa idea”. Entonces, si pensamos que a lo mejor nuestro curso es hacer una espiral de cómo aproximarnos y analizar distintos contenidos, algunos lo haremos en clase, lo haremos con ellos, los evaluaremos, haremos retroalimentación y en otros casos les diremos que es responsabilidad de ellos. Aprender, estudiar, explorar ese otro contenido y poder evaluarlo y probablemente retroalimentarlo. A través de la evaluación, creo yo que la manera de entregarle la responsabilidad al estudiante es justamente retroalimentar, modelar la manera en que vamos a estudiar este fenómeno, retroalimentar cómo lo ha hecho y luego decirle: “Ahora es tu turno”, fuera del tiempo de clase; porque, si hacemos este proceso muy profundo, probablemente no podremos ya enseñar de esta manera, a lo mejor, las quince grandes ideas que teníamos: tendremos que enseñar ocho y dejar las otras siete al estudio autónomo de los estudiantes.

■ MARÍA TERESA MORENO

Yo quisiera, llegados a este punto, comentarles algo sobre lo cual me parece interesante

que podamos continuar nuestro diálogo esta mañana. El tema de la formación por competencias tiene que ver con el tema de *life-long learning*, ¿no? Aprender a lo largo de la vida. Creo que algo que se tocó ayer puede contribuir a hacer esto realidad, y es este tema de la “metacognición”. Sin embargo, nosotros, en la universidad, hemos tenido algunos problemas con este asunto. Y es una de estas cuestiones de cuando miramos nuevamente la educación básica regular. Sinceramente, creo que la educación básica regular, con respecto a los avances que ha tenido en términos del rol del docente, de las metodologías de enseñanza de aprendizaje, tiene varias sugerencias que darle a la formación universitaria; sin embargo, son dos ámbitos diferentes. Entonces, la universidad tendrá que pensar también cómo se plantea una formación por competencias en la universidad, que es un espacio de formación distinto. Y ahí les cuento un poquito —y esto se los cuento de primera mano— lo que ha ocurrido en la educación básica regular con el tema de la metacognición. Como es un tema complejo, algunos docentes creyeron que, si al final de la sesión de clase le preguntaban a los estudiantes: ¿qué has aprendido?, ¿cómo lo has aprendido?, ¿para qué crees que lo puedes usar?, entonces ya están haciendo metacognición. Esto de pronto podría funcionar en estudiantes pequeños, pero el riesgo que se corre allí es creer que eso es y de eso va la metacognición. Hace poco hablaba con una profesora de Estudios Generales Letras de

la PUCP, que me decía: “No quiero escuchar esa palabra, María Teresa”. Uno ve con pena cómo una idea tan potente de pronto, al no trabajarse bien, ya no se quiere recibir. Pero ayer ustedes hicieron un excelente trabajo mencionando tres componentes que creo son centrales en la correcta comprensión de la metacognición. Hablaron de “autojuicio” —lo ha ido mencionando un poco Dieudonné también—, “autodiagnóstico” y “autorregulación”. Un poco como si fuésemos en etapas, de un “autojuicio” inicial a ser capaces de hacer un “autodiagnóstico”, para finalmente lograr que estas personas se “autorregulen”. Podemos profundizar un poco de manera tal que podamos todos aquí entender la potencia enorme que tiene la metacognición en el aprendizaje cuando la entendemos bien.

■ DIEUDONNÉ LECLERCO

Cuando dices que al final de un curso los profesores pueden pedir a los estudiantes que tomen un trozo de papel diciendo lo que han aprendido, sí, es metacognición; pero el reto es que todos los estudiantes lo hagan a un nivel suficiente de profundidad, porque no sirve de nada que escriban: “He aprendido varias cosas sobre Biología y cómo se escucha”. Eso es metacognición, pero no a un nivel suficiente. Para que esté a un nivel suficiente, debemos de tocar los procesos mentales mismos y esto se puede lograr solamente si consideramos situaciones precisas. Cuando vas a ver una pelí-

cula, ¿qué recuerdas? Es demasiado vago. Pero si les decimos: “Hemos visto una película”, ahí cuál es la probabilidad de que va a recordar eso, eso, eso o que va a comprender. Vamos a hacer preguntas de comprensión: ¿cuál es la probabilidad de que va a comprender eso, eso, qué va a ser capaz de aplicar eso y eso? Estamos a un nivel de juicio adecuado porque se puede verificar que lo ha aprendido o no. Podemos verificarlo y él también. Leandro me contó que les dijo a sus estudiantes: “Pregunta en Internet sobre las estrellas” y él le pide a sus estudiantes que le expliquen. Los estudiantes leen la pantalla. Él les dice: “No, no leas en la pantalla, explícame”. Que sea el libro o que sea la pantalla, lo saben leer, pero no lo saben explicar. Hay una confusión. Nivel apropiado. El segundo nivel es autodiagnóstico. ¿Por qué me equivoqué? Me equivoqué porque he leído mal, porque... Tenemos el deber de entrenar a los estudiantes a autodiagnosticarse. No sirve de nada que su autodiagnóstico sea “he pensado mal” o “no he leído bien la pregunta”. ¿Qué no han leído bien? ¿Qué palabra? ¿En qué sentido no lo han leído bien? Debemos entrenarlos a hacer buenos diagnósticos, como un médico. No es suficiente que un médico le diga al enfermo: “Tú tienes una enfermedad grave”. ¿Qué? Grave. ¿Ves?, debemos entrenarlos a ser precisos y, por supuesto, no es innato. Lo deben aprender. También en la regulación; es decir, cuando saben lo que pasa, las causas, debemos entrenarlos. ¿Y ahora qué hacemos?: intervenir.

■ ÁLVARO CABRERA

Quisiera decir dos cosas más generales. Primero, la metacognición es algo tan importante que, a mi juicio, la capacidad de metacognición es lo que define a un profesional. Es decir, un profesional es una persona que ejerce una tarea, cualquiera que esta sea, y es capaz de monitorear lo que está haciendo, de saber si va bien o mal, de corregir. Nosotros, cuando hacemos clase, vamos haciendo esta reflexión probablemente durante ella, vamos monitoreando qué es lo que está pasando con nuestros estudiantes, hacemos cambios; pero también lo hacemos a la mitad del semestre o al final. Entonces, la metacognición es algo tan importante como decir que define la capacidad de actuar profesionalmente, ¿verdad? De actuar en la ambigüedad, de resolver problemas complejos. No simplemente, a lo mejor, de seguir un protocolo y si algo falla ir a ver un manual y saber qué es lo que pasó.

■ MARÍA TERESA MORENO

Álvaro, discúlpame que te interrumpa en este punto, pero por eso hacía la pregunta, porque siendo tan clave necesitamos comprenderla bien en su real dimensión para luego ver cómo la trabajamos.

■ ÁLVARO CABRERA

Exacto. Y el otro punto es que tenemos que entender que nuestros estudiantes están

desarrollando la metacognición permanentemente. El desafío nuestro ahora es que todos lo hagan, ¿verdad? Al menos a un cierto nivel. Otros lo harán por naturaleza o por ciertas capacidades, lo harán con profundidad. Y ahí está, ya se ha mencionado, hay una definición operacional que el profesor Leclercq ha realizado, diciendo que la metacognición tiene tres procesos principales: el juicio, el diagnóstico o análisis casual, la regulación. Parece muy lineal, pero por supuesto en nuestras propias mentes esto puede ocurrir, podemos partir por un análisis, después llegar al juicio, si queremos aproximarnos a esto de manera analítica, lo podemos dividir en estos tres aspectos. Y el desafío es poder generar momentos y a lo mejor instrumentos para conscientemente ir trabajando con los estudiantes estas tres capacidades. Ahí a lo mejor podríamos muy brevemente tratar de mostrar un ejemplo que ha desarrollado el profesor Leclercq, que es lo que llama los “tests espectrales metacognitivos”, aplicados a un instrumento de evaluación muy particular que nuevamente son las pruebas. Pueden ser pruebas de respuestas breves, de elección múltiple, como mostrábamos ayer en la presentación. Yo mostraba particularmente esto de los grados de certeza. ¿Qué tan seguro estoy de mi respuesta? Eso es el juicio. No desarrolla el diagnóstico, no desarrolla la regulación; pero al menos desarrolla el juicio. Si nosotros hiciéramos una siguiente etapa, en la siguiente clase, luego del examen, y eso implica tarea extra porque implica entregarles cada vez a los estudiantes

el examen y volver a generar otras preguntas cuando venga a la siguiente clase, el siguiente año. Pero la retroalimentación, sin que el estudiante pueda tener las preguntas después, no es posible. Si en la siguiente clase revisamos con ellos este examen y les pedimos que vayamos conversando sobre aquellas respuestas que son correctas o incorrectas y por qué lo son, y ellos ven ahí, en caliente, inmediatamente, escribiendo, sobre todo aquellas preguntas en que se equivocaron y estaban muy seguros, estamos frente a un problema grave; también aquellas respuestas que tenían correctas, pero sobre las que estaban muy poco seguros. Entonces, ¿por qué dudaron, verdad? Les pedimos que vayan escribiendo y haciendo un diagnóstico rápido —que llamamos diagnóstico en caliente— y finalmente, como ya mencionaba el profesor Leclercq al final del semestre podemos decirle: “Usted reúna este material que surgió en una sesión de retroalimentación de las tres, cuatro pruebas del semestre y haga un pequeño informe metacognitivo leyendo todo esto que le fue pasando en el semestre, para que pueda decir dónde estuvieron sus mayores dificultades, por qué estas ocurrieron, qué es lo que debería usted cambiar a lo mejor con respecto a cómo se aproxima a estos contenidos, a estas teorías”. De esa manera, ocupando un poco más de tiempo, pero de un tiempo que es de aprendizaje de calidad, podríamos tener un mecanismo concreto muy claro para desarrollar tanto el juicio, el diagnóstico y la regulación.

■ PATRICIA ESCOBAR

Quizá uno de los aspectos que yo considero más difíciles de evaluar es justamente el proceso metacognitivo. He tenido una experiencia en maestría, no en pregrado, que particularmente considero que ha sido una linda experiencia con alumnos de Pronabec que fueron becados para seguir una maestría, que vinieron de distintas regiones del país y que querían obtener una maestría en currículo. Entonces, yo dictaba con ellos el curso de evaluación del aprendizaje y lo primero que hicimos fue preguntarles, porque era un curso electivo, ¿por qué querían llevar ese curso? Ellos decían porque el tema más difícil de trabajar en la educación básica regular es el de la evaluación. Comenzamos preguntándoles: ¿qué quieren saber de la evaluación?, ¿qué quieren aprender de la evaluación? Entonces hicieron una lista de temas. Más que temas, una especie de compromisos; es decir, “yo quiero aprender a seleccionar mis técnicas e instrumentos”, “quiero saber qué es la evaluación auténtica”, “quiero saber cómo se hace la retroalimentación”. Una serie de aspectos. Entonces, yo les dije: “Muy bien, esto que quieren aprender, lo vamos a trabajar, pero esto va a ser el compromiso que ustedes van a lograr al finalizar el ciclo. ¿Y cómo me van a evidenciar estos compromisos? Llamémosle resultados de aprendizaje —que luego lo tradujeron en indicadores—. Lo vamos a trabajar con el portafolio”. Entonces, ellos iban trabajando todo lo que desarrollábamos en el curso, desde lecturas,

investigaciones, textos, *power points*. Todo lo que surgió del curso iba alimentando su portafolio, pero al mismo tiempo ellos iban reflexionando, mirando ese compromiso que habían asumido. Por ejemplo, tuvimos una prueba con materiales. Ellos se sorprendieron y dijeron: “Es la primera vez que nos toman una prueba con materiales”. Nunca habían tenido una prueba así. ¿Por qué? Porque obviamente... las preguntas... podías utilizar el material porque eran casos, aplicación de casos. Y luego ellos, en sus portafolios, reflexionaban sobre cómo les había impactado el hecho de que sea una prueba con materiales; pero, además, ellos ya estaban pensando que eso lo podían hacer con sus alumnos. De esa forma, ya no solamente lo veían desde su propio punto de vista, sino la aplicabilidad. Estaban reflexionando sobre: “¿Por qué no hacer esto en mi clase? ¿Por qué siempre tengo que tomar un examen o una prueba sin materiales?”. Entonces, este proceso es muy importante y creo que el portafolio puede ser un instrumento que contribuye, que colabora a ello.

