

www.universitaria.cl

Dieudo LECLERCQ



Álvaro CABRERA MARAY



UNIVERSIDAD
DE CHILE



Directores de la publicación:

Dieudonné Leclercq
Universidad de Liège (ULg)

Álvaro Cabrera Maray
Universidad de Chile (UCH)

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

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

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

Dieudonné Leclercq

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

Álvaro Cabrera Maray

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

Contenidos del libro IDEAS:

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

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

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

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

La lista de los capítulos y el resumen de cada uno

aparece a continuación después de este capítulo.

CAPÍTULO XI

Evaluar la contribución de cada participante a un trabajo grupal

DIEUDONNÉ LECLERCO, PIERRE GILLET, MARIE ERPICUM Y ÁLVARO CABRERA

Parte 1: Conceptos y desafíos

A. Colaboración y cooperación

A.1. Definiciones

La mayoría de los autores, al definir colaboración y cooperación, coinciden en que en ambos casos existen objetivos compartidos por el grupo.

Al distinguir entre ellas se plantean las siguientes diferencias:

- *Cooperación*: las tareas y responsabilidades (fácilmente identificables) son distribuidas entre los integrantes del grupo: cada uno tiene algunos asuntos específicos que atender y tareas específicas que hacer, diferentes de las de los demás integrantes.
- *Colaboración*: se produce una co-realización, una co-acción constante en el tiempo, en todos o la mayoría de los aspectos del trabajo común.

A.2. Objetivos comunes

Los docentes que adoptan el método de trabajo en grupos, a menudo para llevar a cabo proyectos (ver Capítulo 10) o para resolver problemas (ver Capítulo 7), persiguen objetivos comunes. Éstos se relacionan con desarrollar competencias asociadas al *Life Long Learning*, al “aprendizaje más allá del diploma”⁶⁵, y son difícilmente alcanzables usando otros métodos. Algunos de estos objetivos son:

Tabla 1: Objetivos comunes a la colaboración y la cooperación

- | |
|---|
| <ul style="list-style-type: none"> - Llevar a cabo un proyecto (común) o la resolución de un problema (común) - Tomar conciencia de la fortaleza de un trabajo en equipo, derivada de la complementariedad y sinergia entre las competencias, los estilos, las fortalezas y las debilidades individuales de sus integrantes. - Reforzar la autoestima por haber enfrentado con éxito un desafío, haber vivido una experiencia singular, haber cumplido una meta con una producción de calidad. - Desarrollar sentido de pertenencia (de afiliación): <i>happiness is being one of the gang</i>. - Enfrentar desafíos sociales: no decepcionar las expectativas de los otros integrantes del grupo; hacer respetar principios de trabajo en común, relativos a los derechos de cada uno, de distribución equitativa de las responsabilidades, de solidaridad. - Desarrollar sus habilidades comunicacionales (con el grupo). - Experimentar en forma vicaria (a través de la observación de otros) estrategias eficaces e ineficaces. |
|---|

⁶⁵ Boud et al. (1999); Boud y Falchikov (2005).

A.3. Objetivo específico de la Cooperación

Tabla 2: Objetivo específico de la Cooperación

– Explotar las fortalezas de cada uno en beneficio del objetivo común, a través del resultado de su fracción del trabajo [“no se le confía al tartamudo el primer rol en la obra de teatro, aunque es él quien probablemente más se hubiera beneficiado de tal decisión”⁶⁶].

A.4. Objetivo específico de la Colaboración

Tabla 3: Objetivo específico de la Colaboración

– Reforzar a los estudiantes más débiles, empoderándolos [empowerment], al hacerles asumir su rol social: colocarse al servicio de los demás, jugar el rol de guía [peer-coaching] en la Zona de Desarrollo Próximo (ZDP) de los otros, entregar retroalimentaciones recursivas [por ejemplo, con re-lecturas críticas de capítulos escritos por los demás, en lo que puede llamarse “escritura a cuatro manos”].

B. Desafíos

B.1. ¿Por qué evaluar la contribución personal?

– *Razón 1: mejorar la calidad Deontológica (equidad) de la evaluación.*

Muchos autores⁶⁷ se han preocupado por la injusticia (con frecuencia expresada por los propios estudiantes) que se puede cometer al hacer que todos los miembros de un equipo reciban el mismo puntaje por el trabajo común. Willmot y Crawford (2004) reconocen que existe el peligro de que “un estudiante holgazán se beneficie de los esfuerzos de los demás miembros del equipo y/o que un estudiante particularmente diligente pueda ver sus esfuerzos diluidos por miembros débiles del grupo”.

En esta línea de razonamiento no se debe permitir que reciba una nota no merecida:

– un estudiante *injustamente favorecido* (que se ha beneficiado de un efecto de halo, producto de la alta calidad del trabajo común), porque intencionadamente o no, ha dejado a los otros hacer el trabajo, o ha hecho solo lo más fácil (no siendo su trabajo representativo de la competencia que se buscaba evaluar).

– un estudiante *injustamente desfavorecido*, porque

a) fue víctima de daños colaterales. Por ejemplo, su parte del trabajo no pudo ser valorada porque el trabajo de otro miembro del grupo no estuvo listo a tiempo.

b) fue víctima de un convenio que le desfavoreció. Por ejemplo, para mantener la buena armonía en el grupo tuvo que renunciar a sus ideas propias o tuvo que darles una importancia menor a la merecida o a la que hubiera deseado.

⁶⁶ Formulación de Philippe Meirieu.

⁶⁷ Conway et al. (1993); Kaufman et al. (2000); Felder y Brent (1994), citados por Rodilla (2007, p. 8).

c) se vio acorralado entre dos presiones: por un lado, su compromiso con ser solidario con el grupo, y la defensa de su propio interés (incluyendo la denuncia de inequidades) por el otro.

– *Razón 2: mejorar la calidad Informativa (diagnóstica) de la evaluación.* Es posible que el docente necesite saber cuáles han sido las contribuciones exactas de cada miembro del grupo,

- para ayudar a cada estudiante a mejorar sus capacidades y sus competencias,
- para mejorar sus métodos de enseñanza-aprendizaje.

B.2. ¿Por qué no evaluar la contribución personal?

Puede ser que los objetivos en torno a la formación sean más importantes que aquellos en torno a la evaluación. Por ejemplo, si se privilegia el objetivo de solidaridad en el grupo, de apoyo a los más débiles, el foco del curso puede estar en la potenciación de la sinergia colectiva sin motivaciones extrínsecas (como la nota), y en ese caso la evaluación (y calificación) de las contribuciones individuales no es un tema prioritario.

