

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](http://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](mailto: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](mailto: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 XXI

## Retroalimentación al estudiante sobre sus evaluaciones

DIEUDONNE LECLERCO, MILTON DE LA FUENTE Y ÁLVARO CABRERA

## A. Teorías relativas a la retroalimentación

## A.1. Las retroalimentaciones ¿siempre tienen buenas consecuencias?

Desde las teorías sobre la ley del efecto (Thorndike, 1913) y el condicionamiento “operante” (Skinner, 1958), sabemos que las retroalimentaciones son un componente importante del proceso de aprendizaje. La teoría del constructivismo de Piaget (1937) insiste sobre las interacciones entre el ambiente y el individuo, que son permanentes y que permiten al aprendiz verificar sus hipótesis, es decir, verificar la eficacia del mundo interior que se ha construido como la representación del mundo exterior. El individuo observa las reacciones (retroalimentaciones) del mundo frente a sus acciones.

Se puede pensar que las retroalimentaciones tienen siempre buenas consecuencias. Sin embargo, el mismo Thorndike ya había observado que algunas retroalimentaciones escolares, especialmente cuando tienen referencias normativa (comparando la performance del estudiante con la de los demás), pueden perjudicar el aprendizaje, porque desmotivan y desvalorizan la persona misma del aprendiz. Las retroalimentaciones **NO SIEMPRE** tienen buenas consecuencias.

Además, sabemos que si bien el ambiente natural (la naturaleza) envía retroalimentación sobre cómo funcionan las cosas, esta es insuficiente pues a menudo no explica por qué lo hacen como lo hacen. Las leyes de Newton en física son ejemplos de conceptos (peso, masa, atracción, movimiento, inercia, gravedad) y de sus relaciones (entre peso, gravedad, masa y distancia, por ejemplo) que constituyen modelos explicativos que la Humanidad ha tardado milenios en poder formular. No se puede pensar que los estudiantes van a redescubrirlos solo gracias a sus interacciones con el ambiente natural. Se requiere que personas más expertas y con más experiencia faciliten este proceso de redescubrimiento, estableciendo vínculos. Eso se llama “comunicación intergeneracional” o “instrucción”.

Estas consideraciones permiten vislumbrar algunos de los factores que explican los efectos de las retroalimentaciones. Estos factores los clasificamos en dos categorías: la estructura de la retroalimentación (sus componentes y contenidos), y la dinámica de la retroalimentación, desarrollada por Kluger y De Nisi (1996) en su *Feedback Intervention Theory*. Lo anterior se combina en un modelo que llamaremos CAIRO (Concepción y Análisis de Intervenciones con Retroalimentaciones según Objetivos).

## A.2. ¿Cuáles son las relaciones entre "Reforzamiento", "Retroalimentación" y "Certeza"?

El *reforzamiento* skinneriano tiene un efecto sobre las respuestas correctas, considerando los errores como una pérdida de tiempo y desaprovechando su potencial para el aprendizaje. La *retroalimentación* entrega información al aprendiz, y su efecto depende mucho de cómo este la utiliza y de su grado de certeza al inicio (como lo han demostrado Kulhavy y Stock, 1989). Cuando se le informa que su respuesta es correcta (verificación o confirmación) el aprendiz aumenta su certeza en sus concepciones, lo que aumenta la durabilidad del aprendizaje, y dedica poco tiempo a la retroalimentación. Si, por el contrario, se le informa que su respuesta con alto grado de certeza es falsa, ocurre una "disonancia" (*discrepancy* o sorpresa). Mientras más grande esta disonancia, más esfuerzos y tiempo para entender y remediar dedica el aprendiz. Cuando la certeza es baja, aunque la respuesta sea correcta o no, la retroalimentación funciona como si se tratara de una enseñanza nueva, y el tiempo que el estudiante dedica a analizar la retroalimentación es intermedio entre los casos anteriores. Los cambios (momentáneos o permanentes) que son producto de la regulación (ajuste) basada en el análisis de las retroalimentaciones no solo se aplican a elementos de conocimiento, sino también al desempeño de tareas y a la postura personal (*Self*) que el estudiante adopta en relación con los conocimientos, las tareas, los estudios, e incluso a la vida.

## A.3. El modelo FIT (Feedback Intervention Theory) de Kluger y De Nisi (1996)

La Feedback Intervention Theory, o modelo FIT, tiene tres principios claves:

- la *fente* de las retroalimentaciones es importante. Existen dos grandes tipos de fuentes:
  - el mundo "*natural*" (R-inf-n), como consecuencia de ensayos y errores, descubrimientos y otros eventos.
  - las *reacciones humanas*, especialmente en un ambiente escolar, que Kluger y De Nisi llaman "intervenciones de retroalimentación" (*Feedback Intervention*). Los ejemplos más emblemáticos de este tipo de retroalimentación son las comunicaciones sobre los resultados en las pruebas o tareas escolares.
- existen tres niveles del impacto de la retroalimentación, que llamaremos *profundidades de penetración*:
  - el *detalle* de la ejecución de una tarea
  - la *tarea*
  - el *Self*
  - el *foco de atención*, que normalmente está ubicado en una posición intermedia (la tarea), puede desplazarse a un extremo u otro de profundidad, ya sea sobre los detalles de ejecución de la tarea como sobre el *Self*. Esto último es lo que ocurre cuando la retroalimentación inflige una herida al narcisismo del apren-

diz (hiere el *Self*). Es también lo que pasa cuando el aprendiz, viendo que tiene éxito, incrementa sus expectativas (modifica el *Self*).