■ MARÍA TERESA MORENO

Llegados a este punto, quiero comentarles que, tal como Leandro mencionó esta mañana, la idea de estos dos conversatorios que vamos a tener es que podamos conversar, pero que también ustedes puedan hacer sus preguntas y conversar también con nuestros invitados. Entonces, yo voy a pasar ahora a leer sus preguntas. Hasta ahora hemos

hablado de la “triple concordancia”, de los dispositivos de evaluación de aprendizajes, como un conjunto coherente de instrumentos de evaluación que pueden permitir evaluar un curso o un plan de estudios. Por eso hemos dicho que hay que tomar decisiones todo el tiempo sobre por qué y para qué quiero evaluar. En esa línea, la importancia de las evidencias que son las que, de alguna manera, hablan de esos aprendizajes. Hemos hablado también de trasladar, a partir de la evaluación del aula, la responsabilidad por el aprendizaje al estudiante de cómo se desarrolla la autonomía a través de la evaluación. Finalmente, llegamos a este rico tema de la metacognición, cuya relevancia ha levantado Álvaro y nuestros tres invitados. Le hemos dedicado un tiempo de juicio, de reflexión, de diagnóstico, de autoregulación, para poder entenderla en toda su dimensión. Vamos a pasar a las preguntas del público.

Aquí hay una pregunta que plantea lo siguiente: “¿Qué porcentaje (proporción de evaluaciones sumativas y formativas) deben estar incluidas en un buen DEA?”.

■ DIEUDONNÉ LECLERCO

Sabiendo la buena noticia de que el mejor modo de evaluar es formativamente, diré que la mayoría debe ser evaluación formativa y que puede ser que no haya sumativa, que es diferente de la suma del conjunto de las formativas. Puede ser la evolución. Podemos decir que hemos visto la evolución y que no se ha

logrado el nivel que esperábamos. Pero deben ver venir el problema, tener tiempo para corregir y también digo que no me preocupa si un profesor me dice que no tiene una evaluación que sea solamente sumativa.

■ ÁLVARO CABRERA

Yo me atrevería a decir que la distinción más famosa en la literatura es entre evaluación sumativa o formativa; pero nosotros defendemos que esa no es una distinción correcta porque lo sumativo y lo formativo no están en el mismo eje. Nosotros hablamos de una evaluación con función sancionadora, para colocar una nota —en ese sentido, no en el sentido de castigo— no *versus*, pero en el mismo espectro de una evaluación con intención formativa, sin nota y para contribuir en el aprendizaje. En cambio, lo sumativo tiene que ver con la información que nos entregan a nosotros y al estudiante: solo una nota *versus* la cantidad de información, *versus* una evaluación diagnóstica informativa. Dicho eso, yo creo que toda la evaluación tiene que perseguir una intención formativa, incluso aquella que tiene nota. Debiésemos retroalimentarla toda y es bien importante a lo mejor tratar de diferenciar la idea de la evaluación de la idea de calificación. Probablemente nosotros estamos evaluando de forma permanente nuestra clase: cuando hacemos una pregunta, cuando abrimos la sala y vemos qué es lo que opinan nuestros estudiantes, cuando vemos qué es lo que no preguntan, también tene-

mos que estar haciendo una evaluación de cómo avanza el aprendizaje en esa sala. Y no siempre estamos obligados a colocar una nota, una calificación, ¿verdad? Entonces es probable que siempre estemos evaluando y ojalá lo hagamos con intención formativa en el sentido que luego devolvamos, retroalimentemos a nuestros estudiantes acerca de cómo vemos que son sus avances. En algunos momentos del semestre, colocaremos una nota. La tendencia que tenemos ahora es de no colocar una sola nota o dos durante el semestre, sino dar más oportunidades y tener una cierta visión. Tampoco colocaremos diez notas, habrá que encontrar allí un equilibrio. Yo diría que la evaluación con calificación tendrá que ser aquella adecuada, viable de realizar, pero correcta para reflejar con calificación todos aquellos aprendizajes que es importante reflejar con la calificación. Y la evaluación con intención formativa tendrá que ser (ojalá) toda la que realicemos durante el semestre. Entonces, yo no lo vería como una proporción, sino como una intención permanente de que nuestra evaluación cumpla también una función formativa.

■ MARÍA TERESA MORENO

Gracias. Pregunta dirigida a Dieudonné Leclercq: “Una vez implementado un plan de estudios por competencias, ¿cómo se logra que los docentes que no son pedagogos se capaciten en didáctica y en evaluación acordes al nuevo enfoque?”.

■ DIEUDONNÉ LECLERCO

Pienso que actuando. Me parece que el ideal, para los que son voluntarios, es que puedan comprometerse en dar un paso, un paso significativo en la dirección. Lo importante es que un profesor dé un paso. Puede ser adelante en esta dirección, que sea en términos de evaluación, que sea en términos de método y en términos de la visión clara de cuáles son los retos, cuáles son las competencias que queremos... un paso. Me gustaría que en cada universidad, cada año, cada profesor pueda decir: "Este es el paso que en este año di"; el año siguiente: "El segundo paso que di". Porque necesita tiempo, necesita reflexión. Y nosotros, como profesores, tenemos que esperar al año siguiente para hacer los cambios; porque, cuando hemos dicho las reglas, las leyes de este año, debemos seguir aplicándolas durante todo el año, no podemos cambiar las reglas del juego a la mitad.

■ MARÍA TERESA MORENO

Pregunta dirigida al profesor Álvaro Cabrera: "Para la enseñanza de las ciencias básicas, con gran enfoque en la innovación y en la investigación, en vuestra experiencia, ¿qué criterios se pueden considerar para un sílabo por competencias?".

■ ÁLVARO CABRERA

Si estamos hablando de ciencia básica, se me ocurre que tal vez la competencia más impor-

tante o una competencia fundamental que podríamos intentar visualizar desde el inicio de la formación es la capacidad de aplicar el método científico y eso lo podemos entender perfectamente como una competencia. La capacidad de observar, de formular una hipótesis, de hacer un diseño experimental para testear esa hipótesis, de redactar los resultados de una manera que sea comunicable, de comunicar esos resultados, es un desempeño complejo que podríamos intentar ir desarrollando en los estudiantes en paralelo a todo aquello que consideramos saberes fundamentales que ellos también tienen que ir adquiriendo. Conversábamos ayer que en la concepción de competencias que estamos manejando no existe competencia sin ciertos contenidos, ¿verdad? No es opuesto a los contenidos. En mi visión —y esto puede ser un poco controvertido, hemos tenido discusiones con otros colegas—, el estar formando por competencias y el estar generando formas de evaluar esta actuación compleja no nos exime como profesores de la responsabilidad de también evaluar que ciertos saberes fundamentales hayan sido adquiridos por los estudiantes. Por eso también la variedad de instrumentos. Podría tener, a lo mejor, instrumentos que me permitan ver si los estudiantes efectivamente conocen ciertas cuestiones fundamentales. La diferencia con lo que hacemos ahora es que decimos: "Esto es válido, pero no es suficiente". También tengo que tener otro tipo de instrumentos que me entreguen evidencias de que el estudiante, por ejemplo,

va comprendiendo el método científico y va adquiriendo la capacidad de desenvolverse en él, de aplicarlo de manera autónoma, más guiada al principio, con más autonomía después, en un ámbito como, por ejemplo, el de las ciencias básicas. Por lo tanto, ahora nuestra responsabilidad es—una responsabilidad grande— evaluar tanto que el estudiante vaya siendo capaz de estas cuestiones complejas, como también que sepa y conozca lo que tiene que conocer, porque son recursos que deberá movilizar en esta actuación compleja; ya que sin ellos no va a haber una actuación compleja que sea competente.

■ MARÍA TERESA MORENO

Patricia, preguntan: “¿Cuál es el nivel de competencia que debe lograr un egresado de la educación superior? ¿Cuándo se determina que es competente?”. De pronto, si lo puedes relacionar con los niveles de progresión de competencias, ¿cómo afrontarías esta pregunta?

■ PATRICIA ESCOBAR

Yo les comentaba que nosotros venimos aplicando ya desde el año 2013, en la Facultad de Educación, un currículo por competencias y no ha sido fácil, sino todo un reto. ¿Por qué? Porque nuestra primera promoción recién egresa el próximo año. Lo que nosotros hemos venido haciendo es ir viendo cómo estas competencias

se están logrando paulatinamente haciendo algunos cortes. Es decir, si nosotros pensamos que al cuarto ciclo los alumnos deberían haber logrado ciertas competencias, deberíamos ir evaluándolas, ¿no? ¿De qué manera se están logrando estas competencias? Y luego estamos pensando —porque todavía no hay una promoción de egreso— hacer la evaluación al cierre. Esto también nos ha hecho ver que, cuando hemos hecho esta mirada, en las reuniones, decimos: “Bueno, hay cosas que hay que corregir”. Hay cosas que nos hemos dado cuenta que no están funcionando bien y hemos tenido que tomar medidas, rectificando algunos aspectos. Los profesores demandan también mirar el tema de la evaluación. Yo creo que este tema también es fundamental porque, siendo una facultad —incluso de Educación—, podemos ser muy innovadores en la metodología y en la enseñanza; pero todavía seguimos manteniendo en algunos aspectos los rezagos de una evaluación muy tradicional. Queremos cambiar eso y también tener un equipo o una persona dedicada a la evaluación, una gestora del currículo, que es la que va a ir midiendo, la que va a ir mirando cómo se está logrando este perfil. Nosotros, al finalizar cada ciclo, cada semestre, mandamos un informe en el cual tenemos que ver cómo se ha trabajado esos resultados de aprendizaje en función de esos desempeños y de esas competencias y la gestora del currículo lo que va haciendo es ir creando una matriz donde se va viendo cómo se está trabajando y se van logrando

estas competencias. Pero sí tiene que haber una evaluación gradual, cada cierto tiempo, porque no podemos esperar al final cuando decimos que ya no se puede hacer nada si no lograron estas competencias. Tiene que haber una evaluación progresiva.

■ MARÍA TERESA MORENO

Dieudonné, esta pregunta es para ti: "Considerando las diferentes formas de evaluación y estilos de aprendizaje, ¿se debe desaprobar al alumno? Si se desaprueba, ¿quién estuvo mal: el alumno o el profesor?".

■ DIEUDONNÉ LECLERCO

Pienso que los profesores deben hacer lo que deben hacer, pero no son los responsables de todo. El estudiante debe aportar su parte del trabajo. No podemos aprender en lugar del estudiante. No estamos en su cerebro. No estamos en su cronograma. Su gestión del tiempo es su problema, pero los estudiantes que son adultos tienen responsabilidad. El problema es que debemos darles oportunidad de saber hacia dónde vamos, de tener los instrumentos y los métodos para llegar ahí y también bastante retroalimentación para que sepan dónde están y si están lejos o cerca de esto o cuáles son sus debilidades. En la competencia, por ejemplo, de diagnóstico o una entrevista con un paciente médico, puede ser que al final del primer o segundo año, el estu-

dante sabe hacer 'a', 'b' o 'c'; pero no ha preguntado 'd'. Ha necesitado demasiado tiempo, ha sido demasiado invasivo inútilmente, y eso se debe mejorar. Tenemos la responsabilidad de dar a los alumnos bastante información para que la aprovechen y también hacer como una presión para que satisfagan criterios. Si no lo hacen y hemos hecho nuestra parte del trabajo, no somos responsables de todo lo malo del mundo.

■ MARÍA TERESA MORENO

Pregunta dirigida a Álvaro Cabrera: "Formar en competencia exige un trabajo de seguimiento sistemático, periódico y cercano —lo están citando un poco—, demanda una dedicación significativa en el marco de los cursos y del plan de estudios, ¿cómo equilibramos todas estas exigencias en el marco temporal del que disponemos?".

■ ÁLVARO CABRERA

Es la pregunta del millón de dólares. No hay una receta y nunca va a haber, creo yo. La gracia de tener profesores con la capacidad que tienen es que pueden ir generando distintas respuestas. Creo yo que la clásica respuesta que, en este caso, uno podría dar: "Bueno, tenemos que —lo voy a decir de manera muy bruta— enseñar menos contenidos", por decirlo de alguna manera. Pero eso no significa dejar de tratar de enseñar cuestiones que son fundamentales.