Cada docente o cada instancia de evaluación debe elaborar un término medio entre cuánto valorar, en la calificación, la evaluación del trabajo grupal y la contribución individual de cada integrante. Este punto medio debe ser propio y pertinente a los resultados de aprendizaje que se pretende desarrollar. Las siguientes secciones del capítulo describen métodos que se pueden utilizar cuando se ha decidido que sí es importante saber, con la mayor exactitud posible, cuál fue la contribución de cada miembro del grupo al resultado colectivo.

C. El contrato al inicio

Un punto central en los trabajos grupales y su evaluación es generar un contrato inicial. Para evitar ambigüedades es necesario que este contrato inicial no solo esté bien escrito, sino que el docente se asegure de que todos los estudiantes lo han entendido bien antes de firmarlo. La validez consecuencial del dispositivo de evaluación depende mucho de estas precisiones iniciales, que los estudiantes requieren y demandan:

Hay evidencias de que los estudiantes que entran en proyectos en grupo estarán preocupados sobre el cómo serán atribuidos a cada individuo los puntos de la calificación, sobre la base del producto elaborado colaborativamente (Kennedy, 2006, p. 1).

Tabla 4: Componentes de un contrato al inicio de un proyecto grupal

Al inicio del curso, un CONTRATO debe ser aclarado y aceptado por todos:

- sobre los productos esperados del grupo ["deliverables"]:
 - el contenido (escala, extensión, estilo del resultado)
 - cronograma (fechas claves)
- sobre la estructura del grupo:
 - los roles (jefe/a, secretario/a, etc.)
 - el carácter fijo o rotativo (un turno cada uno) o flexible de los roles
 - el carácter impuesto (por el docente) o autodeterminado por el grupo (y cómo: ¿con votos? ¿Con qué clase de mayoría?) de la estructura y los roles
- sobre el ingreso y la salida de un miembro del grupo:
 - si es posible (o no) elegir: sus compañeros, el proyecto, y según cuáles reglas
 - si es posible salir del equipo: cómo y en cuáles condiciones
- sobre la evaluación:
 - si se evaluará el resultado o el proceso, o ambos
 - si existirá una evaluación de la contribución (o participación) individual, y según cuáles principios y ponderaciones
 - quién evaluará (docente / pares / expertos externos), y con cuál ponderación
 - si habrá una autoevaluación y con cuál ponderación
 - si se tendrá en cuenta la contribución de cada miembro al desarrollo de la competencia de los otros miembros
 - si la cantidad de tiempo de preparación del trabajo presentado será tomada en cuenta en el puntaje. Habitualmente no lo es, pero pueden existir excepciones. En general, solo se considera la calidad final.

(existen al menos dos razones para eso:
- dedicar un tiempo excesivo a un trabajo puede ser una evidencia de baja eficiencia.
- es difícil verificar las declaraciones de cantidad de horas dedicadas al trabajo en casa; aun las más honestas declaraciones pueden ser víctimas de mala estimación.)

Parte 2: Métodos para evaluar la contribución individual

Los métodos que siguen fueron concebidos para enfrentar uno o varios desafíos. Por ejemplo:

Desafío 1: evaluar la contribución personal (de cada individuo) al trabajo del grupo

Desafío 2: evaluar las formas que adoptó la colaboración

Desafío 3: asegurarse de que cada integrante del grupo conoce y entiende todas las secciones o aspectos del trabajo, cualquiera haya sido el autor de cada sección

Desafío 4:

La Tabla 5 se presenta como un método para apoyar la toma de decisiones sobre el tipo (o los tipos combinados) de evaluación a elegir. No está completa pues se pueden considerar otros desafíos además de estos tres, y sus celdas están vacías para que se escriban las opiniones de cada docente o grupo que la utilice.

Tabla 5: Matriz de ventajas y peligros de varios métodos para evaluar la contribución individual a un trabajo de grupo

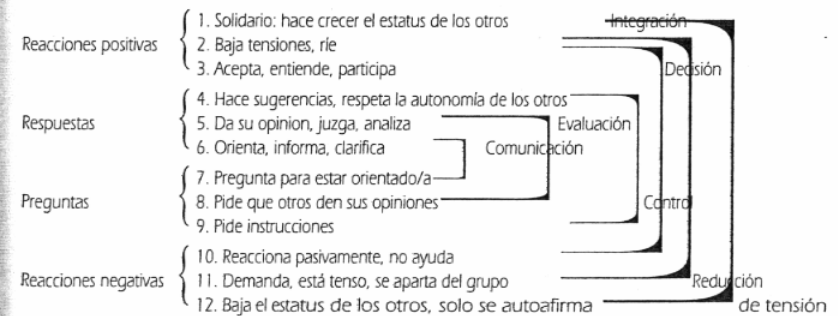
Desafío	Método	Las secciones personales	La reflexión personal añadida	La defensa oral colectiva	La declaración de participación	La observación continua
		n° 1	n° 2	n° 3	n° 4	n° 5
1. Evaluar la contribución personal						
2. Evaluar las formas de colaboración						
3. Asegurar que cada miembro entiende todo						
4.....						

D. La pauta de Bales (desafíos 1 y 2)

Para observar y luego calificar la contribución de cada miembro al "espíritu" del grupo, y a su funcionamiento, Bales (1950) propuso una pauta de 12 "eventos" (ver Tabla 6, columna central) que se puede analizar según las categorías de la columna izquierda (preguntas y respuestas; reacciones positivas o negativas), o según las de la derecha (comunicación, evaluación, control, decisión, integración).

Si se utiliza para elaborar un análisis secuencial, este tipo de pauta permite "dibujar" la dinámica de un grupo y su evolución en el tiempo. Para esto se necesita que el observador esté presente durante todos los intercambios del grupo, situación que no siempre es posible para el docente a cargo del curso. Sin embargo, los integrantes del grupo sí cumplen este requisito, por lo que con frecuencia se propone practicar "evaluaciones por los pares".

Tabla 6: Pauta de Bales para codificar las intervenciones en grupo



La "evaluación de pares" o "por los pares" no se trata en un capítulo específico en este libro. Sí se entrega un ejemplo en el Capítulo 10, donde el docente adaptó una pauta para "evaluación por el profesor", de forma que sirviera para la evaluación por los pa-

res de la contribución de cada integrante del equipo al buen funcionamiento grupal. A continuación (Tabla 7) presentamos otro ejemplo de pauta que, para el mismo efecto, puede ser utilizada por pares y/o por docentes-observadores.