## A.4. La Autoestima (Self Esteem) según William James

William James (1890) expresaba matemáticamente la autoestima (*Self Esteem*) como una *ratio* entre los desempeños (o éxitos) y las expectativas (o pretensiones) de la persona. Se puede expresar de esta forma:

$$\text{Autoestima} = \frac{\text{Desempeños (éxitos)}}{\text{Expectativas (pretensiones)}}$$

En esta teoría las personas se decepcionan cuando sus éxitos están por debajo de sus esperanzas. Para conservar un nivel aceptable de autoestima existen dos estrategias:

- aumentar los éxitos (trabajando más)
- disminuir las expectativas (considerando importantes solo aquellas tareas en las cuales tenemos talento o probabilidades de éxito)

## A.5. Nuestra representación gráfica del modelo FIT

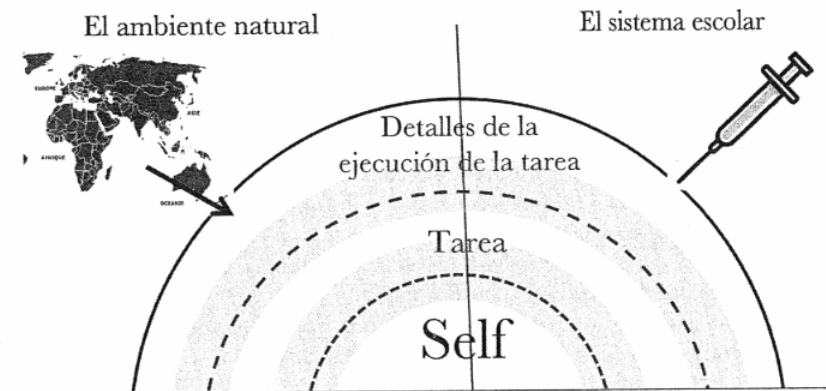


Figura 1: FIT representada por Leclercq

Las diferencias esenciales entre las dos fuentes de retroalimentación son las siguientes:

Tabla 1: Características de las dos fuentes de retroalimentación

Retroalimentaciones del ambiente natural: <i>No se puede controlar...</i>	Intervenciones de Retroalimentación: <i>Se puede controlar...</i>
...la precisión de la profundidad del impacto	
...el momento de la retroalimentación	
...el contenido y la intensidad	

Los símbolos de la Figura 1 quieren expresar las siguientes ideas:

Las líneas punteadas indican que los niveles no existen desconectados entre sí. Entre ellos existe más que una “porosidad”: hay vías de tránsito abierto. De este modo, si el profesor ha concebido una retroalimentación al nivel de los detalles de ejecución de la tarea, puede ser que su impacto penetre más profundamente, aunque no haya sido esta su intención.

Lo anterior se puede ilustrar con lo que ocurre en un curso de idioma extranjero (inglés, por ejemplo), cuando el estudiante habla con la intención de comunicar algo (tarea). Al hacer esto, comete un error de gramática inglesa. Si el profesor interrumpe al estudiante para informarle de este *detalle de ejecución de la tarea*, puede ser que el foco de su atención se desplace a este nivel, perdiendo el hilo y la intención misma de comunicar. Si, además, el profesor le dice que es el único estudiante de la clase que comete este tipo de error, el foco de atención puede desplazarse sobre el nivel del *Self*.

Otro ejemplo es un diálogo entre una madre y su niño, donde la madre está focalizada sobre un detalle de la ejecución de una tarea, pero el niño se apropia del mensaje desde su *Self*, y piensa que su madre ya no lo ama. Es necesario conocer a cada estudiante para saber cuál será su “apropiación” personal de una retroalimentación. Es una de las razones por las que en las pruebas de progreso (PdP), en algunas facultades de medicina de Holanda, hay disponible un gran número de datos sobre los resultados, y es el estudiante mismo el que elige cuáles quiere recibir (ver Capítulo 18).

Kluger y De Nisi (1996, p. 263) observan que “Cuando una tarea ha sido bien practicada, traer la atención a detalles de esta tarea puede interrumpir la ejecución del proceso automatizado y disminuir la performance”.

La *jeringa* simboliza no solo la posible precisión de la inyección (el grosor de la aguja, la profundidad de la penetración y la cantidad inyectada), sino también posibles errores en estas tres dimensiones (error de grosor, de profundidad, de cantidad), además de posibles errores de momento y de duración de la inyección.

Las *zonas sombreadas* alrededor de cada línea punteada ilustran que las distancias entre los niveles varían de un aprendiz al otro, y el observador exterior a menudo no conoce estas distancias, lo que aumenta el riesgo de error de profundidad de la “inyección” de retroalimentación. Hay estudiantes en los cuales el *Self* se ve rápidamente afectado. Hay otros a quienes las retroalimentaciones sobre detalles de la ejecución de la tarea no les distrae fácilmente de la tarea misma.

No podemos impedir que los estudiantes interpreten las retroalimentaciones. Como existen las desviaciones de interpretación, debemos esforzarnos en prevenirlas, porque corregirlas es mucho más difícil, costoso y arriesgado.