Entonces, hay tareas que son colectivas. Por ejemplo, todos fuimos estudiantes y todos vivimos ciertas duplicaciones dentro de un curso y otro, ¿verdad? Entonces, poder detectar aquellos puntos en que hay duplicaciones de contenidos y poder organizar al colectivo de profesores para decir: "Bueno, esto lo vas a tratar tú, lo vas a tratar tú, esto lo voy a ver yo" y que, por lo tanto, disminuyamos, a lo mejor, la cantidad de contenidos que cada uno debe hacerse responsable y luego escoger aquellos contenidos que son realmente fundacionales en la disciplina que queremos entregar, cuestiones sobre las cuales el resto del conocimiento va a ser construido por los estudiantes y enfocarnos en eso en profundidad. Recorrer, como decía antes, el camino del análisis de esos temas con los estudiantes y dejar algunos otros temas para un aprendizaje más autónomo. Ayer, Rafael nos mostraba también otras técnicas, a lo mejor más innovadoras o que nos cuestan un poco asumir. Por ejemplo, hay profesores de matemáticas que en lugar de dedicar su clase a resolver en la pizarra una serie de ecuaciones durante una hora y media, ahora hacen videos, tal y como lo harían en la clase, en un iPad; van dibujando, van explicando, lo suben a YouTube y el estudiante lo puede ver trescientas veces y, en la clase, lo que se hace es resolver problemas similares, equivalentes, colectivamente. Y el profesor se pasea por las distintas mesas de estudiantes apoyando, colaborando. Entonces, encontrar formas de despejar, a lo mejor, ciertos tiempos que tal

vez no son tan bien aprovechados en clase, utilizar otros soportes para esos tiempos fuertemente expositivos y utilizar la clase para un momento irremplazable. A veces tenemos estudiantes que dicen: "No voy a ir a esta clase porque leo el capítulo del libro y es lo mismo. Para qué voy a ir a clase". Bueno, lo que tenemos que hacer es que nuestra clase sea un momento irremplazable, un momento en donde efectivamente se discute, se analiza, nos apoyamos unos a otros y se construye ese aprendizaje. Claro, esto requiere de otras tareas nuevas y necesariamente nos va a poner en la labor, entonces, de ver qué cosas hacíamos antes que podemos despejar o realizar con otros soportes para darle cabida y tiempo a estas tareas de monitoreo permanente, sobre todo de retroalimentación, diría yo, entendiéndola como una actividad de aprendizaje central y no como algo que nos distrae de nuestra enseñanza, sino como parte de nuestra enseñanza, como una parte fundamental.

■ DIEUDONNÉ LECLERCO

Me parece que un concepto que utilizamos ahora en Europa es el de "créditos". ¿Ustedes lo conocen? La parte interesante de este, me parece, es su definición: mi curso tiene un peso de dos créditos. Significa que, si le doy a un profesor un curso de dos créditos, él puede utilizar 2×24 horas en las actividades del curso. De modo que hay una parte de esas horas que son clases. Vienen durante treinta

horas en mi curso y como este tiene un peso de dos veces veinticuatro, son dieciocho horas más. ¿Qué van a hacer? Leer, revisar. Y eso es muy importante porque, gracias a este peso, a esta medida de lo que es un crédito, tenemos interesantes diálogos con los estudiantes diciéndoles: “¿Cuántas horas dedica usted a mi curso, que son dos veces veinticuatro (cuarenta y ocho horas)?”. “¿Yo, a su curso? ¿Cien?”. “¿Qué haces? Leyendo el libro”. Y el profesor dice: “Leer este libro, yo que lo he escrito, te digo que lo puedes leer en tres horas”. Pero el problema fundamental de la formación universitaria es que el estudiante no es el profesor, el estudiante no ha escrito el libro. El estudiante normalmente tiene que pararse en cada palabra, tiene que buscar en el diccionario, hacer una llamada por teléfono, consultar, etc., para comprender. El profesor dice que para eso lo necesitan a él. A menudo dicen: “Este concepto lo enseñé en diez segundos. Yo puedo enseñarles álgebra en diez segundos; si hay una pizarra, lo haré”. El problema es que después de los diez segundos, todos los estudiantes fracasan, no comprenden, porque enseñar y aprender no es lo mismo. Y es muy interesante discutir con los estudiantes qué tiempo necesitan para comprender. Por supuesto, hay estudiantes que son más rápidos, que tienen más recursos y los que tienen facilidad aprenden más rápidamente. Pero este concepto de cuántas horas necesitan de trabajo, y el trabajo, como lo ha dicho Álvaro, puede ser de varias formas: ver videos, escribir informes, asistir y tener apuntes de una clase; pero

debemos tener eso un poquito, como si ustedes fuesen empleadores: hay un contrato con el trabajador. En Bélgica, no se puede pedir a un obrero que trabaje más de cincuenta horas por semana. Es lo mismo. Debemos tener un diálogo claro sobre las exigencias en términos de compromiso de tiempo.

Y también, la clase. Por ejemplo, a menudo son colegas como yo los que tienen una audiencia normalmente de cuatrocientos y hay cincuenta o sesenta presentes. Hay cuatrocientos en la lista. ¿Cómo hacer para que los estudiantes vengan? Les contesto. Tú debes concebir tu curso de manera que tenga un valor agregado. Si a los estudiantes les parece que ir y escuchar la clase no tiene ningún valor agregado, mejor que se queden en la cama.

■ MARÍA TERESA MORENO

Patricia, te hacen esta pregunta que creo va a generar una primera reflexión. Tiene que ver con la heterogeneidad de los estudiantes de Educación: “¿No debiera de estar también la valla alta a la entrada sabiendo finalmente que lo que deseamos son estudiantes de calidad? ¿Cuál es su reflexión?”.

■ PATRICIA ESCOBAR

Bueno, lamentablemente en nuestro país la carrera de Educación no es una carrera que sea socialmente reconocida; por lo tanto, muchos de los postulantes, cada vez menos creo yo,

suelen ingresar o postular a Educación pensando, no en querer entrar a esa carrera, sino porque piensan que, como para Educación se necesita menos puntaje, entonces esa es la vía para hacer el trampolín. Eso no es real porque luego, cuando ingresan, se dan cuenta de que tienen que tener ciento cincuenta créditos para poderse trasladar y son totalmente infelices siguiendo una carrera que no quieren, para la que no tienen vocación. Entonces, uno de los aspectos a trabajar a futuro obviamente es el tema de cómo nos sentimos como educadores. Creo que todos los que estamos aquí somos educadores, no de primera profesión, pero sí de segunda profesión, porque están enseñando, ¿cierto? Están enseñando a los futuros profesionales. Entonces, el tema de la valla es un tema que no lo vamos a poder evitar mientras el tema de la profesión esté reconocida como en otros países. En otros países, ser educador es ser casi igual que un médico y, por lo tanto, postulan diez, quince veces a seguir dicha carrera, lo cual no ocurre en nuestro país. Ahora, si eso lo asociamos al nivel de exigencia, este tiene que ser igual —llamando “exigencia” a qué es lo que nosotros trabajamos en las clases—. Tiene que ser igual para el profesional que es educador, para el profesional que es ingeniero y para el profesional que es médico, así como para los futuros profesionales. La enseñanza tiene que ser la misma, no tiene que variar. Y es un tema realmente de mucha complejidad. A nosotros, como educadores, sí nos llama a reflexión y nos agrada mucho, por ejemplo, que ya desde hace dos años tenemos promociones

de Beca Vocación de Maestro, que pueden ser alumnos que tengan algunas dificultades de adaptación a lo que es el mundo universitario, pero que tienen algo que es fundamental: la vocación, quieren ser verdaderamente educadores. Trabajar las falencias que pueden traer desde la escuela es un reto para nosotros porque implica mucha exigencia, mucho trabajo, mucha cercanía con los estudiantes, mucha tutoría también. La labor de tutor-profesor es fundamental en este acompañamiento, pero nos agrada mucho que son profesoras o futuros profesores en general, que tienen vocación de ser educadores. Entonces, para nosotros es fundamental tener la vocación y luego lo otro se va trabajando paulatinamente. Que es un reto, que es difícil, definitivamente, nada es sencillo y nada es fácil.

■ MARÍA TERESA MORENO

Y comprometidos todos con el tema de revalorizar la carrera docente. Dieudonné, ya es tal vez la última pregunta. Tú hablabas ayer de “diferenciación”. Tiene un poco que ver también con la pregunta que le han hecho a Patricia, aunque la de ella iba muy enmarcada en el ámbito de estudiantes que van a Educación. Saliendo de ese marco y extendiéndolo en general a los estudiantes de la universidad, la pregunta es: “Los alumnos no llevan las asignaturas en orden y los grupos que se conforman son heterogéneos respecto a las competencias que van logrando, ¿qué experiencias de evaluaciones diferenciadas o tratamiento

diferenciado de la enseñanza-aprendizaje conoces y qué implicaciones han tenido?”.

■ DIEUDONNÉ LECLERCO

¿Serviría adaptar la evaluación a cada estudiante, a cada grupo de estudiantes? No estoy seguro de haber entendido.

■ ÁLVARO CABRERA

No solo la evaluación, sino también el proceso de formación.

■ DIEUDONNÉ LECLERCO

En Teacher’s College, en Columbia, tenía un amigo —la mayoría de mis amigos han fallecido— que ha analizado la noción de aptitud y la ha definido como el tiempo que necesita cada persona para lograr un nivel dado. Por ejemplo, mi aptitud de aprender el piano es un fracaso (tengo uno), porque necesitaría años y siglos para aprender. Otros necesitan solamente cinco o seis años. Bien: aptitud. Es verdad que hay algunos que no tienen aptitud en nada y hay un reto por saber hasta dónde (o qué) debemos compensar. Una fórmula es preguntar a los que tienen la mayor aptitud para ayudar a los otros, pero es una visión del mundo. Si estamos en un mundo de competición y hay países y lugares en que son muy competitivos, que se matan. Voy a dar un ejemplo. En una universidad de Europa —no voy a decir cuál—, un profesor de Derecho había desarrollado pruebas de selec-

ción múltiple de modo que los estudiantes puedan entrenarse y había puesto en la biblioteca las respuestas correctas, de modo que los estudiantes pudiesen —era en los tiempos antes de las computadoras— ver cuáles eran las respuestas correctas. Y el colega observó que uno de los estudiantes robó las respuestas. Entonces, puso cadenas, de modo que no pudiesen robarlas y observó que uno cambió las respuestas correctas para que otros se equivocaran. Depende del mundo en el que estemos.

Benjamin Bloom (a quien ustedes conocen), de Chicago —que he conocido, he trabajado con él—ha promocionado la idea de éxito de clase. La clase tiene éxito si todo el mundo tiene éxito. Entonces, vamos a ayudar a los que lo necesitan, como en un equipo de fútbol. Lo ha llamado “la pedagogía de dominio”, pero en su pensamiento era un dominio colectivo. No basta que unos tengan el nivel de dominio, sino la mayoría y para eso se organizan. Te vas a ocupar de esto, vamos a hacer equipo. Depende del mundo en que queremos que aprendan y también que vivan después.

■ MARÍA TERESA MORENO

Bueno, llegados a este punto, quiero agradecer a nuestros invitados y, de manera muy especial, a Patricia Escobar, que se unió esta mañana, a Dieudonné Leclercq y a Álvaro Cabrera. Muchísimas gracias a los tres por todo lo que han compartido, toda su experiencia, además vivida y no solo teórica.

Segundo conversatorio: Sistemas de evaluación en la educación superior: evaluación del plan de estudios y los estudios de egresados

■ LEANDRO GARCÍA-CALDERÓN

En esta segunda parte del conversatorio, nos acompañan los profesores José Rafael López Islas, del Tecnológico de Monterrey; Juan Fernando Vega Ganoza, de la Pontificia Universidad Católica del Perú; y Sandro Paz, del Departamento de Ingeniería de esta casa de estudios. El profesor Paz es además directivo del SINEACE. También está con nosotros el profesor Óscar Pain, coordinador del área de evaluación académica de la dirección de asuntos académicos.

Dejo con ustedes al profesor Oscar Pain, quien va a conducir este segundo bloque del conversatorio de esta mañana.

■ ÓSCAR PAIN

Gracias, Leandro. Buenos días a todo el público asistente. Buenos días a Rafael,

Juan Fernando y Sandro que nos acompañan en este conversatorio. El día de ayer hemos conversado bastante sobre el tema de la evaluación de competencias desde el ámbito del aula, que corresponde, en términos técnicos, a la evaluación del nivel microcurricular. En el caso de Rafael, nos expuso temas asociados al nivel mesocurricular, específicamente ligados al seguimiento de los planes de estudio. Juan Fernando planteó cómo hacer seguimiento de los egresados de un carrera; es decir, la reflexión del perfil de egreso a un nivel que se denomina macrocurricular. Entonces, la primera pregunta, como para ir aclarando y volviéndonos a familiarizar con los temas y profundizar en ellos, sería: ¿cómo ven ustedes la articulación entre las evaluaciones que se hacen a los planes de estudio y los sistemas de egresados? ¿Cómo se comunican estos sistemas de evaluación?