Tabla 7: Extracto de <http://www.scribd.com/doc/18261604/Group-Work-Peer-Assessment> (consultado el 4 de enero de 2014)

Estuvo presente en los encuentros del grupo	
Se preparó para estos encuentros	
Propuso ideas, contenidos	
Manejó bien el tiempo	
Se focalizó en la tarea	
Fue capaz de solucionar problemas	

Este tipo de pautas con frecuencia se utiliza junto a escalas del tipo:

Tabla 8: Tipos de escalones verbales frecuentemente usados

Nunca o "se rehusaba a..." o "no lo intentó"	Escasamente	A veces u "ocasionalmente"	Habitual- mente	Siempre, o "de forma sistemática"
--	-------------	-------------------------------	--------------------	--------------------------------------

Un problema de este tipo de "escalones" es su falta de precisión debido a la interpretación que cada persona puede dar a las palabras, sobre todo en los escalones intermedios (que no son "siempre" o "nunca"). Esto puede hacer disminuir la validez informativa de las observaciones, y dificultar que estas permitan respaldar una calificación.

Para respaldar una calificación son necesarias pautas de observación y métodos de notación más precisos. Ejemplos de esto pueden encontrarse en las secciones F y G, que describen en detalle dos casos de cursos que evaluaron la contribución individual al trabajo del grupo. La calidad informativa de las observaciones que pueden lograrse aplicando sus instrumentos (como los que se ven en las tablas 13 y 14), comparadas con las observaciones logradas usando pautas como la de Bales, nos hace pensar que es preferible que los docentes elaboren sus propias pautas y escalas detalladas, de modo que evalúen realizaciones o productos específicos de sus cursos, en lugar de limitarse al uso de pautas generales. Esto puede ser hecho adaptando ejemplos de otros profesores.

E. Tres métodos para tres desafíos

E.1. Las secciones personales (desafío 1)

En esta fórmula el trabajo del grupo (por ejemplo, un informe) es firmado por todos los autores, pero los capítulos (o secciones) son firmados cada uno por un solo autor. Se considera la posibilidad de que algunos capítulos sean firmados por varios (o todos

los) autores. La calificación se hace para cada capítulo por separado, y es recibida por cada autor.

E.2. La reflexión personal añadida al trabajo común (desafío 2)

El trabajo principal es firmado por todos los autores. Además, cada uno de ellos escribe una breve nota individual que puede tratar del contenido (por ejemplo, un comentario de minoría sobre algún punto), del proceso o de otros aspectos (como los significados personales que cada uno asigna al trabajo común, o los aprendizajes individuales extraídos). Esta reflexión puede recibir una calificación individual, que puede combinarse con la calificación grupal otorgada al trabajo principal.

E.3. La defensa oral colectiva (desafío 3)

En el momento de la evaluación cada estudiante tiene en la mano una copia del trabajo común. El docente tiene su propio ejemplar, con las notas manuscritas sobre las que quiere interrogar al grupo. Siguiendo un orden circular, y empezando con el estudiante a su izquierda, el docente dirige su primera pregunta al primer estudiante, después la siguiente al segundo estudiante, y así sucesivamente.

Dependiendo de la respuesta del estudiante, son posibles algunas variantes:

- cuando un estudiante no puede contestar a una pregunta, el docente
 - designa al siguiente, que es invitado a responder
 - o designa al azar quién será invitado a responder
 - o permite que conteste el estudiante que quiera (siendo esto menos aleatorio)
- cuando un estudiante puede contestar, el docente
 - pasa a otra pregunta (y consecuentemente a otro miembro del equipo)
 - o invita a añadir algo a quien lo desee
 - o invita a un miembro que ha designado (al azar o según el orden) a comentar la respuesta del estudiante (sin haber dicho si es correcta o no)
- cuando un estudiante recibe una pregunta que concierne a la sección que ha preparado, puede ser interrogado sobre aspectos más específicos y profundizar en el tema.

En el contrato educacional inicial se estableció que, en esta evaluación, se verificará que cada integrante del grupo conozca y entienda todas las secciones del trabajo colectivo, cualquiera haya sido el autor. En consecuencia, la única respuesta que nunca será aceptada es: "no puedo contestar porque no fui yo quien hizo esa parte del trabajo".

Lo anterior obliga a los estudiantes a generar formas de comunicar las reflexiones y procesos detrás de la elaboración de una sección del trabajo, para que sea conocida

y entendida por todos, pues es altamente probable que no hayan trabajado en forma conjunta todas las secciones (lo que es normal en un trabajo grupal). Se estará así promoviendo la construcción de una narrativa común del trabajo grupal que puede fortalecer el sentido de pertenencia con el proyecto/trabajo común y aumentar la motivación.

La calificación de las respuestas de cada estudiante se combina o puede afectar la calificación grupal otorgada al trabajo principal.

F. La declaración de participación: el caso de los Proyectos de Animación Recíprocos Multimedia (PARMS)

F.1. El principio de los PARMS

Este método se basa en que cada estudiante declara al docente (a los evaluadores) cuál ha sido la importancia de su contribución particular en cada una de las secciones del trabajo. Así, cada grupo entrega una declaración que detalla la participación de cada estudiante. Esta declaración ha sido prevista en el contrato inicial y debe mantenerse oculta al docente hasta después de que este ha entregado los puntajes asignados al trabajo. Al mismo tiempo, las tareas y los criterios de evaluación deben haber sido definidos y comunicados a los estudiantes desde el inicio del curso o unidad de aprendizaje.

Un ejemplo que ilustra estos principios fue aplicado por Leclercq *et al.*, en la Universidad de Liège, en el año 2002: los llamados *PARMS* -Proyectos de Animación Recíprocos Multimedia. Fue aplicado en un curso masivo de 200 estudiantes, donde 150 fueron voluntarios para participar en este método y trabajaron en 30 equipos de 5 estudiantes cada uno. En un *PARM* cada equipo anima al curso durante un periodo fijo (en este caso, 20 minutos seguidos de 5 minutos de preguntas y respuestas), con una puesta en escena acerca de un tema relacionado con los resultados de aprendizaje del curso. En el ejemplo de los *PARMS* 2002, cada equipo había sorteado al azar una sección (3 o 4 páginas) del libro que servía como base del curso, y debía exponer los conceptos involucrados al resto de sus compañeros.