## A.6. El modelo CAIRO

Inspirándose en la síntesis de Mason y Bruning (2001), Leclercq propone una estructura de los componentes de la retroalimentación *individual* en tres facetas: retroalimentaciones sobre (1) la pregunta, (2) la(s) respuesta(s) del estudiante, y (3) la metacognición y la regulación consecutivas a la retroalimentación.

En la Figura 2 estas tres facetas (a la derecha) se combinan con la teoría de Kluger y De Nisi.

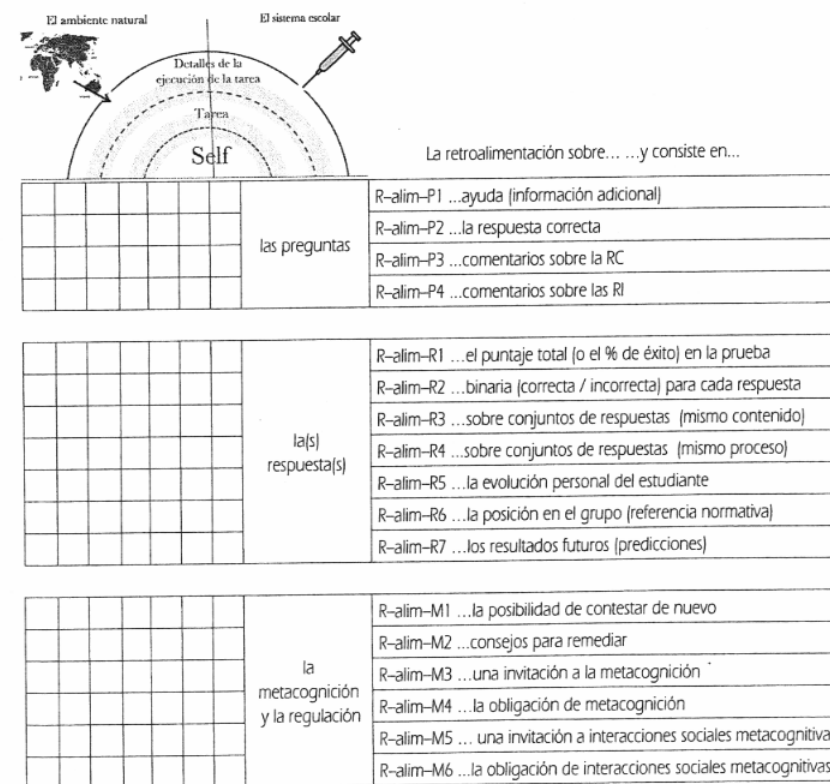


Figura 2: Modelo CAIRO para la Concepción y el Análisis de Intervenciones INDIVIDUALES de Retroalimentación según Objetivos de formación

## A.7. Los componentes de las Intervenciones de Retroalimentación individual

Los componentes pueden estar agrupados en tres grandes grupos, según lo muestra la Tabla 2.

Tabla 2: Lista de componentes de retroalimentaciones y ubicaciones de ejemplos

La retroinformación ... es sobre...	...y consiste en...	Código	Ejemplo y número de capítulo donde es descrito	Capítulo
la Pregunta	ayuda (información adicional)	R-alim-P1	Cascada convergente y facilitación progresiva - AFC	7B
	la respuesta correcta	R-alim-P2	Publicación en diario mural, acción mínima para evitar el peligro de que los estudiantes memoricen distractores como si fueran correctos	13D1
	comentarios sobre la RC	R-alim-P3	por computador (UCH) con debate (TEM)	21D 9E
	comentarios sobre las RI	R-alim-P4	PACELBRO: al nivel 1: binario (0/1) por computador (UCH)	3A3b2 21D

la(s) respuestas(s)	el puntaje total (% de éxito) en la prueba	R-alim-R1	Limitarse a esto es correr los riesgos denunciados en sección A1	(*)
	la calidad (binaria o espectral) de la respuesta	R-alim-R2	PACELBRO: al nivel 2: tipo de error (0/1) MOHICAN: espectral	3A3b2 21C2, C4
	sobre conjuntos de respuestas (que miden un mismo contenido)	R-alim-R3	En las Pruebas de Progreso (PdP) los estudiantes reciben detalles según contenidos variados. Idem UCH	18O2 21D
	sobre conjuntos de respuestas (que miden un mismo proceso)	R-alim-R4	RESSAC 3A3b4 Fig. 3, 4, 5. UCH (espectral)	3A3b4 21D
	la evolución personal del estudiante	R-alim-R5	En las PdP los estudiantes reciben la curva de sus progresos	18O2
	la posición en el grupo (referencia normativa)	R-alim-R6	En las PdP En MOHICAN	18 Fig 7 21C2
	la evolución futura del estudiante (predicción)	R-alim-R7	En las PdP los estudiantes reciben predicciones	18O3, O6

la Metacognición y la regulación	... la posibilidad de contestar de nuevo	R-alim-M1	PRM-B5 y Proyecto Spa (azul - rojo).	15E3 21C3
	... consejos para remediar	R-alim-M2	Por computador (UCH)	
	... una invitación a la metacognición	R-alim-M3	TEM: metacognición "en caliente"	9E3 y E4
	... la obligación de metacognición	R-alim-M4	TEM: Informe metacognitivo retrospectivo	9F2
	... una invitación a interacciones sociales metacognitivas	R-alim-M5	Invitación a consultar un centro de apoyo a los métodos de estudio	(*)
	... la obligación de interacciones sociales metacognitivas	R-alim-M6	Obligación de consultar un centro de apoyo a los métodos de estudio	(*)