■ RAFAEL LÓPEZ

Creo que la evaluación que se hace del programa para ver el cumplimiento del perfil de egreso es una parte muy importante de la evaluación de la efectividad que está teniendo la institución en ese programa en particular. La otra parte es el seguimiento de los egresados porque, como decía yo ayer, lo que estamos haciendo con una evaluación del programa es asegurarnos que, justo antes de que el estudiante vaya a salir de la universidad, realmente tenga la formación que nosotros nos comprometimos a darle. Pero, más adelante, nosotros ya no tenemos acceso a esta persona como estudiante, sino que se va al mercado laboral, a la vida. En el seguimiento a egresados realmente es donde podremos ver si esa persona, en su primer empleo, en el siguiente, en los próximos diez años, en los próximos veinticinco o cuarenta años está realmente desarrollándose para aquello que lo formamos. Entonces, creo que la evaluación que se hace a nivel del programa y después el seguimiento que se realiza por parte de la institución a los egresados son dos componentes fundamentales del aseguramiento de la calidad universitaria.

Creo que al principio, con la evaluación del programa, podemos ver la parte más formativa del corto plazo y la otra parte de largo plazo le tenemos que dar seguimiento a través de los otros estudios. Yo creo que son ambas cosas. En el caso particular de mi institución, todo eso forma parte del modelo. Lo que yo ayer compartí con ustedes fue obviamente

solo la parte de evaluación de programas, pero también tenemos la responsabilidad del seguimiento, el cual hacemos desde que están próximos a graduarse, a los tres meses, al año, a los cinco años, a los diez, quince, veinte, veinticinco años. Entonces, tenemos una información de desempeño de nuestros egresados a lo largo de por lo menos veinticinco años que nos permite ver si eso que estamos buscando se está cumpliendo no solamente al momento de graduarse, sino a largo plazo.

■ JUAN FERNANDO VEGA

Yo, por mi parte, ayer traté de hacer un acto de prestidigitación. Traté de mostrar cómo es que lo que estamos intentando hacer en el sistema de seguimiento de egresados en la PUCP apunta en la dirección del seguimiento de las competencias. Es un acto de prestidigitación porque mucho de lo que debemos hacer no lo estamos haciendo aún, lo que estábamos tratando es de marcar el territorio. Sobre esta demarcación del territorio, debo empezar declarando que Óscar se encarga de la evaluación de las competencias y yo me encargo del seguimiento de egresados. Sobre este último, quiero decir que estamos haciendo una parte, porque nuestro proceso de implementación de la formación por competencias ha comenzado después. Como les decía ayer, nuestro sistema de seguimiento de egresados comenzó, de alguna manera, antes y desfasado. Quisiera añadir un dato que ayer no pude desarrollar y que les anuncié que tenía y es que el sistema

de seguimiento de egresados no debería mirar solo ni principalmente a lo que piensan sobre las competencias los egresados. En el mundo hay ya otras experiencias y, por supuesto, una parte del sistema de seguimiento de egresados requiere que consultemos qué piensan de estos los empleadores. No acaba ahí la historia: también es bueno preguntar qué piensan los profesores de los egresados. Y sobre eso no tenemos experiencia local, pero sí hay experiencia norteamericana y lo que está documentado es que los profesores piensan que los egresados están súper calificados: los egresados piensan que tienen algunas fortalezas y los empleadores dicen que no. Entonces, este tipo de ajuste es otro que deberíamos procurar hacer.

El dato que les quería ofrecer —publicado en el libro de Cinda que les mencioné— es que, así como hay diversidad de especialidades (esto fue el *core* de mi intervención de ayer), hay diversidad de competencias y las competencias tienen distintas prioridades. Esto lo ha hecho un estudio del proyecto Alfa Tuning, que define veintisiete competencias genéricas y donde se puede distinguir unas competencias que son muy importantes o más importantes para el conjunto de las personas consultadas (empleadores, titulados, académicos y estudiantes) y descubrimos que compromiso ético, capacidad de aprender y actualizarse permanentemente, compromiso con la calidad, capacidad de aplicar conocimientos en la práctica, capacidad de identificar, plantear y resolver

problemas, capacidad de formular y gestionar proyectos, capacidad de abstracción, análisis y síntesis, capacidad de tomar decisiones y capacidad de trabajo en equipo, están arriba. El orden no es el mismo entre las distintas comunidades. ¿Qué está abajo? Capacidad de comunicación en segundo idioma, capacidad de investigación, habilidades interpersonales, capacidad de motivar y conducir hacia metas continuas, compromiso con la preservación del medioambiente, valoración y respeto de la diversidad y multiculturalidad, habilidad para trabajar en contextos internacionales, responsabilidad social y compromiso ciudadano y con su medio sociocultural.

¿Por qué señalo esto? Porque el registro no es suficiente. Al registro hay que hacerle un juicio y sobre ese juicio de diagnóstico hay que tomar decisiones, algunas de las cuales van a ser, bueno, lo siento, señores encuestados, están equivocados. Para nosotros es muy importante y quizá más importante la atención del ambiente, la atención de la multiculturalidad, lo cual genera una tensión.

■ ÓSCAR PAIN

Es un punto importante el que acaban de mencionar Fernando y Rafael. Quizá Sandro podría complementar. Al momento en que yo pregunto a empleadores, estudiantes y profesores sobre las competencias y viéndolo desde las profundidades que conversábamos ayer respecto a cómo evaluarlas, en general,

estamos más en un nivel superficial porque nos están reportando lo que ellos creen o qué tan satisfechos están con la formación recibida. Y aquello que tú comentabas, lo que los empleadores creen que es más importante, ¿Sandro, cómo has visto tú esta experiencia en ingeniería en el trabajo con los empleadores?

■ SANDRO PAZ

Justo quería, en esta bisagra, establecer puentes entre el perfil de egreso, cómo se hace el seguimiento del egresado y lo que se debe medir allí. Hay un actor muy importante que en el modelo de acreditación que tenemos en el SINEACE: los grupos de interés. Estos terminan siendo los empleadores, pero no solamente ellos, porque también pueden estar los mismos docentes o los colegios profesionales o alguna otra institución que pueda tener relevancia sobre la carrera o sobre la formación en particular. Lo interesante es que cada uno de ellos tiene una mirada diferente, cada uno de ellos hace un énfasis diferente en lo que quiere del egresado. La misma ley universitaria plantea aspectos como, por ejemplo, la investigación que es sumamente clave para la licenciatura y el bachillerato que a lo mejor los empresarios dicen: “Para qué investigar, yo quiero que el estudiante venga a resolver los problemas”. Y los padres de familia a lo mejor pueden decir: “Quiero una carrera corta porque no me va a costar tanto”, mientras que los docentes pueden decir: “Un ingeniero en menos de cinco

años, no termina, en tantos créditos no se puede formar a un ingeniero”.

Esta diferente mirada que tienen estos grupos de interés, a los cuales finalmente cada universidad debe acercarse, hace que las universidades sean diversas y que tengan enfoques diferentes, lo cual creo que es sumamente saludable para el desarrollo de un país: que no todas se dediquen a lo mismo. El ingeniero mecánico lógicamente tiene que saber de mecanismos y todos los ingenieros mecánicos tendrán que saberlo, pero hay matices que cada una de las universidades puede reenfocar en función de cómo se está enfocando en esos grupos de interés. Una de las cosas que, por ejemplo, vimos en las mediciones que hicimos de los objetivos educacionales en ingeniería cuando estábamos buscando la acreditación con CEAB (*Canadian Engineering Accreditation Board*) y luego con ABET (*Accreditation Board for Engineering and Technology*) es que, ante prácticamente las mismas preguntas, si es que los objetivos educacionales planteados para cada una de las ingenierías eran logrados por los estudiantes, estos —en una medición indirecta— sentían que no lo estaban logrando tan bien como ellos hubieran querido, como se les había ofrecido. Pero, por otro lado, los empleadores decían que sí, que los egresados estaban por encima de las expectativas que ellos tenían. Entonces, esa mirada se repitió en casi todas las carreras de ingeniería cuando se hizo este análisis, por lo menos la primera vez, en donde yo estuve participando más de cerca. ¿Cómo calzar esas

dos percepciones? Yo creo que no es una tarea sencilla.

■ RAFAEL LÓPEZ

Yo diría que en esto la misión de la institución juega un papel central porque define para qué existe la universidad y es en el momento de definirla en donde se escucha a estos grupos de interés y se dice: "Hay muchas cosas, muchas necesidades, muchos puntos de vista, pero nosotros decidimos qué vamos a enfocarnos en esto. Para eso fuimos creados, para eso existimos y ese es el compromiso que asumimos". Nuestro rector dice que elegir es renunciar. De esta lista enorme, no podemos desarrollar todas las competencias; entonces, si escogemos cuatro, cinco o seis de esas, seleccionamos esas y renunciamos a las otras. No porque no son importantes, ni porque no podamos o no queramos hacerlo, simplemente tenemos que ser realistas y enfocarnos en lo que es nuestro centro.

Entonces, creo que, en ese sentido, la misión como un eje estratégico para entender el quehacer de la universidad, para definir su planeación, para orientar los planes y programas educativos, me parece esencial. A veces pareciera que la misión es un documento que está ahí guardado o que está en un afiche por ahí, en una pared, y que ya está amarillo y un poco deteriorado por el paso de los años, pero porque se convirtió en letra muerta. Creo que

en ese sentido, esa parte, por lo menos para nosotros, ha sido muy importante.

■ ÓSCAR PAIN

Ese quizá sea un primer marco interesante de acuerdo o de ancla. Si todo sistema de evaluación o todo proceso de evaluación finalmente implica emitir un juicio de valor y tomar una decisión y tenemos estas distintas fuentes (de los grupos de interés, de la información que haya en los planes de estudio), este juicio de valor no puede estar en el aire. Es decir, no se toma sobre base de la nada y creo que una primera perspectiva sería esa: revisar nuevamente nuestra misión, hacia dónde vamos, y darle ese matiz a la hora de tomar decisiones entre los juicios de los distintos grupos. ¿Estarían de acuerdo?

■ LOSTRES

Sí, claro.

■ ÓSCAR PAIN

Podemos tomar decisiones desde ese punto de vista; pero no es exactamente lo mismo si esta información proviene de los distintos grupos de interés, que si proviene de los ciclos de mejora continua o de la evaluación de los planes de estudios. Es decir, la información que Sandro nos comentaba hace un momento, en la cual los estudiantes percibían que habían

desarrollado menos capacidades y, en cambio, los empleadores estaban contentos, nos lleva a la pregunta de cómo relaciono la información que proviene del seguimiento a los logros del plan de estudios con la información sobre el avance de los estudiantes.

■ RAFAEL LÓPEZ

Creo que nuevamente es un tema muy importante de entender. A fin de cuentas lo que estamos haciendo cuando generamos información es poner un insumo en la mesa, con el cual hay que trabajar. Los resultados, como decía Juan Fernando, los datos por sí mismos no tienen sentido, hay que dárselo. Me parece que esa parte, a veces, se deja como algo obvio cuando en realidad se tiene que hacer. Entonces, una vez que se genera información, hay que darle sentido, hay que ponerle el contexto y, a fin de cuentas, es una propuesta. Lo que tú planteas como una conclusión es una propuesta, un argumento que haces. Nosotros creemos que vamos bien, creemos que nos falta esto o que nos está saliendo bien esto otro, que tendríamos que hacer algunos cambios. Es un argumento, no es una verdad absoluta, sino un trabajo de argumentación. Los padres de familia te dicen algo, los alumnos te dicen algo, los empleadores, tus datos te arrojan algo; entonces hay que ponerlo sobre la mesa, discutirlo. Somos académicos, esto supone que se nos da bien entendernos, analizar y poder entonces sacar conclusiones.