F.2. Las definiciones de las tareas y sus ponderaciones

Leclercq *et al.*, concibieron una serie de criterios de evaluación que a la vez permitían estructurar el *PARM*. Los identificaron con la sigla *DECLAR*: Desafío – Exponer – Criticar – Ligar – Animar – Responder. La Tabla 9 muestra una breve definición de cada criterio y los pesos o ponderaciones de cada uno en la calificación final.

Tabla 9: Definición y peso de las tareas según fueron anunciadas

Objetivos (y criterios de evaluación) de este Proyecto de Animación Recíproco Multimedia <i>DECLAR</i>	Peso
Desafío: explicar para qué sirve el contenido que van a presentar; ¿Cuál(es) pregunta(s) intenta responder?	10
Exponer (con apoyo multimedia de PowerPoint o video)	Forma 10 Fondo 10 Medios 10
Criticar el contenido (expresar su propio punto de vista y contraponer ideas)	10
Ligar (vincular) este contenido con ideas o documentos de la biblioteca o de la Web	10
Animar al grupo masivo (fomentar el aprendizaje activo) con alguna actividad: juego, preguntas, etc.	Fondo 10 Forma 10
Responder a las preguntas de los compañeros del grupo masivo	Fondo 10 Forma 10
	<i>Total</i> 100

F.3. La tabla teórica de los puntajes máximos

Si cada estudiante asume la misma carga de trabajo y responsabilidad en cada una de las secciones del *PARM*, la tabla teórica de los puntajes máximos sería:

Tabla 10: Distribución teórica de puntajes máximos si cada estudiante (E) asume la misma carga de trabajo en cada uno de los criterios *DECLAR*

	E1	E2	E3	E4	E5	Peso
Desafío:	10	10	10	10	10	50
Exponer:	Fondo 10 Forma 10 Media 10	10 10 10	10 10 10	10 10 10	10 10 10	50 50 50
Criticar:	10	10	10	10	10	50
Ligar (vincular):	10	10	10	10	10	50
Animar:	Fondo 10 Forma 10	10 10	10 10	10 10	10 10	50 50
Responder a las preguntas:	Fondo 10 Forma 10	10 10	10 10	10 10	10 10	50 50
	<i>Total</i> 100	100	100	100	100	500

Sin embargo, la particularidad de este método es que la repartición de estos puntajes máximos puede diferir entre estudiantes, de acuerdo con las contribuciones específicas que cada uno de ellos haya hecho a cada una de las tareas.

F.4. La repartición efectiva de las tareas en el grupo declarada por los estudiantes⁶⁸

En esta repartición se expresan las contribuciones específicas de cada estudiante a cada aspecto del PARM, y se puede hacer según –al menos– dos modalidades. Cada modalidad supone un tipo distinto de límites para los puntajes por criterio y por estudiante.

Modalidad 1:

Límite 1: El total para cada criterio debe ser 50 (10 x el número de estudiantes).

Límite 2: El total para cada estudiante debe ser 100.

La Tabla 11 muestra la repartición de puntajes de uno de los 30 grupos, siguiendo esta modalidad:

Tabla 11: Repartición de las tareas en el grupo n° 29 de la experiencia de PARMs 2002 en Liège

Grupo	Grupo	sobre	Estudiante 1		Estudiante 2		Estudiante 3		Estudiante 4		Estudiante 5		Total implic. máx. 50
			Implicación	Puntos	Implicación	Puntos	Implicación	Puntos	Implicación	Puntos	Implicación	Puntos	
Definición del Desafío	Fondo	10	10		10		10		10		10		50
Exponer	Fondo	10	10		10		10		10		10		50
	Forma oral	10	25								25		50
	Forma medios	10				25		25					50
Criticar	Fondo	10			50								50
Ligar - vincular	Fondo	10				25		25					50
Animar - activar	Fondo	10	10		10		10		10		10		50
	Forma	10	25								25		50
Respuestas a las preguntas de los compañeros	Fondo	10	10		10		10		10		10		50
	Forma	10	10		10		10		10		10		50
	Total		100		100		100		100		100		

Aun cuando Wilmot y Crawford (2005, p. 4) observan que en la mayoría de sus grupos se logra “un amplio acuerdo sobre cuáles integrantes han hecho las mayores contribuciones” al resultado grupal, permanece la amenaza de *colusión*⁶⁹ entre los integrantes del grupo (Pond *et al.*, 2007), y por ende el desafío, para el docente, de evitarla. Un ejemplo de colusión es que los estudiantes decidan declarar la misma tasa de participación para todos los integrantes, en todos los criterios, aunque no sea verdad. Esta

⁶⁸ En F.6 y F.7 veremos que esta repartición no debe ser revelada a los evaluadores (jueces) antes de que califiquen la calidad de las tareas del grupo.

⁶⁹ Lejk y Wyvill (2001), Pond *et al.* (2007), Robinson (2006), Loddington (2008, pp. 11 y 13).

colusión amenaza con ser aún más alta cuando los estudiantes ya conocen los puntajes que el docente ha asignado a los criterios –ya saben la cantidad “a repartir”–, para evitar tensiones y conflictos en el grupo.

¿De qué depende la resistencia de los estudiantes a declarar las contribuciones reales, aunque sean diferentes para cada miembro del grupo?

Rodilla (2007) utilizó el principio de “evaluación por los pares” en Escocia (Universidad de Aberdeen) sin encontrar el problema de la colusión y provocando la satisfacción de sus estudiantes, quienes no querían que algunos poco comprometidos (*free riders*, *hitchhikers*) se aprovecharan de los esfuerzos de los otros. Cuando repitió la experiencia con sus estudiantes españoles (Universidad Cardenal Herrera de Valencia) las reacciones fueron muy diferentes. Estos pusieron en duda la pertinencia (utilidad, valor moral) de las evaluaciones de sus pares, que, en ese caso, tenían la connotación de “soplonaje”. En las primeras aplicaciones observó (tal como Leclercq en la Universidad de Liège) que los estudiantes habían decidido un “pacto previo” (*prior pact*) para otorgar la misma calificación a todos⁷⁰. Rodilla (2007, p. 9) explica estas diferencias psicológicas y filosóficas por la ausencia de familiaridad con este tipo de evaluación por pares. Esta puede resultar de la mayor dificultad para organizar trabajos en grupo que existe en España, donde la flexibilidad curricular que tienen los estudiantes (múltiples opciones a elegir) redundan en escasos horarios comunes fuera de las clases, en comparación con Escocia donde los estudiantes están en el mismo lugar desde las 9:00 a las 17:00 horas, diariamente.