(\*) Esta es una práctica tan generalizada que no la hemos ilustrado

## B. Retroalimentaciones para desempeños complejos

### B.1. El asunto de "¿Sobre qué?"

Frente a la pregunta sobre qué retroalimentar, se puede partir del siguiente supuesto: un error inicial produce errores subsecuentes. Así, se puede pensar que la retroalimentación más apropiada es "identificar este error de inicio, y no mencionar las consecuencias (que van a desaparecer una vez que el error "inicial" sea corregido)"<sup>193</sup>.

### B.2. El asunto del "¿Cuándo?"

Junto a Corbett y Anderson (2001), pensamos que "la mejor retroalimentación es la que ocurre inmediatamente después de la respuesta del estudiante, pero no antes de que esté listo (disponible) para hacer ajustes en su performance o en su comprensión". También compartimos con Butler y Winne (1995) que "la mejor retroalimentación es la que conduce al estudiante a revisar la tarea que ha motivado la retroalimentación, y su comportamiento en esta".

### B.3. El asunto del "¿Cómo especificar?"

Por supuesto, retroalimentaciones vagas no pueden ser diagnósticas. Kluger y De Nisi (1996) comentan la observación de Thorndike (1913, p. 286): "las calificaciones pueden derrotar al aprendizaje", y explican así este hecho: "su vicio no es solo su relatividad (la comparación con otros), sino también su indefinición (bajo nivel de especificidad)". Para ayudar a concebir retroalimentaciones más informativas para los estudiantes, Kluger y De Nisi dan un ejemplo de dos tipos de retroalimentaciones de naturaleza diferente, que pueden seguir a una tarea:

Tabla 3: Correspondencias de vocabulario entre dos formulaciones del concepto "especificidad de la retroalimentación"

En los términos de Kluger y De Nisi	Ejemplo de retroalimentación	En nuestros términos
Conocimiento del resultado	"Su rapidez promedio de dactilografía es de 100 palabras por minuto"	% del total o R-inf-R1
Intervenciones de retroalimentación	"Ud. no utiliza su pulgar para teclear"	Diagnóstico o R-inf-R4: sobre conjuntos de respuestas (proceso)

Bangert-Drowns *et al.* (1991), en su meta-análisis sobre los efectos de las retroalimentaciones, muestran la superioridad de las retroalimentaciones específicas sobre las generales. Enseñan en el gráfico de la Figura 3 la comparación entre la Amplitud de Efecto de las retroalimentaciones específicas (histograma en gris, a la derecha), y la de las re-

<sup>193</sup> <http://www.psy.gla.ac.uk/~steve/feedback.html#How> (consultado el 10 de enero de 2014).

troalimentaciones generales (histograma de la izquierda), que tienen una Amplitud de Efecto (*effect size*) promedio cercano a cero.

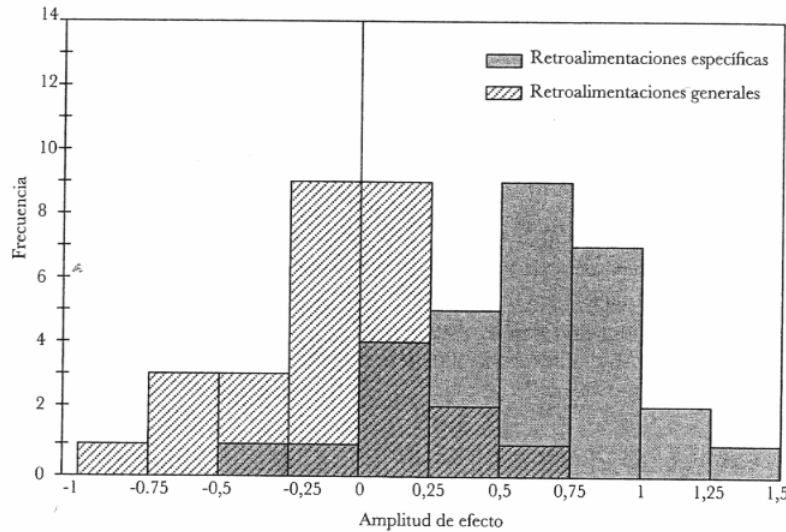


Figura 3: Las dos distribuciones de las Amplitudes de Efecto observadas en dos grupos de investigaciones resumidas por Bangert-Drowns *et al.* (1991) en su meta-análisis

#### B.4. El asunto de “¿Cómo decirlo?”

El Oxford Learning Institute (OLI) recomienda los principios que siguen (Tabla 4):

Tabla 4: Cuatro consejos del OLI

1. Comente positivamente	Cada vez que sea posible, intente dar retroalimentación positiva. Eso hará que las negativas sean más fáciles de soportar.
2. Sea específico y claro	Por ejemplo: “Pienso que su informe necesita más correcciones aquí y aquí” en lugar de “su escrito todavía es un borrador”.
3. Asuma la responsabilidad de la retroalimentación	Utilice frases desde el “Yo” en lugar de desde el “Tú” o “usted”. Por ejemplo: “Pienso que su descripción es confusa”, en lugar de “usted es confuso en esta parte”.
4. Ofrezca ideas constructivas	Por ejemplo: “Si usted utiliza la palabra xxx en lugar de yyyy, su texto será más fácil de leer.”