En ese sentido, ver este proceso en el contexto de ciclos de mejora me parece que

ayuda mucho porque es bueno; concretamente, de todo esto que estamos observando, ver qué podemos hacer para movernos. Y a lo mejor esa decisión no es la mejor, a lo mejor nos estamos equivocando, pero nos estamos moviendo de buena fe y me parece que eso es lo que se espera. Es cierto, a veces quisiéramos tener un plan maestro de evaluación antes de tomar una decisión: se quisiera un análisis de impacto a gran escala; pero, ¿sobre qué impacto, si todavía no actúas? Actúa y entonces mides el impacto. Me parece que tenemos que verlo. Y, en ese sentido, los ciclos de mejora ayudan muchísimo a ir avanzando, a ir viendo el rumbo. Es un poco como decir: “Nos vamos a meter a un bosque que no tiene rutas marcadas, pues tengo que ir abriendo brecha”, y a lo mejor de repente digo: “Aquí ya llegué a un barranco, mejor me voy para atrás y ni modo”. Pero hay que ir caminando, no esperar a que lleguen. Poden todos los árboles y entonces digan: “Ahora sí, ahora ya camino sobre seguro”.

■ JUAN FERNANDO VEGA

Yo quisiera retomar la idea de que hay que ajustar en función del norte. Y tenemos en el afiche del evento una rosa de vientos que hemos usado poco y que ahora pensaba que podríamos haberla usado un poco más. Yo, entonces, tratando de usar la rosa de vientos, diría que los sistemas de seguimiento de egresados tienen que incorporar más “quiénes” para también definir otros “cómo” y “cuándo”. ¿A qué me

refiero? Una de las cosas que les mostré ayer como el hallazgo principal de nuestra reflexión respecto a la utilidad que podría tener lo que estamos haciendo para evaluar competencias, es que no diseñamos el sistema para evaluar competencias, sino para tener un tablero de mando. Estamos en el segundo nivel de la descripción de Dieudonné y Álvaro sobre en qué consiste esto de gestionar las competencias. Queríamos información para que el piloto pudiera tomar decisiones. Entonces la investigación era sobre el mercado, sobre cómo se sentían los alumnos respecto de algunas competencias; incluso habría que declarar que teníamos una lista de competencias en 2010 y ya la hemos cambiado. Y ese cambio va a tomar tiempo en que lo incorporemos en el sistema de seguimiento de competencias, porque los cambios en la estructura de las preguntas hace la información no comparable. Pero entonces el principal resultado es que tenemos diferentes especialidades y hay que dar cuenta de información a ellas.

Ese ha sido el norte del proceso que hemos seguido y eso ha significado meterse en problemas cada vez más complicados, porque no podíamos reportar para las especialidades pequeñas, ya que no teníamos un número suficiente de casos. Además, hemos tratado de entender cómo es que había algunas regularidades que teníamos la intuición de que existían, pero de las que no teníamos evidencias y entonces hemos construido análisis de comunalidades, constatando que hay diferentes demandas. Los ingenieros no demandan lo mismo que los

literatos o los matemáticos o los químicos o los físicos. Están como en diferentes juegos. ¿Qué es lo que dice Becher²? Todos los especialistas sobre disciplina dicen: “No hay la disciplina, hay las disciplinas”. Como yo suelo decir, tampoco hay la universidad, hay las universidades. Yo tengo una opinión muy fuerte sobre cómo es que exigir investigación en todas las universidades es una tontería; pero hay competencias de investigación que se tienen que desarrollar ahora para todo el mundo y no está cantado que dicha competencia pertenezca al juego de algunas especialidades.

■ ÓSCAR PAIN

Con lo cual el tema de cómo ajustamos esos ciclos de mejora se vuelve más relevante, porque va a haber una serie de cambios y de ventanas de tiempo en las cuales tengo que introducir y adaptarme a estas distintas demandas. ¿Cómo lo has visto esto tú, Sandro, en el caso de las carreras de ingeniería?

■ SANDRO PAZ

La experiencia que he podido vivir es que muchas veces toda esta información que se recoge —ya sea de las mediciones que se

² Nota del editor: Juan Fernando Vega hace referencia el libro de Becher titulado *Tribus y territorios académicos: la indagación intelectual y las culturas de las disciplinas* (Barcelona, Gedisa, 2001).

hacen directamente en las competencias que va logrando el egresado evaluando el perfil de egreso, digamos, y toda la información que se puede recibir de los estudios de egresados—, muchas veces termina siendo información que los profesores la ponen a un lado y plantean mejoras sin necesariamente seguir un análisis de aquella información. Y entonces varias veces nos ha tocado decir: “Perdón, profesores, aquí tenemos toda esta información”.

¿Cómo la analizamos? Tenemos que hacerlo sobre la base de lo que nos dicen otras personas y obviamente nuestro juicio y la experiencia que tenemos como docentes. Pero muchas veces el profesor cree que él es el que más conoce de esto, que la pregunta que se les planteó a los empleadores estaba mal hecha y que la medición que hicieron en verdad no debió ser al final de este curso, sino al final de este otro y por eso es que salen esos resultados. Muchas veces nos ha tocado enfrentarnos a los profesores al decirles: “Por favor, analicemos la información que recibimos”. Sí, seamos críticos con ella, pero usemos esa información para mejorar. También tendremos soluciones que vienen de nuestra propia experiencia, pero toda esta información es valiosísima. Y ahí hay un tema que mencionaste, Juan Fernando, retomando el de los pequeños grupos de promociones, a veces con los cuales es complicado establecer conclusiones representativas por el tamaño de muestra y todo esto; pero termina siendo información válida que la carrera, considerando que no es

una muestra representativa, puede analizarla y establecer conclusiones. O sea, todo termina alimentando estos ciclos de mejora.

Algo que siempre tuve duda es con qué frecuencia se deben hacer estos ciclos de mejora. Si debe ser algo estructurado—es decir, todos los febreros nos reunimos para plantear cuáles van a ser las actividades de mejora para el siguiente año o dos veces por año—o debería ser algo natural, frecuente, cada mes o cada semana, o que los profesores, sin ninguna estructuración sino, digamos, naturalmente, puedan ir contribuyendo con los planes de mejora en cualquier momento. Y creo que esta segunda opción es la más conveniente.

■ RAFAEL LÓPEZ

Yo quisiera comentar aquí algo basado también en la experiencia que me toca vivir trabajando obviamente con profesores y directores de programa, pero también con la rectoría, incluso con los consejeros. Se relaciona con el tema de los ciclos de mejora y la periodicidad. Cuando trabajamos en la educación, el efecto de lo que hacemos es de largo plazo. Y ahí, a veces, existe poca sensibilidad en el sentido de que se espera un resultado, un cambio que se note de manera inmediata. Si decidimos que vamos a hacer un nuevo currículum, nos toma uno o dos años diseñarlo. Imagínense: lo decidimos hoy, en 2018 empezamos a impartirlo, en 2023 se gradúan los primeros, en 2024-2025 podríamos ir viendo ya, más o menos, cómo les está yendo

allá fuera. Pero la evaluación que te van a pedir que hagas del nuevo plan de estudios te la van a pedir ya para febrero de 2017, cuando apenas estás empezando a ver quién va a trabajar en esto, y no digamos en febrero de 2019 cuando ya pasaron seis meses con el nuevo currículo. Por qué no tenemos tantas competencias, por qué a nuestros egresados que se graduaron hace diez años no se le notan las competencias; pues porque esas criaturas se graduaron antes de que habláramos en la universidad de las competencias. Entonces, ese efecto de largo plazo, pues hay que vivir con él.

Yo sé que muchas de las cosas que estoy haciendo actualmente no las voy a ver. O sea, espero sí verlas, desde el asilo (risas); no pienso morirme pronto, pero voy a estar seguramente retirado de mi institución cuando muchas de estas cosas estén entrando en vigor. Tenemos que tener esa paciencia y por eso los ciclos de mejora me parece que son importantes como una actividad continua. Cuando digo ciclos de mejora, me refiero a algo continuo, como se dijo antes, que tiene que fluir de forma natural, lo cual ayuda obviamente a tener cierta periodicidad que la marcan los planes de evaluación. Si dices: “Aplicamos una evaluación”, inmediatamente, cuando se tenga la información, convocamos a una reunión y revisamos para ver qué cosas podemos ajustar, porque a fin de cuentas esa sí es una respuesta inmediata sobre la base de la información de corto plazo que tenemos, asumiendo que, si vamos haciendo las correcciones a lo largo

del camino, podemos llegar en el mediano o largo plazo a tener un mejor impacto. Esa es una realidad con la cual tenemos que vivir en la educación.

■ ÓSCAR PAIN

Juan Fernando, querías complementar la idea.

■ JUAN FERNANDO VEGA

Rafael me ha dado la entrada a una metáfora que a mí me parece muy potente y la usé en otra presentación, cuyo título era: “¿Qué puede aprender la educación superior de la balística?”. ¿En qué consiste la balística? Es aquella ciencia que estudia la trayectoria de un proyectil. La descripción que ha hecho Rafael de cuánto tiempo toma desarrollar un proceso antes de tener resultados estables es importante de tener en cuenta; porque, además, la primera generación no es la normal, hay que esperar que se gradúe o se regule. Entonces estamos hablando del año 2030, más o menos.

■ SANDRO PAZ

Cuando ya es necesario hacer otro cambio.

■ JUAN FERNANDO VEGA

Entonces, lo que estamos haciendo es, al viejo estilo, lanzar un obús lejos a ver si llegamos. El problema es que el mundo ha cambiado. Ese era el sistema de balística para cañones fijos y

edificaciones que no se movían. ¿Pero qué pasa cuando la edificación ya no está fija, sino que se mueve? Ahí se creó un problema. ¿Y cuando se mueven los dos? ¿Y cuándo sube la velocidad? Imagínense la dificultad de abatir un caza de última generación. Ya nadie dispara un obús, lanzas un misil que tiene sistemas de autocorrección. A veces nos representamos la educación como el obús, para diseñar hoy, mirando la experiencia de hace cinco años, un obús para que le dé al blanco dentro de... ¿quince? Necesitamos diseñar un misil que incluya procedimientos de autocorrección. Ahí está el tema del tamaño de los ciclos de realimentación. ¿Cómo se comunica eso con los sistemas de seguimiento de egresados? Habrá que ver. No se me ocurre todavía, pero hacia ahí es donde tenemos que apuntar, porque de lo contrario no la logramos. Vamos a tener sistemas cada vez más complejos que van a ver la diversidad y van a poder decirle a los de la especialidad de matemática que sus dos matemáticos que se graduaron el año pasado “opinan que la matemática...”, pero no van a tener la capacidad de mirar lo suficientemente lejos y tener retroalimentación lo suficientemente rápido.

■ ÓSCAR PAIN

Todo sistema automático de redireccionamiento, finalmente, le da información a alguien que tiene que tomar una decisión. Un aspecto que quizá se puede perder de vista es que quienes solemos producir información, por nuestra sobrecarga en el proceso mismo

de producir informes, enviarlos a los departamentos, a los especialistas, etc., podemos perder con facilidad el seguimiento a los que se deberían generar para que los actores se empoderen de esa información.

Si no tenemos hitos en los ciclos de evaluación, si no tenemos un corte de tres o seis años, entonces no se genera un espacio o, por lo menos, un compromiso en el corto plazo con los resultados observados, como decía Rafael. Y quizá por ahí es que uno debería tratar de generar esos espacios para poder tomar esas decisiones. Todo esto es difícil porque, si los empleadores y la tecnología empiezan a avanzar tan rápidamente y esto afecta tanto a las ciencias e ingeniería como a las humanidades, si la cantidad de información y *papers* que uno encuentra en bases de datos es increíblemente difícil de procesar para cualquiera, entonces qué decisiones y con qué rapidez se deberán tomar durante estos ciclos de evaluación. Lo que tú proponías en el TEC (a Rafael) y lo que propone también el sistema de ABET es centrarse en los cursos, priorizar ciertas competencias, ciertos resultados del estudiante y eso hacerlo cíclico en un espacio de seis años. No siempre estás midiéndolo todo. ¿Qué ventajas y desventajas tiene esto en la vida práctica?

■ SANDRO PAZ

Una de las cosas que puede traer no muy buenas consecuencias es el hecho de estar

haciendo mediciones en periodos muy cortos, porque eso da la tentación de hacer cambios tan pronto ocurren algunos puntos extraños en las mediciones que se pueden hacer del logro de las competencias. Por eso, una de las cosas que siempre estuvo en discusión es cada cuánto tiempo se debe medir el logro de una competencia, si es que se debe hacer en todas las promociones, si se debe hacer cada dos promociones o cada año. Entonces, una de las cosas que nos ha preocupado siempre cuando diseñamos el sistema de medición de los *outcomes* para ABET —y algo similar que hay para CEAB—, es cuál es la rúbrica adecuada, cuál es el curso adecuado donde se debe hacer esta medición directa y con qué frecuencia. Porque, si es que se hace, como decía, de manera muy consecutiva, al momento de hacer ese análisis, de buscar ese espacio para que los profesores analicen, tan pronto ven algo preocupante en un curso extraño, dicen: “Tenemos que hacer algo con ese curso, tenemos que hacer algo con ese profesor, tenemos que hacer algo con esta materia”, sin que necesariamente la acción tenga efecto en la siguiente medición, sino que probablemente se requieran dos o tres mediciones más. Creo que una de las cosas que es sumamente clave para este proceso de mejora es recoger información, que esa información sea la oportuna, en el momento adecuado, y que en general, como decía hace un momento, los profesores usemos esa información para tomar decisiones de mejora.