Modalidad 2:

Límite 1: El total para cada criterio debe ser 50 (10 x el número de estudiantes).

Límite 2: El total para cada estudiante debe ser entre 80 y 180.

Una debilidad de la modalidad anterior es que impone que, en la suma final, la participación de cada estudiante deba ser la misma (100 puntos), aunque distribuida de forma diferente en cada criterio. La modalidad 2 (que no ha sido aplicada) permite que se expresen algunas diferencias entre estudiantes en el puntaje global, indicando quiénes han trabajado más y quiénes han trabajado menos.

El límite 2 significa que se puede encontrar una situación donde hasta 4 de los 5 estudiantes se han implicado en 80 (en lugar de 100), y uno de ellos lo ha hecho en 180 (en vez de 100), con todas las variaciones posibles, manteniendo siempre que el total de las implicaciones de los estudiantes sea 500.

Al momento de calificar, es decir, aplicar esta ponderación de implicación en el trabajo a los puntajes que los evaluadores han otorgado a cada criterio, no es posible simplemente multiplicar estos últimos por 0,8 o 1,8, pues el máximo posible podría quedar desfasado. Es necesario decidir una regla del tipo: “Un estudiante que asuma sobre 1,5 del criterio recibirá un puntaje superior al del grupo (en un punto, un nivel)”.

⁷⁰ Loddington (2008, p. 13) demostró que cuando la declaración es anónima se debe temer menos que ocurra este fenómeno.

F.5. La performance del grupo

El día de la presentación-animación oral (20 minutos + 5 minutos de debate), el grupo debe entregar en 4 ejemplares (uno por cada evaluador), una hoja con formato específico que incluye:

- las fotografías y los nombres de cada integrante del grupo
- un resumen de cada tarea DECLAR: Desafío, Exponer, Criticar, Ligar, Animar, Responder

Foto: rostro Estudiante 1	Foto: rostro Estudiante 2	Foto: rostro Estudiante 3	Foto: rostro Estudiante 4	Foto: rostro Estudiante 5
Nombre Estudiante 1	Nombre Estudiante 2	Nombre Estudiante 3	Nombre Estudiante 4	Nombre Estudiante 5

Definición del problema:
La formación por OBSERVACIÓN - IMITACIÓN se opone a la PEDAGOGÍA POR PROYECTOS, que demanda creación.

Exposición: ver anexo E.

Critica: A pesar de esta oposición aparente, estas dos nociones tienen más complementariedades que divergencias.

Links:
<http://www.geoscopie.com/creation/index.html>
<http://www.mcli.dist.maricopa.edu/pbl/problem.html>
<http://www.reseauproteus.net/therapies/creativi/>
interpsychonetv2.free.fr/Marie/developpement9.pdf
 Formateur d'adultes, J-P Martin et E. Savary, Chronique Sociale, Lyon, 1999 (pp 99,100)
 Training with NLP, J. O'Connor & J. Seymour, Thorsons 1994 (p. 141)
 Les Grandes Notions de Pédagogie, M.T. Chaduc, Bordas, 1999 (257-260)
 Autonomie et apprentissage, M.-J. Barbot et G. Camatari, PUF, Paris, 1999 (pp 93, 98)

Animar - activar:
 Plegar un origami "Pequeño barco",
 Proyectar una película: "El Profesor Spirelli ¿Un genio o un holgazán?"

Respuestas a las preguntas:

Figura 1: Ejemplo de la hoja de presentación del grupo y su trabajo

Durante la presentación del PARM no todos los estudiantes del grupo toman la palabra. En el ejemplo del grupo 29 solo hablaron dos estudiantes (1 y 5). En la Tabla 11 se puede ver que se han repartido entre ellos los 50 puntos que se asignan a cada uno de estos dos criterios: Exponer / forma oral; Animar / forma oral. Este hecho se hace evidente en ese momento para los evaluadores, que siguen ignorando quiénes en el grupo se han encargado del contenido y de la mediatización.

Uno de los 5 estudiantes debe asegurar que esté visible el nombre de la etapa que está desarrollando el grupo, escribiéndolo en la pizarra o proyectándolo en una pantalla. El orden es impuesto y el mismo para cada grupo: Desafío-Exponer-Criticar-Ligar-Animar-Responder. Los evaluadores califican cada criterio asignando un puntaje y anotando observaciones y reflexiones que apoyan este juicio.

En los complementos multimedia de este libro se pueden encontrar fotografías que ilustran los actores y momentos claves de un PARM: el curso (200 estudiantes en el caso descrito), los evaluadores (4 en este caso: el docente del curso y 3 asistentes), los 5 momentos DECLAR, y la deliberación de los jueces para acordar puntajes para cada criterio y aunar comentarios.

F.6. La retroalimentación por escrito de los evaluadores

Debido a las restricciones de tiempo en las presentaciones –cada grupo tiene media hora para realizar su PARM, y son 4 grupos en el periodo de 2 horas–, los evaluadores no entregan retroalimentación en el momento. La única excepción es cuando el docente se ve obligado a re-explicar una noción mal presentada por el grupo, pues los estudiantes no pueden salir del aula con ideas falsas.

Durante el breve lapso (5 minutos) para el cambio de grupo, los evaluadores intercambian oralmente (en voz baja) sus impresiones y el docente toma apuntes de ellas. Los evaluadores acuerdan un puntaje (una calificación) a otorgar al grupo, para cada uno de los 10 criterios. Los escriben cada uno en su hoja de evaluador y el docente las recolecta.

Durante la semana que sigue el docente redacta una hoja de retroalimentación al grupo como la que se muestra en la Tabla 12 (de modo que cada semana se deben redactar 4 hojas, pues son 4 grupos). La hoja también presenta la calificación o puntaje del grupo para cada criterio, sobre un máximo de 10 puntos.

F.7. El intercambio de documentos

La Figura 2 muestra un momento crucial: el intercambio PÚBLICO de los dos documentos que influirán en la calificación final de cada estudiante. El grupo de estudiantes entrega el detalle de la participación de cada uno, según las tareas de los 10 criterios (ver Tabla 11). El docente entrega la hoja de retroalimentación (ver Tabla 12) con la síntesis de los comentarios de los evaluadores y los puntajes para los 10 criterios. Este intercambio público asegura que cuando los estudiantes repartieron las cuotas de implicación en las tareas no conocían los puntajes otorgados por los evaluadores. Simétricamente, cuando los evaluadores decidieron los puntajes no sabían quiénes en el grupo se habían encargado de cuáles tareas (excepto en el caso de quienes se hicieron cargo de la presentación oral).