El cuarto punto de la Tabla 4 ilustra el principio defendido por Kulhavy y Stock (1989), quienes distinguen la *Verificación* (se comunica el carácter correcto o incorrecto de la respuesta) y la *Elaboración* (se comunican elementos que ayudan al aprendiz a lograr la respuesta correcta).

#### B.5. Combinar la retroalimentación con la certeza

Kealy y Ritzhaupt (2010) concibieron un sistema informatizado para entregar retroalimentación sobre las respuestas (R) de los estudiantes en una prueba. Las preguntas (P) de la prueba estaban concebidas para evaluar la comprensión y la memorización de textos, y las respuestas (R) debían ser redactadas (de desarrollo). A partir del uso de este sistema concibieron un dispositivo experimental con 4 etapas y 4 grupos. La primera etapa es la lectura del texto, la segunda son las respuestas de los estudiantes, la tercera es la etapa de la retroalimentación (en la que difieren los 4 grupos), y la cuarta etapa (no ilustrada en la Figura 4) es una prueba de comprensión y memorización (post test).

En la Etapa 3, de retroalimentación, los estudiantes del Grupo 1 recibieron cada una de las preguntas con la sección (párrafo, frase) del texto (T) en la que se puede encontrar la respuesta. En el Grupo 2, antes de recibir las secciones de texto, los estudiantes fueron invitados a dar un grado de certeza a cada una de sus respuestas (dadas en la Etapa 2). En el Grupo 3, recibieron las preguntas y sus propias respuestas, antes de dar su grado de certeza y conocer la sección de texto pertinente. El Grupo 4 no recibió ninguna retroalimentación. El Grupo 3 logró el mejor índice de ganancia en el post-test aplicado en la etapa 4.

##### Etapa 1 (Aprendizaje)

Lectura de Texto

##### Etapa 2 (Test - Prueba)

P1	P2	P3	P4	P5	P6	P7 etc.
R1	R2	R3	R4	R5	R6	R7

##### Etapa 3 (Retroalimentaciones)

Grupo 1: Pregunta + texto (sección)

P1	P2	P3	P4	P5	P6	P7 etc.
T1	T2	T3	T4	T5	T6	T7

Grupo 2: Pregunta + Certeza  
Texto (sección)

P1	P2	P3	P4	P5	P6	P7 etc.
C1	C2	C3	C4	C5	C6	C7
T1	T2	T3	T4	T5	T6	T7

Grupo 3: Pregunta + Respuesta + Certeza + Texto (sección)

P1	P2	P3	P4	P5	P6	P7 etc.
R1	R2	R3	R4	R5	R6	R7
C1	C2	C3	C4	C5	C6	C7
T1	T2	T3	T4	T5	T6	T7

Figura 4: Dispositivo experimental de Kealy y Ritzhaupt (2010)

### C. Retroalimentaciones sobre performances aislables

#### C.1. Ejemplo de R-alim-P3

Esta es la retroalimentación que se da acerca de la respuesta correcta.

Tabla 5: Ejemplo de R-alim-P3<sup>194</sup> (sobre la Respuesta Correcta)

PREGUNTA	RESPUESTA CORRECTA: A.
3. ALU significa a. Arithmetic Logic Unit b. Array Logic Unit c. Application Logic Unit d. Ninguna de las anteriores	Explicación: ALU es una unidad en la CPU (Central Processing Unit) de un sistema computacional, responsable de los cálculos aritméticos y las operaciones lógicas. Además de ALU, la CPU contiene MU (Memory Unit) y CU (Control Unit).

#### C.2. Ejemplo de R-alim-R2 (calidad de cada respuesta) y R-alim-R6 (posición en el grupo)

Esta es la retroalimentación sobre el porcentaje de éxito en la prueba y sobre la posición del estudiante en el grupo. En el proyecto MOHICAN (Capítulo 15, sección C3), una semana después de que 4.000 estudiantes rindieron diferentes tests, Leclercq, Georges *et al.* (2003) les enviaron retroalimentaciones personalizadas en un correo individual. La Figura 5 presenta un extracto de la retroalimentación a un estudiante para el test de Vocabulario francés: las 12 primeras preguntas de una prueba sobre 45 palabras. La calidad espectral (con grados de certeza) de cada una de sus respuestas se posiciona sobre el eje de calidades que va de -100 (error con certeza 100%) hasta 100 (éxito con certeza 100%). En el ejemplo, este estudiante nunca se equivocó con un grado de certeza más alto que 60%. Además, a la izquierda de cada palabra está la tasa de éxito de su grupo de referencia (de su facultad) o Porcentaje de Éxito de la Cohorte (PEC).

#### C.3. Focalizar la atención

Inspirándose en Cross y Fray (1977)<sup>195</sup>, Leclercq, Laguesse y Henrotay (2012) y Laguesse (2012) propusieron a estudiantes de último año de educación secundaria contestar las 20 preguntas de una prueba constituida de PSM con cuatro Soluciones Generales Implícitas: Ninguna, Todas, Faltan datos, y Absurdo (ver Capítulo 13, sección A7).