■ RAFAEL LÓPEZ

o creo que, en ese sentido, a veces, si en una evaluación se logran buenos resultados, posiblemente no sea necesario mantener una frecuencia muy alta. Quizá en algunos puntos donde pudimos haber identificado alguna debilidad notoria vamos a tomar acciones y posiblemente decidamos tener una mayor frecuencia para ver si las acciones que estamos tomando nos van a llevar en esa dirección. Creo que es la flexibilidad con la que se tiene trabajar y entonces yo comparto esa noción que decía Sandro hace un momento de que sea un ciclo natural, más que algo planificado, que se lleve con flexibilidad para ir respondiendo conforme lo requiera el problema.

■ JUAN FERNANDO VEGA

Me quedé pensando en que quizá no estaba poniendo suficiente énfasis en la importancia del cambio en el entorno que tenemos que enfrentar. ¿Conocen a Watson? Búsquenlo en Google: IBM Watson. IBM ha diseñado un sistema que es de inteligencia artificial, en el viejo sentido de la palabra; pero que IBM ya no lo denomina así, sino de “inteligencia aumentada”. Watson es una computadora con un tipo de programa que realiza cierto aprendizaje (*machine learn*) y que en este momento hace trabajo humano. Hay abogados que trabajan con Watson para mejorar su capacidad de leer la increíble profusión de reportes legales, leyes, cruces, para poder tener la mejor opción

al momento de atender al cliente. Sobre Watson se ha creado plataformas de diagnóstico médico que ya superan a los médicos expertos en eficacia en el diagnóstico de algunos cánceres. Solemos hablar de realidad aumentada y eso es el Pokémon. Bueno, una de las competencias con inteligencia aumentada que necesitamos ya porque está funcionando sistema en todo el mundo. ¿Hay alguien que haya incorporado esa competencia?

■ ÓSCAR PAIN

En mi clase, por lo menos la mitad de estudiantes, juegan Pokémon.

■ JUAN FERNANDO VEGA

Esto es realidad aumentada, no inteligencia aumentada. ¿Cuántos son capaces de lidiar con Watson? Watson además ha derrotado al experto en ajedrez y al experto en Go. El futuro ya no es como antes.

■ ÓSCAR PAIN

En esa línea de la tecnología —y pasando a las preguntas del público, que son cuantiosas—, hay varias preguntas para Rafael acerca de cómo es que el Tec de Monterrey ha incorporado todos estos cambios de tecnología en su trabajo y en la mejora de los planes de estudios, en la demanda tecnológica de los profesionales egresados.

■ RAFAEL LÓPEZ

Creo que son varios aspectos. Uno muy particular tiene que ver con el sistema que platicábamos ayer. Es un sistema que se hizo pensando en cómo dar seguimiento a toda la información que se genera en un proceso de evaluación, cuando tienes que evaluar a todos los programas de estudio y sobre todo con tantos campos como tenemos nosotros. Entonces, desarrollamos un sistema al que pusimos como condición de entrada que se conectara con los sistemas de administración escolar existentes, de manera que pudiéramos tener toda la información y que esta pudiera fluir en ambos sentidos, permitiendo conectar los datos de estudiantes, profesores y grupos con este sistema que llamamos SAEP (Sistema para Administrar la Evaluación de Programas Académicos).

Esta conexión permite simplificar muchísimo el trabajo porque ya no se tiene que hacer cargas de información; entonces, si yo voy a hacer una evaluación de los resultados, simplemente me doy de alta. En el sistema está registrado que este semestre, por el plan de evaluación, me toca evaluar cierto curso o, sobre la base de cierto curso, me aparecen los datos de los estudiantes, sus trabajos, la rúbrica. Obviamente, el trabajo intelectual de evaluar lo tomo yo, pero toda la información está ahí y, al momento que doy guardar, la información queda almacenada y después se puede generar los reportes. En ese sentido, el sistema ha sido muy importante y por eso también hemos

tenido dificultad para encontrar una versión que esté en el mercado que pudiera sustituirlo. Existen sistemas muy poderosos, pero necesitamos evaluar el costo-beneficio de quedarnos con este con las inversiones adicionales que hay que hacer o migrar a otro sistema. Eso por el lado de apoyo informático para este proceso que es importante para poderlo administrar. De otra manera, sería un trabajo humano enorme, nuestro equipo es muy pequeño y en buena medida podemos hacer esto porque tenemos la información toda sistematizada con este gran proyecto informático.

Por otro lado, ayer hacía referencia también al impacto que ha tenido todo lo que hemos estado haciendo en términos de cambios en la forma de enseñar, demandas para que nuestros salones y en general cualquier espacio de nuestro campus esté habilitado con acceso tecnológico a Internet, etcétera. Esto sí es importante porque, si nosotros le decimos al alumno que la expectativa es que cuando venga al salón haya visto materiales en Youtube, tenemos que abrirle canales, tenemos que darle posibilidades de que lo haga y al profesor también, tenemos que equiparlo tecnológicamente, pero eso es una inversión más que un costo regular. O sea, es parte del entorno tecnológico porque, si necesitamos personas, jóvenes que en pocos años tengan que lidiar con esa inteligencia aumentada, necesitamos tener ese equipamiento, de otra manera no se va a poder. En ese sentido, la tecnología también ha sido un aliado muy

importante. Tenemos una gran experiencia trabajando en modelos de educación a distancia que ahora hemos cambiado para que sea un modelo más híbrido, estos *blended learning* en donde el profesor graba cosas, se pone en repositorios para que se pueda tener acceso a la información. Entonces, hay una experiencia por ese lado que también está sobre la mesa y que va en apoyo de lo que se está haciendo.

■ ÓSCAR PAIN

Sandro, tú querías complementar.

■ SANDRO PAZ

Más que complementar, una pregunta: ¿estos cambios tecnológicos que han habido en la universidad, vienen por una decisión política, por una mirada en la visión de la institución o porque los profesores comienzan a empujar y es necesario que ocurran estos cambios?

■ RAFAEL LÓPEZ

Yo creo que son ambas vertientes. Hay algunos aspectos que se derivan de la propia planeación, de la propia forma de concebir la visión y cómo la debemos de poder implementar; pero también hay una demanda muy fuerte por parte de los profesores acerca del tipo de equipamiento que se requiere para poder trabajar en el tipo de espacios que se están buscando. Confluyen ambos aspectos.

Lo que sí está muy claro es que la opinión de los estudiantes, no tanto de los profesores, me parece que jamás la vamos a satisfacer. Hemos hecho a veces cambios en los que, de un día para otro, se instala un nuevo sistema y se incrementa en diez veces el ancho de banda y ese mismo día por la tarde llegan las quejas de los alumnos de que es muy lenta la red. Esto es algo que no tiene fin. Entonces, el director de informática se jalaba el pelo, de hecho ya le queda poco por eso (risas), porque dice: “Cómo va a ser posible si ayer teníamos una décima parte de capacidad y llegan las quejas”. ¿Por qué? Pues porque hay más demanda y hay más inteligencia—quisiéramos pensar que inteligencia aumentada— que requiere grandes cantidades de información. La demanda está ahí. Ese es un poco el tema.

■ ÓSCAR PAIN

Aquí hay otra pregunta para Rafael, pero la voy a poner para el panel en general. Se han quedado asombrados varios asistentes en la capacidad que tiene el Tec para seguir a los egresados hasta por veinticinco años. La pregunta es: ¿qué hacen con esa información cuando ustedes encuentran que sus egresados, siguiéndolos a lo largo del tiempo, tienen algunas falencias o dificultades? ¿Solo sirve para mejorar el sistema actual? ¿Ofrecen algún tipo de formación continua? ¿De alguna manera es eso responsabilidad o no de la universidad?

■ RAFAEL LÓPEZ

Es un punto muy importante porque es un riesgo que se corre. Si busco a un egresado de hace veinticinco años y le pregunto cómo le va y resulta que le va muy mal, entonces ¿qué tendría que hacer? ¿Le invito que venga gratis por un reemplazo porque el producto salió mal o qué hago con ellos?. Es un riesgo que se corre por andar preguntando. ¿Para qué preguntas si no quieres conocer la respuesta? Creo que a todo el mundo le queda claro que —lo decía ayer— la universidad forma y lo hace para el largo plazo, no solamente para los primeros tres meses; entonces, en ese sentido, lo que hacemos como seguimiento es realmente ver o asegurarnos de que esa misión se está dando en esa línea. Por ejemplo, los aspectos claves de nuestra misión: el liderazgo. Para nosotros es muy importante ver que nuestros egresados a los cinco, a los diez, a los quince, a los veinte, a los veinticinco años realmente ocupan puestos o funciones—no necesariamente puestos— de liderazgo, que tienen sentido humano, que trabajan en pro de la comunidad.

Ese tipo de evidencias son las que buscamos y obviamente recibimos también algunas críticas de los egresados diciendo: “Quisiéramos que fuera mejor”. No podemos actuar ya con ellos, pero sí podemos actuar con los alumnos actuales para evitar que ese tipo de cosas que estamos viendo se sigan dando. Generalmente son cosas que ya atendimos, pero por ese rezago no se alcanza a ver. A esas personas no les tocó. Me parece que en términos de

“cómo hacemos para tener ese seguimiento”, yo diría que lo que planteaba Juan Fernando ayer es muy parecido a lo estábamos haciendo nosotros allá por los años noventa. Empezamos de una manera similar. Lo escuchaba y decía: “Esta historia, película ya la vi”, pues de repente te das cuenta de que los muestreos no te sirven con poblaciones pequeñas. Esa parte la vivimos porque así iniciamos y después, igual que ustedes, hemos trabajado muy de cerca con el área que se encarga de dar seguimiento por el lado de la relación con los egresados, no el seguimiento académico o de investigación. Y lo hacemos también así para no ser juez y parte, para separar la parte analítica de la parte de vinculación. Estamos trabajando de manera muy sincronizada porque identificamos a alguien que no está en su base de datos, la compartimos y al revés.

Eso nos ha ayudado bastante. Pero también me parece que algo muy importante —y este es el punto que yo veo medular y como una visión a futuro— es la mayor vinculación de nuestros egresados, no solamente dando dinero, que es la forma tradicional que muchas universidades ven la vinculación del egresado con su universidad, sino cómo se vincula en el aprendizaje y ahí es donde veo un gran potencial. Estamos teniendo cada vez más proyectos en donde preguntamos: “¿Cuál va a ser esa situación compleja a la que queremos exponer al alumno, como parte de la formación?”. Tenemos egresados que están haciendo cosas interesantísimas afuera, tocamos base con ellos, y quieren ser tutores y estar pendien-

tes; así, regresan a la universidad o reciben a nuestros alumnos y nuestros profesores en sus organizaciones y se trabaja de manera continua. Entonces, creo que de repente ellos son los mejores tutores que podemos tener para nuestros alumnos y creo que es una relación que me parece que fortalece el vínculo de la universidad, no solamente con la persona durante los cuatro o cinco años que es nuestro estudiante, sino a lo largo de toda la vida.

■ ÓSCAR PAIN

Antes de otra intervención, voy a agregar a la misma pregunta tres más en combo que tienen que ver con el sistema de egresados: ¿cómo hacemos esto: institucional o por carreras?, ¿en cuánto tiempo?, ¿cuánto espaciamos las cohortes o seguimos a todas?, ¿y si introdujéramos certificaciones periódicas, eso también nos daría información para el seguimiento de egresados?

■ JUAN FERNANDO VEGA

Quizá en un plazo nosotros vamos a estar siguiendo a veinticinco años, pero ahora estamos siguiendo hasta cinco años. Y la idea de quedarnos ahí, de hacer solamente dos cortes, es que a medida que pasa el tiempo la importancia de lo que nosotros hacemos en la universidad va siendo sustituida por la importancia de lo que ellos hacen desde afuera, en su experiencia de trabajo. Creemos que con cinco años podemos tener una idea clara sobre

de qué manera nosotros hemos influido este aterrizaje.