Tabla 12: Hoja de retroalimentación escrita por el docente para entregarla al grupo de los 5 estudiantes la semana siguiente al PARM

Notas	sobre	Micro PARM del grupo 14 sobre <i>ww</i> (Interacción entre Aptitud y Método) para la formación. Cap. 2 14-15.
Desafío (motivo para enseñar el contenido)	10	Banal (obvio): ¿Qué implicancias tiene <i>AVT</i> -Aptitude Treatment Interaction- para la formación?
Exponer:	8	"Enfoque" inter-individual e intra-individual. La palabra "enfoque" no es suficiente para lograr la comprensión. Los ejemplos de suplantación (zoom para la focalización; flashback para recordar <i>ly no para memorizar</i>), pero el problema no se resuelve si las diferencias son ESTABLES.
Exponer:	10	MUY CLARO. Buena comparación con la pirámide de competencias (la compensación de las debilidades es del nivel demultiplicativo). Buena síntesis después de las actividades. Buena utilización de la frase de Cronbach y Snow ("aquellos que están retrasados").
Exponer:	8	Imágenes bien contrastadas: cada persona es clasificada en una aptitud vs aquella que tiene varias. Las láminas son claras, pero los textos no aparecen de manera progresiva, sino que por desgracia en bloque.
Crítica	8	En un contexto puntual (ejemplo, el interés del jefe), favorecer el corto plazo: OK.
Ligar	9	Buenos vínculos con la pedagogía del dominio (maestría) y con la remediación.
Actividad:	9	Una Hyper media es una aproximación no lineal. La TV es zapping. Su actividad ("En una semana usted rendirá un test sobre El libro de la Selva. ¿Qué media elige usted? 1... 2... 3... 4...") ¿Y si el test fuera en 2 horas? ¿Y si fuera por placer?
Actividad:	9	Entre aquellos que han elegido el libro ¿Si quedan muchas preguntas sobre la biografía de Kipling?. (Entonces la web)
Respuestas:	9	Pregunta 1: ¿Y por la educación, el corto plazo?. Buena respuesta Pregunta 2: ¿Qué métodos elegir?. Buena respuesta: "Variar"
Respuestas:	7	OK sobre el fondo, pero muy larguilll
Suma	82	100

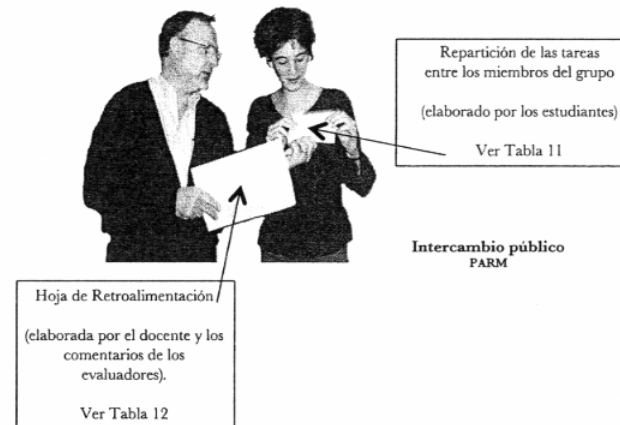


Figura 2: Intercambio público de los dos documentos (Repartición de las tareas en el grupo y Puntajes por criterio)

F8. El resultado en términos de puntajes individuales

La Tabla 13 muestra los resultados individuales de los 5 integrantes del grupo, los que resultan de la multiplicación entre el puntaje asignado al criterio por los evaluadores y la tasa de implicación que cada estudiante tiene para ese criterio.

Tabla 13: Puntajes (Columna "grupo") multiplicados por las tasas de implicación (números grandes) resultan en los puntos para cada estudiante (columnas "puntos") y en la nota final (sobre un máximo de 20)

Fecha	Grupo	29		Estudiante 1		Estudiante 2		Estudiante 3		Estudiante 4		Estudiante 5		Total Implic. máx. 50
		Grupo	sobre	Implicación	Puntos	Implicación	Puntos	Implicación	Puntos	Implicación	Puntos	Implicación	Puntos	
Definición del Desafío	Fondo	4	10	10	40	10	40	10	40	10	40	10	40	50
Exponer	Fondo	8	10	10	80	10	80	10	80	10	80	10	80	50
	Forma oral	7	10	25	175		0		0		0	25	175	50
Crítica	Forma medios	6	10		0		0	25	150	25	150		0	50
	Fondo	8	10		0	50	400		0		0		0	50
Ligar - vincular	Fondo	6	10		0		0	25	150	25	150		0	50
Animar - activar	Fondo	8	10	10	80	10	80	10	80	10	80	10	80	50
	Forma	7	10	25	175		0		0		0	25	175	50
Respuestas a las preguntas de los compañeros	Fondo	8	10	10	80	10	80	10	80	10	80	10	80	50
	Forma	8	10	10	80	10	80	10	80	10	80	10	80	50
Total				100	710	100	760	100	660	100	660	100	710	
					14,2		15,2		13,2		13,2		14,2	

Se ve que dos estudiantes reciben 13 puntos, dos reciben 14 y uno logra 15. Este ejemplo ilustra una observación de Kennedy (2006), donde refiere que los puntajes finales difieren poco del promedio del grupo (14 puntos). Aquí, esto resulta del hecho que estos 5 estudiantes declararon haber asumido muchas tareas con una distribución igualitaria del trabajo.

F.9. Las declaraciones de participación de los 30 grupos

El principio de declarar la repartición de las tareas entre los miembros de un equipo no formaba parte de la cultura de estos estudiantes. Lo anterior resultó en que casi la totalidad de los grupos contestaron que “todos participamos de forma equivalente en todas las tareas”. Vemos en esto el desafío de abrir un debate sociocognitivo entre especialistas, docentes y estudiantes, que ahonde en las razones de esta reacción y explore formas de superarla en beneficio de los aprendizajes.