En un *primer momento* los estudiantes dieron sus respuestas (R) y sus grados de certeza (C) con un bolígrafo azul. En un *segundo momento*, el profesor-experimentador<sup>196</sup> insistió sobre un determinado proceso mental. Para el profesor de matemáticas (P.

<sup>194</sup> <http://www.psexam.com/mcq-Collection-for-Fundamentals-of-Computer/computer-fundamentals-objective-questions-mcq-with-solutions-and-explanations-set-1.html> (consultado el 10 de enero de 2014).

<sup>195</sup> Ver capítulos 4.B.2 y 16.B.2.

<sup>196</sup> P. Henrotay en matemáticas. H. Gohy-Guillet en idioma flamenco.

Henrotay) fue “Considerar en primer lugar (antes de contestar) si la Pregunta tiene sentido o si es un absurdo”. Para la profesora de lengua flamenca (H. Gohy-Guillet) fue “Formular su propia respuesta antes de considerar las soluciones propuestas”. El *tercer momento* consistió en contestar de nuevo (incluyendo nuevos grados de certeza) las mismas preguntas, con un bolígrafo rojo.

ESPECTRO DE CALIDAD DE LAS RESPUESTAS													PEC	
-100	-80	-60	-40	-20	-0	OM	0	20	40	60	80	100		
													73	corroborar
													28	prerrogativas
													35	estigmatizado
													54	subyacentes
													4	latentes
													76	divergentes
													37	denegación
													57	descrédito
													35	inferido
													50	inherentes
													55	jerarquización
													14	apresurar
respuestas incorrectas						respuestas correctas								
PEC = Promedio de éxito de la cohorte														

Figura 5: Retroalimentación individual en la prueba “Vocabulario” del Proyecto MOHICAN para las 12 primeras preguntas (sobre 45)

Tabla 6: Dispositivo experimental de las dos respuestas

MOMENTO 1			Focalización de la atención	MOMENTO 2			MOMENTO 3		
Preg 1	R1 azul	C1 azul		Preg 1	R1 rojo	C1 rojo			
Preg 2*	R2 azul	C2 azul		Preg 2*	R2 rojo	C2 rojo			
Preg 3	R3 azul	C3 azul		Preg 3	R3 rojo	C3 rojo			
Preg 4	R4 azul	C4 azul		Preg 4	R4 rojo	C4 rojo			
Preg 5*	R5 azul	C5 azul		Preg 5*	R5 rojo	C5 rojo			
etc...	etc.	etc.		etc...	etc.	etc.			
Pred = pregunta			R = respuesta			C = certeza			



Este diseño permitió medir las mejoras específicamente en las preguntas “sensibles”. Por ejemplo, en la prueba de matemáticas solo en tres preguntas (sobre 20) la respuesta correcta era “Absurdo”. En la Figura 6 se ve que la mejora, en términos de respuestas correctas, fue solo de 9% (pasando de 22% a 31%), pero que los niveles de certeza 80% y 100% entre las respuestas correctas han crecido.

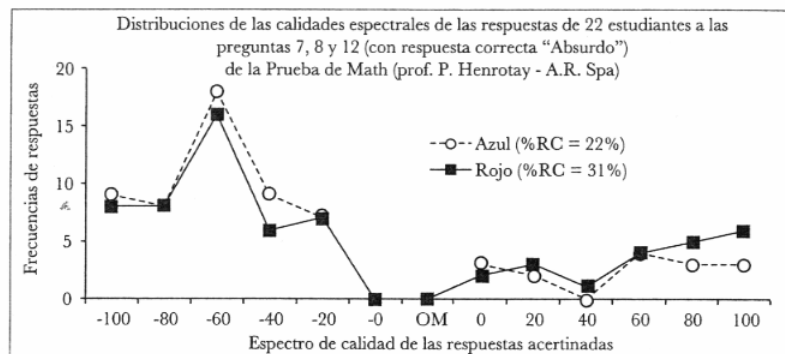


Figura 6: Distribuciones de las respuestas acertadas en la experiencia de matemáticas en Spa

El número de *falsas alarmas*, es decir, dar la respuesta “Absurdo” a otras preguntas *no varió* entre las respuestas “azules” y “rojas”.

#### C.4. Ejemplo de R-alim-R2 (binaria y espectral) para docentes

Esta es la retroalimentación acerca de procesos comunes a un conjunto de respuestas. En el ejemplo se aborda la tasa promedio de éxito, la confianza promedio de los estudiantes, y la estimación promedio que hacen profesores del porcentaje (%) de éxito de los estudiantes. Para cada una de las 45 preguntas del test de vocabulario del proyecto MOHICAN<sup>197</sup>, Leclercq, Gilles y Dupont (2003) compararon tres valores representados en la Figura 7, donde:

- el eje horizontal representa la tasa promedio de éxito de un grupo de 200 estudiantes de psicología (testado en octubre de 1999), para 13 de las 45 preguntas (psm) de vocabulario francés.
- los círculos son las estimaciones promedio que 13 profesores de estos estudiantes hicieron (en mayo de 2000) sobre la tasa promedio de éxito que habrían obtenido en octubre de 1999.
- los rombos son las certezas promedio de estos estudiantes en octubre de 1999, lo que permite observar sobrestimaciones (por ejemplo, el ítem *inherente*) y subestimaciones (por ejemplo, el ítem *corroborar*).