La siguiente cosa que quiero declarar es que nosotros hemos decidido hacerlo continuo. Censal y continuo para poder acumular. Porque necesitamos retroalimentar a las especialidades. Los núcleos vivos de la formación y los núcleos vivos de la investigación dentro de la universidad no son las facultades, son las especialidades. ¿Qué puede representar un promedio de facultad para que incluya a psicología, literatura, historia y geografía? Es poco probable que esto les permita retroalimentar a los de geografía, sobre todo si, demográficamente y para hacerlo representativo de la facultad, hemos tomado proporcional a la población, entonces, la mayoría son psicólogos. El promedio está corrido a los psicólogos. Esa es la segunda idea respecto al sistema de seguimiento de egresados.

La tercera idea tiene que ver con algo que no quisiera dejar de decir, que se vincula con una cosa que dijo Álvaro: todos sabemos que hay alumnos brillantes que, si los guardamos en un closet, cuando salgan un año después van a saber más. No nos necesitan. Esa era la vieja educación. Nosotros tenemos un paradigma que no se mueve, que se resiste a morir, que es el de la universidad elitista. Ese paradigma se agotó. No podemos garantizar que todos logran lo mismo. No se puede y sería tonto. Lo que tenemos que garantizar es que la varianza no sea tan grande y que además haya un corte, de tal manera que no graduemos a gente que

no tiene las capacidades. Ese es el esfuerzo en el que estamos involucrados. Para eso se necesita mucho más en el sistema de seguimiento de egresados que preguntarles si en el trabajo les demandan la competencia o si sienten que están preparados, porque esta opinión además nosotros sabemos que cambia en el tiempo: a los dos años dicen que esto era un desastre y a los cinco años dicen: "Qué bien, la hicieron".

■ SANDRO PAZ

Yo creo que son varios los mecanismos que puede tener una universidad para hacer ese seguimiento de egresados. Acercarse a las empresas o a los centros de trabajo para buscar tutores y que estos terminen siendo los mismos egresados de la universidad, es un modo de relacionarse con ellos; es más, nos permite hacer hasta una medición directa de cómo se están desempeñando. Las encuestas indirectas son otra forma de obtener información. Creo que se debe establecer un sistema de seguimiento de egresados que tenga varios componentes y que cada uno de estos brinde una información que se use de diferente manera. Uno de esos componentes puede ser la certificación: establecer sistemas nacionales de certificación, así como hay en México un examen al final de la carrera que permite saber cómo se van logrando algunas competencias de conocimiento, digamos. Pero esquemas similares se pueden hacer a los tres, cinco o siete años.

El punto es, en este esquema de la garantía, hasta dónde afecta a la formación que le

damos en la universidad. ¿Hasta los próximos tres años? ¿Hasta los próximos cinco años? Porque ocurre en un momento, ahora que es muy común la formación continua, que después de tres años esa persona, que ha trabajado mucho tiempo en un área en particular, termina siendo un experto en ella y se olvidó de todo el resto de conocimientos que pueda haber tenido. O esta persona, que se convirtió en un analista, es el mejor, pero nunca ha tenido oportunidad de desarrollar sus capacidades de liderazgo o de trabajo en equipo porque su trabajo es analizar datos mañana, tarde y noche. Es el mejor haciendo eso; pero, digamos, no ha tenido oportunidad de ejercer su liderazgo o su emprendimiento porque lo que está haciendo es exitoso.

Ahora que es tan fácil que los egresados estudien maestrías o tengan especializaciones dentro de la misma empresa, me parece bastante complicado poder medir cuál es el impacto de la universidad en un egresado de cinco, siete, veinticinco años más allá. Si es que ese impacto es resultado de nosotros o si es que ese impacto resulta de la experiencia que pueda haber ganado en el medio. Lógicamente, las preguntas que se van a hacer a un egresado de veinticinco años no son: “¿Te acuerdas de las integrales?, ¿de la integral doble?”. Existen varios mecanismos que me permitirían hacer un seguimiento de egresados, cada uno de los cuales ellos me da información diferente y se puede aplicar en un plazo de tiempo adecuado.

■ ÓSCAR PAIN

El tiempo nos ha ganado y tenemos una serie de preguntas sin responder. Quiero agradecer y quiero plantear un resumen para que queden en la mesa y creo que después pueden también acceder a cierta información. Hay varias preguntas que tienen que ver con la relación entre perfil de ingreso y perfil de egreso, cómo hacer el seguimiento entre ambos perfiles; otras preguntas que son más específicas a la metodología de la construcción de los sistemas de egresados, en el caso de nuestra universidad, está disponible en la página del vicerrectorado académico. Hay un primer documento que se ha publicado sobre eso donde pueden leer algo más.

Quiero rescatar tres ideas principales. La primera es que un norte sobre el cual nos fijarnos es la propia misión y visión de la universidad al momento en que nos reunimos a tomar decisiones sobre la información analizada de una diversidad de actores y sobre los planes de estudios. Creo que la segunda gran conclusión tiene que ver con —y ya lo decía Dieudonné ayer— que esto es algo que se construye juntos. El construir juntos requiere compromiso de docentes, de la institución y espacios en los cuales los distintos actores puedan aprender sobre esta diversidad de sistemas. Creo que muchas preguntas debieran tener la respuesta de cómo hacer específicamente para incorporar la mejora de una forma en particular y no hay una única respuesta. Esta es la conclusión general de la mesa y que nos va a tomar siempre tiempo y, como dijo Rafael, probablemente

no lo veamos, pero vamos en el camino de la mejora.

■ RAFAEL LÓPEZ

O a lo mejor sí.

■ ÓSCAR PAIN

Quiero agradecer a Sandro, a Juan Fernando y a Rafael con un fuerte aplauso. Muchas gracias.

■ LEANDRO GARCÍA-CALDERÓN

Quisiera, como parte del comité organizador, agradecer muy sinceramente a nuestros invitados, a nuestros amigos, Dieudonné, Álvaro, Rafael y Juan Fernando, por haber colaborado paciente y eficientemente con nosotros.

Ahora, invito al ingeniero Jorge Zegarra Pellanne, director de la Dirección de Asuntos Académicos, para que nos dirija unas palabras y proceda a clausurar nuestro Tercer Encuentro Internacional.

■ JORGE ZEGARRA

Quiero empezar con una cosa absolutamente informativa y práctica. Es recomendarles que entren a la página web de la universidad y busquen vicerectorado académico. En la sección “Documentos descargables”, pueden encontrar el informa de egresados que se mencionaba antes, así como la información y

las ponencias completas del primer y segundo encuentro internacional del año 2014 y 2015, respectivamente, que también fueron sobre el tema de competencias. Ya si van por ahí, pueden ir hacia la sección “Noticias y eventos”, donde van a encontrar un boletín que publicamos de forma periódica prácticamente todas las semanas solo en época de clases. Son alrededor de veinticinco o treinta boletines al año, con información bastante interesante en temas de educación superior. En la sección “Documentos descargables”, van a publicarse pronto las presentaciones que han visto, en formato PDF. Les enviaremos un correo electrónico avisándoles de esta publicación.

El año pasado, por esta fecha, cuando clausurábamos el segundo encuentro e invitábamos al tercero, decíamos que el tema iba a ser “evaluación de las competencias en la educación superior” y teníamos la tarea de seleccionar a los expositores. Justamente, en el mes de octubre, estuvo trabajando con nosotros durante una semana Dieudonné, quien colaboró muy de cerca con el equipo de gestión curricular responsable de este evento. Era absolutamente natural e indicado, como ha quedado demostrado, invitarlo para pedirle que nos acompañe y él fue quien sugirió la presencia de Álvaro Cabrera, que también ha quedado claro que han trabajado juntos y que se complementan muy bien. Un agradecimiento especial a los dos por la forma como han presentado sus ponencias. En el caso de Rafael, el año pasado, en diciembre, estuvo por acá

el doctor Teófilo Ramos, del Tecnológico de Monterrey, quien nos dijo —creo que acertó también—: “La persona adecuada que yo te puedo indicar para hablar de la evaluación de competencias es Rafael”. Y lo invitamos. He tenido la oportunidad, la suerte, en este año, de haber coincidido con él en otras tres reuniones y me quedó clarísimo que él era la persona indicada para hablar del tema y también podía haberlo sido para hablar de acreditación, que era el tema que desarrolló en las otras reuniones. En el caso de Juan Fernando, es de la casa y queríamos mostrar un poco lo que estábamos haciendo en el tema de seguimiento de egresados.

Ha quedado claro, pues, que el proceso de evaluación es un asunto complejo, que hay una serie de decisiones que se tienen que tomar, que hay que hacer cambios. Va a requerir esfuerzo, seguramente, cuando se quiera tomar estas rutas y es cuestión de decidirse y empezar, como mencionaba Rafael. El próximo año, en el mes de agosto, debe haber toda una semana en la universidad dedicada a temas de educación superior universitaria con motivo del centenario de la universidad. Dentro de esa semana, van a estar tanto las actividades a cargo de nosotros, la Dirección de Asuntos Académicos, como actividades a cargo de la Facultad de Educación, del Departamento de Educación y de la Dirección Académica de Profesorado, que somos los responsables de ver

estos temas de educación superior que, repito, van a cubrir los de docencia, entre otros.

Así, agradecemos a nuestros cuatro expositores, nuestros cuatro amigos, por las presentaciones, por su claridad, los materiales que nos han dado, la transparencia con que han compartido con todos nosotros y su disponibilidad para responder las preguntas y contactar con las personas interesadas en estos temas. Muchas gracias a ustedes cuatro por habernos ayudado a aclarar estos temas y dejarnos una serie de pistas a seguir, tanto en las presentaciones del día de ayer como en los conversatorios del día de hoy. Quiero agradecer también de manera especial a la profesora Patricia Escobar y al profesor Sandro Paz, quienes nos han acompañado en el conversatorio el día de hoy, y a Óscar Pain, quien nos estuvo también acompañando como conductor del segundo conversatorio.

Este evento no hubiera sido posible sin el apoyo y el trabajo de varios meses del equipo de gestión curricular que está a cargo de María Teresa Moreno. Además, el evento fue coordinado por Leandro García-Calderón. Mi agradecimiento también a ellos, así como a todos ustedes por su participación, sus preguntas y sus aportes. Dicho lo anterior, no me queda sino declarar clausurado el III Encuentro Internacional Universitario de evaluación de competencias en la educación superior.

Sobre los ponentes

DIEUDONNÉ LECLERCO

Nació en 1944 y se diplomó como docente de primaria; luego, máster y doctor (1975) en Ciencias de la Educación por la Universidad de Liège (ULg) en Bélgica.

Sus estancias de postdoctorado en las universidades de Pittsburgh, UCLA y UCB (U. de Colorado Boulder) le permitieron especializarse en individualización de los aprendizajes, en evaluación y en pedagogía universitaria, respectivamente.

En la ULg, ha desarrollado el Servicio de Tecnología de la Educación; el CAFEIM (Centro de Aprendizaje, Formación y Evaluación Interactiva con Multimedias); el SMART (Sistema de Apoyo a la Realización de Tests); el Servicio Guidance-Estudio (celda de apoyo a los estudiantes); el IFRES (Instituto de Formación e Investigación en la Enseñanza Superior); y el máster en Pedagogía Universitaria (Formasup) que se daba en francés y en inglés, en presencial y a distancia.

En paralelo con su cursos en la ULg desde 1975 (donde sigue dando tres cursos), ha enseñado regularmente en las Universidad de Namur, de Aosta (Italia) y, desde 1984, en la Universidad de París 13 (Facultad de Medicina de Bobigny) de la Sorbonne Paris Cité (donde sigue dando un curso desde 1984).

Sus temas de especialización son: pedagogía universitaria, concepción y evaluación de los programas de formación (basados en competencias), evaluación de los aprendizajes de los estudiantes (incluido uso crítico de preguntas de selección múltiples), psicología de la educación, metacognición (incluido autoevaluación con

grados de certeza), métodos de enseñanza y de aprendizaje (incluido audiovisual) y educación terapéutica del paciente.

Ha sido miembro de la Comisión “Psicología y Educación” del Fondo Nacional de Investigación Científica (FNRS) de Bélgica, presidente de la Comisión Nacional de “Éxito en primer año de universidad” y sigue siendo miembro de varias asociaciones científicas y revisor en varias revistas de ciencias.