F.10. Las opiniones de los estudiantes

Una pregunta en el cuestionario al final del semestre pedía su opinión sobre el método PARMs. De los 200 estudiantes que la contestaron, 190 expresaron una opinión positiva y 10 una opinión negativa. En los comentarios un estudiante hizo una reflexión de sabio: “*Vamos a ver en los exámenes*”. Leclercq compartía el interés por conocer los resultados en los exámenes finales, como una forma de evaluar la efectividad de este método en permitir que los estudiantes alcanzaran los resultados de aprendizaje pretendidos por el curso —considerando que en su mayoría los contenidos fueron presentados al curso por sus propios compañeros a través de los 30 PARMs.

F.11. Los resultados en los exámenes

Ese año los resultados (en comprensión) fueron más bajos que en años anteriores, lo que constituyó una decepción para el docente. Leclercq hizo varias reflexiones sobre estos resultados.

La primera fue: “el profesor enseña de una manera más eficaz que los estudiantes”.

La segunda, más prudente, vino a matizar la primera al considerar que “cuando enseña, el docente sabe cuáles son los conceptos importantes sobre los que va a plantear preguntas en el examen. Esto los estudiantes no lo saben con tal precisión, de modo que no pueden colocar los focos sobre estos conceptos y principios claves”.

La tercera fue que, aunque existe un resultado menor que en años anteriores en términos de conocimiento, puede ser que los estudiantes se hayan desarrollado más en términos de trabajo en equipo, de espíritu crítico, de construcción de vínculos dentro de un tema, de búsqueda de información en la Web, de presentación de contenidos verbalmente y con imágenes, de animación de un gran grupo, entre otros aspectos. Estas competencias transversales eran parte de los objetivos del curso, pero no fue concebido un dispositivo de evaluación para evaluarlas.

La cuarta reflexión fue que tampoco se había trabajado, específicamente, para que los estudiantes desarrollaran estas competencias transversales, de modo que este dispositivo de formación carecía de triple coherencia.

En conclusión, habiendo verificado que algunas de las acciones asociadas al método eran posibles, otras eran difíciles de implementar, y sobre todo había mucho por mejorar —especialmente en términos de triple concordancia (ver Capítulo 1)—, esta experiencia no fue repetida, pero pensamos que puede inspirar a colegas que, con modificaciones, quieran abordar los mismos objetivos.

G. La observación continua: el caso del curso “Legislación en el ámbito de la salud”

G.1. Desafío: la dificultad de observar “en vivo”

Son escasas las ocasiones en que se puede observar en vivo (durante la interacción) la contribución de cada estudiante al trabajo del grupo.

Una primera razón es que a menudo la mayor parte del trabajo se realiza fuera del espacio de co-presencia docente/estudiantes (las clases), ya sea en casa o en espacios de estudio de la facultad.

Una segunda razón es que, cuando se está en presencia de los grupos, si todos los estudiantes trabajan al mismo tiempo es difícil para el docente observar todos los comportamientos (orales, escritos, gestuales) de cada estudiante... y calificarlos. La grabación en video parece ser la solución ideal, pero tiene debilidades:

- multiplica la duración de la observación,
- puede modificar el comportamiento de los observados, especialmente cuando practican *brain storming*, y para eso es necesario asegurar un ambiente donde los estudiantes no se inhiban de decir cosas que puedan parecer “tonterías”,
- si el encuentro no está preparado con la función de observar la contribución de cada uno de los estudiantes, las observaciones carecen de validez teórica y de validez de replicabilidad (fiabilidad).

Una respuesta a estos desafíos se presenta en la siguiente sección: un método donde los objetivos de la observación han sido bien definidos (validez teórica) y donde los procesos de observación han sido estandarizados (fiabilidad).

G.2. Redefinición de un curso: adoptando la orientación por competencias

El curso “Legislación en el ámbito de la salud”, de la Facultad de Medicina de la Universidad de Liège, impartido por el Profesor Pierre Gillet⁷¹, permite ilustrar un método de observación y evaluación de la contribución individual en un trabajo de grupo,

⁷¹ Con el apoyo de sus asistentes Michel Marée y Marie Erpicum.

implementado a la par con una redefinición de los objetivos del curso en términos de competencias.

Considerando que la legislación belga en este ámbito es tan abundante (miles de hojas) y en permanente cambio (cada mes se crean nuevos decretos, leyes, y disposiciones), P. Gillet decidió⁷² que los objetivos no se asociaran a que los estudiantes fueran capaces de conocer la legislación existente (foco en el contenido) sino que fueran capaces de entenderla en profundidad cuando la leen (foco en la competencia). En consecuencia, concluyó que el mejor método para lograr este objetivo era que los estudiantes vivieran el papel de un creador de ley, de un parlamentario, alternativamente de gobierno o de oposición al gobierno. La solución didáctica se basó en la combinación de los eventos de aprendizaje “resolución de problemas (o experimentación)” y “debate en grupo”, junto al método “simulación”, en la forma de juego de roles.

G.3. Un método basado en la simulación del proceso parlamentario

P. Gillet aplicó la Pedagogía Por Proyectos (PPP): grupos de 8 a 14 estudiantes constituyen un partido político que prepara, durante todo el semestre, una ley en el ámbito de la salud conforme al proceso parlamentario de Bélgica. Cada una de las 10 clases semanales de 3 horas se dividió en dos periodos. Durante el primero cada grupo prepara su ley sobre un tema de actualidad. Para eso leen la legislación anterior, contactan especialistas, formulan los argumentos y los textos, entre otros aspectos, como si fueran el gobierno. Durante el segundo periodo de la clase, cada “gobierno” presenta su proyecto “en comisión parlamentaria” a los otros grupos, que juegan el papel de la oposición, formulando preguntas y enmiendas, lo que resulta en un debate que empieza oralmente pero que continúa en la internet. Cada grupo debe escribir sus intervenciones dirigidas a cada gobierno vía un sitio Web confidencial que ha sido concebido para eso.

Durante las presentaciones (del gobierno) y los debates, el docente

- juega el rol del presidente del parlamento, finalizando los debates con su martillo.
- toma apuntes (codificados) de las intervenciones de los estudiantes, en hojas especialmente preparadas. Para esto, cada uno de ellos debe decir su nombre y el de su grupo cada vez que toma la palabra para intervenir; el docente además se apoya en fotografías de los estudiantes, individuales y del grupo.

G.4. Las evaluaciones en el curso

El dispositivo de evaluación de los aprendizajes de este curso consistió en dos instrumentos:

- a) Una prueba escrita estandarizada, que aporta el 40% de la calificación final.
- b) El informe final del grupo, que aporta el 60% de la calificación final.

⁷² Ver Gillet y Marée (2005).