<sup>197</sup> Ver Capítulo 15, sección C.3.

- los rectángulos son las tasas promedio de éxito observadas en octubre de 1999.

Las preguntas se presentan según las estimaciones de los profesores, en un orden creciente de facilidad: desde la estimación de éxito más baja a la más alta. Se puede observar que, en promedio, las tasas de éxito observadas son más o menos 20% más bajas que las estimaciones de los profesores. Estos se habían preparado para lo peor, pero aun así se sintieron decepcionados con el resultado de los estudiantes en el test.

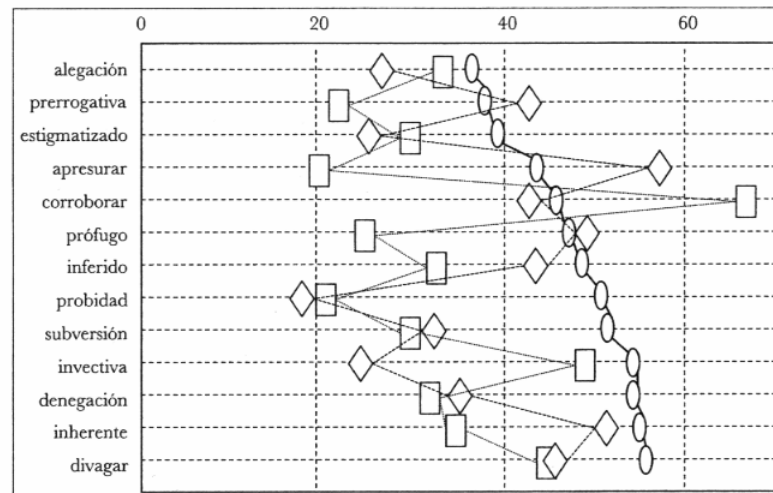


Figura 7: Los tres resultados para cada una de las 13 preguntas del test de vocabulario de MOHICAN

## D. Retroalimentación que apunta a lo esencial: el caso de medicina de la Universidad de Chile

### D.1. El desafío y el contexto

Como hemos visto (sección A), Leclercq (2012) y Kulhavy y Stock (1989) recomiendan utilizar grados de certeza, junto a otras informaciones, para favorecer el uso óptimo de la retroalimentación. Por otro lado, si los estudiantes reciben demasiadas informaciones existe el peligro de que no las lean, o no las entiendan, y en definitiva no las utilicen para regular su estudio a favor de su aprendizaje.

Milton de la Fuente<sup>198</sup>, uno de los autores de este capítulo, tuvo en cuenta esta necesidad de concisión y diseñó un sistema<sup>199</sup> para la retroalimentación automatizada

<sup>198</sup> Profesor del Programa de Fisiología y Biofísica del Instituto de Ciencias Biomédicas (ICBM) de la Facultad de Medicina de la Universidad de Chile.

<sup>199</sup> Usando el compositor de texto LaTeX.

a los estudiantes, basado en las conclusiones de Kulhavy y Stock (1989): los estudiantes dedican más atención y tiempo a las “disonancias”, es decir, a los errores en los que habían entregado una certeza alta. De la Fuente decidió que su sistema de retroalimentación se enfocara en los resultados de aprendizaje *más importantes* (finalidad formativa) tanto para ese grupo curso específico como para cada estudiante por separado. Ha sido implementado desde 2010 en la carrera de medicina de la Universidad de Chile, para la retroalimentación en las pruebas de física y química de los estudiantes de primer año. Estas pruebas consisten en preguntas de selección múltiple (PSM) que utilizan grados de certeza con la consigna recomendada en el Capítulo 16, sección B3, Figura 3. Es decir, los 6 grados de certeza (múltiplos de 20%): 0%, 20%, 40%, 60%, 80% y 100%, en este caso codificados como 0, 1, 2, 3, 4, 5.

La Figura 8 muestra la estructura del sistema de retroalimentación y las retroalimentaciones más importantes.

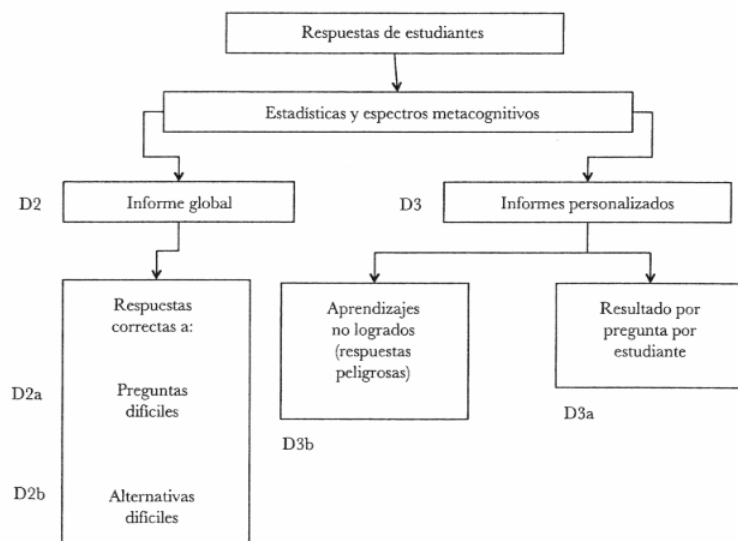


Figura 8: Estructura de las retroalimentaciones “urgentes” (según Milton de la Fuente)

## D.2. Retroalimentación general

Es la retroalimentación que se entrega a todos los estudiantes sobre dos aspectos que pudieran haber aparecido en los resultados de la prueba sobre física de fluidos: preguntas especialmente difíciles (D2a) y alternativas problemáticas (D2b).