Se ha formado (en Barcelona) en auditorías externas de calidad de la enseñanza superior, organizadas por la Asociación de las Universidades Europeas (AUE). Brinda asesorías para instituciones nacionales y regionales. Ocasionalmente da cursos o talleres intensivos (tres a cinco días consecutivos) en Francia, España, Lituania, Argelia, Senegal, Tunes, Colombia, Chile, México y Perú.

Es autor varios libros, así como de artículos de enciclopedias y de revistas en francés, inglés y castellano.

ÁLVARO CABRERA MARAY

Es licenciado en Artes con mención en Teoría de la Música por la Facultad de Artes de la Universidad de Chile (1998). Formó parte de los equipos de gestión del Departamento de Estudios de Pregrado, parte de la Vicerrectoría de Asuntos Académicos de su alma máter. Del año 2006 a 2014, fue director del área de Gestión de la Formación del Departamento de Estudios de Pregrado, parte de la Vicerrectoría de Asuntos Académicos de la Universidad de Chile; cargo que implica dirigir los equipos centrales a cargo de los programas de docencia transversal (cursos de formación general, *English as a Second Language*), apoyo al diseño curricular, desarrollo docente (RED para la excelencia docente) y apoyo al aprendizaje de los estudiantes. Entre 2008 y 2012, integró la Comisión de Aseguramiento de la Calidad del Proceso de Innovación Curricular de la Facultad de Medicina de esa misma institución. Ha sido evaluador de proyectos para el Ministerio de Educación de Chile. En la Universidad de Liège, en 2010, obtuvo un máster en Pedagogía en Educación Superior (Formasup) y, dos años después, realizó una estadía de especialización en evaluación, incluyendo contacto directo con Cees van der Vleuten, de la Universidad de Maastricht, uno de los más reconocidos expertos en evaluación. Actualmente es *Doctoral Student* del *Department of Organization & Leadership* del *Teachers College* de la *Columbia University*.

JOSÉ RAFAEL LÓPEZ ISLAS

Es licenciado en Ciencias de la Comunicación por el Tecnológico de Monterrey, *master of Science in Mass Communication* por Oklahoma State University y doctor of Philosophy in Media Studies por The University of Texas at Austin.

Ha colaborado con el Tecnológico de Monterrey desde 1983 como profesor de planta, director de la carrera de Ciencias de la Comunicación, director de la maestría en Ciencias con especialidad de Comunicación, director de Investigación y Tecnología Educativa y director de programas sociales en la Universidad Virtual del Tec de Monterrey, director de Investigación Institucional y Bibliotecas en la Vicerrectoría Académica, y actualmente es director de Efectividad Institucional en la Vicerrectoría de Normatividad Académica.

Ha participado en proyectos de innovación educativa, tales como el desarrollo y creación del Sistema de Educación Interactiva por Satélite (SEIS), que dio pie a la Universidad Virtual; el desarrollo de los centros comunitarios de aprendizaje y de los contenidos que ahí se imparten; y ha liderado el desarrollo del primer modelo de interacción de los cursos en línea de la Universidad Virtual, así como la creación de la editorial digital del Tecnológico de Monterrey.

Ha sido investigador asociado en Oklahoma State University y The University of Texas at Austin; además, es autor o coautor de treinta y cinco ponencias y publicaciones arbitradas a nivel nacional e internacional. A la fecha, ha asesorado veintitrés tesis de posgrado. Asimismo, ha sido consultor de la Unesco para proyectos de educación en modalidades virtuales en varios países de América Latina.

Dentro de sus responsabilidades actuales, figura la coordinación general de las acreditaciones académicas del Tecnológico de Monterrey, por lo que dirige el auto-estudio institucional para la reacreditación de la Federación de Instituciones Mexicanas Particulares de Educación Superior (FIMPES, 2016), así como el proceso de reafirmación de la acreditación de la Southern Association of Colleges and Schools Commission on Colleges (SACS COC), ante la cual también figura como enlace institucional (*liaison*). Forma parte de la Comisión Permanente de Dictaminación de FIMPES desde junio de 2013 y ha coordinado dicha comisión desde junio de 2014. Asimismo, coordina los procesos de acreditación de los programas académicos de los campus del Tecnológico de Monterrey por parte de agencias acreditadoras,

tanto nacionales como internacionales. Como director de Efectividad Institucional, su responsabilidad primaria es diseñar y facilitar los procesos de mejora continua a lo largo de toda la institución, tanto en el cumplimiento de las competencias de egreso establecidas en cada programa académico como en el logro de las metas de cada área de la organización.

JUAN FERNANDO VEGA GANOZA

Es coordinador en la Dirección de Asuntos Académicos de la Pontificia Universidad Católica del Perú. Lidera la Unidad de Información Académica, coordina el Sistema de Seguimiento de Egresados (evaluación de carácter censal y continuo sobre la experiencia formativa y trayectoria académica y laboral de los egresados) y publica, junto a su equipo, un boletín semanal que recoge las más destacadas noticias nacionales e internacionales sobre educación superior. Además, es profesor a tiempo parcial del Departamento de Teología y de la Facultad de Gestión y Alta Dirección de la misma universidad. Se desempeña también como consultor independiente.

Bachiller y licenciado en Sociología por la Pontificia Universidad Católica del Perú y bachiller en Filosofía por el ALOISIANUM (Pontificia Facoltà di Filosofia, Italia).

Se desempeñó, durante siete años, como Jefe de la Oficina de Planificación Estratégica y Medición de Calidad Educativa en el Ministerio de Educación, de la cual dependían: la Unidad de Programación, a cargo del diseño y seguimiento de los planes de corto, mediano y largo plazo del sector; la Unidad de Medición de Calidad, a cargo del diseño y aplicación de indicadores e instrumentos para la evaluación del sistema educativo; y la Unidad de Estadística, a cargo de recopilar y mantener al día la información del sistema educativo y de un sistema de información geográfica que se convirtió en el estándar para el Estado. Asimismo, fue presidente de la Comisión de Educación Rural, en la que participaron las distintas unidades de línea que actúan sobre ese tema transversal; así como jefe de su Secretaría Técnica, que diseñó y creó las condiciones para la ejecución del proyecto de educación rural.

Fue miembro de la Compañía de Jesús, profesor de colegio secundario, promotor de programas de educación popular, asesor de empresas asociativas y gremios campesinos e investigador en temas de desarrollo regional, organización rural y políticas de salud y educación. Ha liderado también equipos de investigación y

promoción del desarrollo en varias organizaciones no gubernamentales, como el Centro de Información y Desarrollo Internacional de Autogestión, la Asociación Multidisciplinaria de Investigación y Docencia en Población, el Centro de Estudios Rurales Andinos Bartolomé de Las Casas del Cusco y el Centro de Investigación y Promoción del Campesinado (CIPCA) de Piura.

Su experiencia en el sector público, privado y social le ha permitido publicar diversos ensayos, libros y artículos sobre una amplia variedad de temas que van desde los movimientos sociales y el desarrollo regional, hasta la prospectiva y la política educativa.

GEMMA PUIGVERT PLANAGUMÀ

Es doctora en Filología Clásica por la Universitat Autònoma de Barcelona (1996), donde ejerce como profesora titular de Filología Latina desde 1998. Ha sido becaria de investigación (programa de Formación de Investigadores Predoctorales) de la Generalitat de Catalunya (1992-1996) y es máster en Estudios Medievales por la Fédération Internationale des Instituts d'Études Médiévales (FIDEM), programa cursado en Roma durante los años 1992-1993.

Ha centrado sus actividades en torno a dos grandes ejes: por un lado, la filología latina (sus estudios y publicaciones en revistas nacionales e internacionales versan sobre manuscritos de época altomedieval; es autora, entre otros, del libro *Astronomía i astrología al Monestir de Ripoll. Edició i estudi dels manuscrits ceintífics del monestir de Santa Maria de Ripoll*); por otro, la didáctica de la lengua latina (es colaboradora del manual universitario *Introducció a l'estudi de la llengua llatina* [UAB, Barcelona, 1998]; es coautora del manual universitario de la Universitat Oberta de Catalunya [UOC], *Llatí i cultura llatina en el món medieval i modern. De la fi de l'Antiguitat a l'albada del Renaixement* [Barcelona, 2012]). Ha sido investigadora principal de diversos proyectos de innovación docente en enseñanza universitaria y coordinado la *Guía per a l'avaluació de competències en l'àrea de les Humanitats*, financiada y publicada por AQU-Catalunya (2007). Ha participado en el diseño por competencias del grado en Estudios Clásicos de la UAB para el que recibió una ayuda de AQU-Catalunya (2009): *l'Avaluació per competències en les activitats d'aprenentatge del nou Grau en Estudis Clàssics*.

También ejerce como traductora de la obra de Terencio al catalán: *l'Eunuç, La Magrana, Barcelona 1994 i Els germans* (RBA, Barcelona 2015).

A parte de la docencia en la UAB, ha sido docente de la UOC (de la asignatura *El llatí en el món medieval i modern*; cursos 2010-11; 2011-12; 2012-13) y de la asignatura *Latin Médiéval*, que se imparte en un postgrado de la Lumsa (Roma) desde el periodo 1999-2000.

Es también subcoordinadora de las pruebas PAU (pruebas de acceso a la universidad) de Latín desde el curso 2011-2012 hasta el presente.

Ha ejercido de coordinadora de la titulación de Filología Clásica (2000-01 a 2003-04) y de la titulación de Humanidades (2005-06 a 2008-09).

Actualmente, es directora del Departamento de Ciencias de la Antigüedad y de la Edad Media (desde marzo de 2015 hasta la actualidad).

JOSEP ANTONI IGLESIAS-FONSECA

Doctor en Filosofía y Letras (1996), especializado en Historia Medieval, por la Universitat Autònoma de Barcelona y *Premio Extraordinario de Doctorado* (2000). Desde 1992 está vinculado al área de Ciencias y Técnicas Historiográficas de dicha universidad, de la que es profesor desde el año 1996. Ha cursado estudios del Máster de *Archivística y gestión de la documentación* en la misma universidad, así como de Derecho en la Universitat de Barcelona. Ha participado en la redacción del *Libro Blanco del Grado de Historia*, elaborado por la ANECA-Ministerio de Educación y posee también experiencia en el ámbito de la gestión universitaria. Ha dirigido diferentes Proyectos de Investigación concedidos en el marco del Plan Nacional de I+D+i, siendo el más reciente *Retazos de cultura escrita. Fragmentos y membra disiecta de códices en la Cataluña central*. Asimismo, es director del grupo de investigación interdisciplinar sobre cultura escrita *Chartae*, que organiza periódicamente las jornadas *Communicatio: un itinerari històric*.

Sus investigaciones se centran principalmente en la Historia del Libro y de las Bibliotecas, y ha dedicado una atención especial al estudio de fragmentos y *membra disiecta* de códices custodiados en archivos y bibliotecas «menores» (sobre todo

en archivos comarcales, municipales, parroquiales y patrimoniales), algunos de los cuales, después de su estudio paleográfico-codicológico e identificación, han resultado ser el testimonio más antiguo conservado de su tradición textual. También se ha ocupado del estudio detallado de algunas bibliotecas medievales singulares (en particular, las de algunos canónigos, juristas y notarios que destacaron por su bibliofilia o por poseer libros excepcionales), en la estela de su tesis doctoral sobre *Llibres i biblioteques a la Barcelona del segle XV. Les biblioteques de clergues, juristes, metges i altres ciutadans a través de la documentació notarial*. Últimamente, ha iniciado una nueva línea de trabajo centrada en los problemas de conservación y defensa del patrimonio bibliográfico y documental, tanto desde el punto de vista histórico como jurídico. Ha presentado más de treinta contribuciones como ponente o comunicante en congresos, jornadas, seminarios o *workshops* de ámbito nacional e internacional.

Este libro reúne las ponencias presentadas y los conversatorios realizados en el III Encuentro Internacional Universitario: La evaluación de competencias en la educación superior. El evento fue organizado por la Dirección de Asuntos Académicos de la Pontificia Universidad Católica del Perú, y se llevó a cabo los días 11 y 12 de octubre de 2016 en el campus de la universidad.

Participaron en el evento Dieudonné Leclercq, de la Université de Liège; Álvaro Cabrera Maray, del Teachers College, Columbia University; Rafael López Islas, del Instituto Tecnológico y de Estudios Superiores de Monterrey; y Juan Fernando Vega Ganoza, de la Pontificia Universidad Católica del Perú.

ISBN: 978-612-47238-2-7



9 786124 723827