En este informe final todos los estudiantes reciben el mismo puntaje. Sobre esta base común se califica *la contribución de cada integrante*, mediante puntos “menos” (-) y “más” (+) que afectan la calificación individual, de acuerdo con las observaciones directas realizadas por el profesor durante las clases presenciales.

La Tabla 14 muestra un ejemplo de pauta que sirve para tomar apuntes durante la observación de las intervenciones individuales en tres momentos diferentes: (1) el trabajo en grupo (la primera mitad de la clase), (2) las presentaciones orales, y (3) el debate.

Tabla 14: Criterios de la pauta de Gillet y Ercicum (las celdas de cada estudiante son más anchas para tomar apuntes sobre la razón de la nota)

GRUPO 4		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					

Durante el debate (segunda mitad de la clase) se observan las intervenciones de los estudiantes en dos momentos:

- 1) Cuando el grupo juega el papel de la oposición se observan las preguntas y críticas que son planteadas al gobierno.
- 2) Cuando el grupo juega el papel del gobierno se observan las respuestas que dan los individuos, en representación del gobierno, a las dudas planteadas o a las contradicciones denunciadas por la oposición. El grupo que gobierna previamente ha distribuido entre sus integrantes las respuestas a las objeciones que han podido anticipar... pero en ocasiones ocurren sorpresas, y hay grupos que necesitan un "salvador".

El docente utiliza una hoja por cada grupo para anotar sus observaciones y puntajes.

Las intervenciones orales han sido preparadas por los estudiantes de modo que cada estudiante deba intervenir al menos dos veces: una vez (preparada) en la presentación del proyecto de ley de su grupo; otra vez (improvisada) en la respuesta del grupo a una pregunta o una objeción de la oposición.

G.5. El dispositivo de evaluación de la última clase

La última clase es la *culminación* de todo el proceso. En esta se escenifica la sesión de aprobación o rechazo de los proyectos de ley de cada grupo: cada versión final de proyecto de ley se presenta al "parlamento", la oposición realiza sus últimas objeciones o declara las últimas razones para entregar su apoyo, y se procede a la *votación*, que constituye un momento de suspenso (¿será adoptada la ley?).

Con frecuencia hay verdaderos políticos presentes en el aula, invitados especialmente a esta sesión final. Durante los 10 años en que este método se ha desarrollado, varias proposiciones de leyes iniciadas por los estudiantes han sido adoptadas por parlamentarios en ejercicio y se han transformado en leyes del país real.

Este indicador, del cuarto nivel de profundidad de la evaluación de Kirkpatrick (ver Capítulo 1), es decir, de impacto final del curso (en la legislación, en la política pública, en la sociedad), constituye una clara evidencia de la alta *validez ecológica* (es decir, de autenticidad, de cercanía con la realidad) de este dispositivo de formación y evaluación.

H. Conclusión

Todos los métodos que se han presentado equilibran ventajas y desventajas. Cada uno de ellos debe ser evaluado de acuerdo con los ocho criterios de calidad, o dimensiones de validez, de un Dispositivo de Evaluación de los Aprendizajes, que se resume en el modelo ETICPRAD (Capítulo 4). El último de estos criterios, la validez Deontológica —entendida en términos de equidad y justicia en la evaluación—, es particularmente difícil de lograr al evaluar las contribuciones personales a un trabajo grupal.

Referencias

- BALES, R. (1950). *Interaction Process Analysis: A method for the study of small groups*. Cambridge (MA): Addison-Wesley.
- BOUD, D., COHEN, R. y SAMPSON, J. (1999). Peer learning and assessment. *Assessment and Evaluation in Higher Education*. 24 (4), pp. 413-426.
- BOUD, D. y FALCHIKOV, N. (2005) redesigning assessment for beyond Higher education. In *Higher Education in Changing World*. HERDSA. Pp. 34-41.
- CONWAY, R., KEMBER, D., SVAN, A. y WU, M. (1993). Peer assessment of an individual's contribution to a group project. *Assessment and Evaluation in Higher education*. 18 (1), pp. 45-55.
- FELDER, R. y BRENT, R. (1994). Cooperative learning in technical courses. ERIC 377038.
- GILLET, P. y MARÉE, M. (2005). Un exemple de méthode active d'apprentissage: la simulation de séances parlementaires, in *Puzzle*. CIFEN – Université de Liège, n° 18, pp. 17-20.
- KAUFMAN, D., FELDER, R. y FULLER, H. (2000). Accounting for individual effort in cooperative learning teams. *J. Engr Education*. 89, pp. 133-140.
- KENNEDY, G. (2006). Peer assessment in group projects. Is it worth it? Australian computing Education conference 2005. Disponible en: <http://crpit.com/confpapers/CRPITV42Kennedy.pdf> (consultado el 4 de enero de 2014).
- KIRKPATRICK, D. A. (1983). *Practical Guide for supervisory Training and Development*. Reading, Mass: Addison-Wesley.
- LEJK, M. y WYVILL, M. (2001). The effect of the inclusion of self-assessment with peer assessment of contributions to a group project: A quantitative study of secret and agreed assessments. *Assessment and Evaluation in Higher education*, 26 (6), pp. 551-561.
- LODDINGTON, S. (2008). *Peer assessment of group work: a review of the literature*. WebPA Project. eLearning Capital Program. Faculty of Engineering, Loughborough University.
- POND, K; COATES, D y PALERMO, O. (2007). Student perceptions of peer review marking of team projects. <https://dspace.lboro.ac.uk/dspace-jspui/handle/2134/3040>. (Consultado el 4 de enero de 2014).
- ROBINSON, C. (2006). Self and peer assessment in group work. The SEFI MWG Seminar, Kongsberg. Cited by Loddington (2008).
- RODILLA, V. (2007). Peer assessment as an approach to judge group work: Does it work? *Edusfarm*, revista d'educació superior en Farmacia. 1, pp. 1-11.
- WILLMOT, P. y CRAWFORD, A. (2004). Online peer assessment marking of team projects. Intl Conf. on Engineering Education. Gainesville, Florida.
- WILLMOT, P. y CRAWFORD, A. (2005). Validating the assessment of individuals within undergraduate teams. Intl Conf. on Engineering in Education. Gliwice, Poland.

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

Dieudonné LECLERCQ y Álvaro CABRERA MARAY 2014

Resumen de cada capítulo

Los editores y autores principales del libro

p. 11-13

Prologo

Álvaro Cabrera &
Dieudonné
Leclercq

Parte 1: Conceptos clave en educación

p. 15-20

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

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

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

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