**D2a. Concepciones erróneas frecuentes.** En la retroalimentación general primero se identifican las preguntas que muestran peor resultado. Esto se logra a partir de los espec-

tros de calidad de las respuestas, que distribuyen las frecuencias de los grados de certeza (de 0 a 100%) de las respuestas de los estudiantes en dos hemiespectros: uno para las respuestas correctas y otro para las incorrectas. Si el gráfico de una pregunta muestra una importante contribución de respuestas erróneas con alto grado de confianza (como ocurre en la Figura 10), entonces esa pregunta (aquí la Preg 9) será explícitamente retroalimentada.

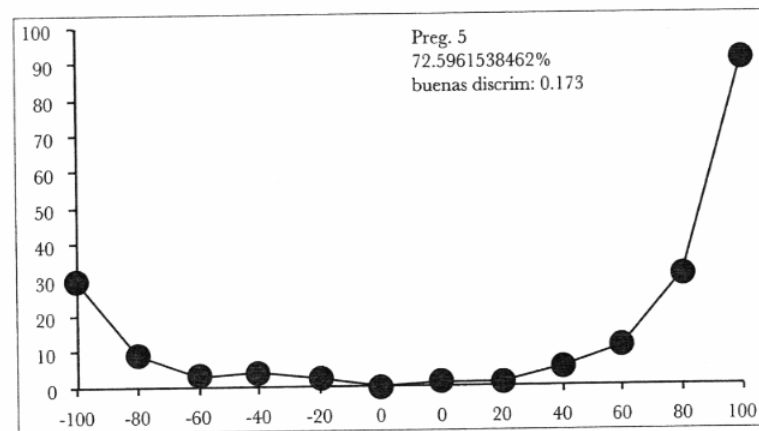


Figura 9: Una pregunta sin problema

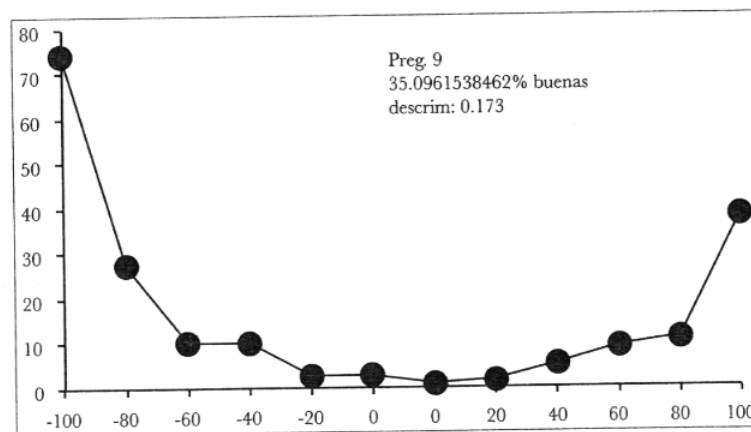


Figura 10: Una pregunta con problemas

Para retroalimentar se da la solución general al problema, sin referencia precisa a las condiciones exactas del enunciado de ese ítem (a menos que una o varias de esas condiciones sean cruciales en la solución).

**D2b. Errores populares.** Un segundo criterio para la retroalimentación general proviene del análisis estadístico de la prueba. Si una alternativa incorrecta de una pregunta fuera escogida masivamente (como en la Tabla 7 a continuación), entonces esa alternativa será discutida y aclarada, o se dará derechamente la respuesta correcta (si corresponde a un tema de conocimiento más que de comprensión o análisis). Por ejemplo, en las siguientes distribuciones de frecuencia de elección de las alternativas para tres preguntas (el asterisco indica la correcta), la alternativa incorrecta B de la pregunta 2 será escogida para su discusión/información:

Tabla 7: Popularidades (tasas de elección) de las soluciones de 3 PSM

PREGUNTA	O	A	B	C	D
1	3,36	11,5	74,0*	4,32	6,73
2	1,92	2,88	44,7	6,73	43,7*
3	9,13	17,3	13,4	34,1*	25,9

La idea es que en todos los casos problemáticos el estudiante salga de la clase con el conocimiento correcto, pero sin afectar indebidamente la confidencialidad que caracteriza a las pruebas de selección múltiple en muchas instituciones, incluido el ICBM de la Universidad de Chile.

### D.3. Retroalimentación individual

**D3a. Vista general.** Antes de retroalimentar, a cada estudiante se le presenta el detalle de los datos entrados al sistema, con el fin de que él/ella pueda verificar —de una fotocopia de su hoja de respuestas— que su prueba fue bien leída y los datos ingresados son correctos. Al mismo tiempo obtiene una visión global de sus resultados, incluyendo cantidad de respuestas correctas e incorrectas, y los grados de certeza de cada una, lo que hace posible calcular índices de dominio subjetivo del conocimiento, como la Confianza y la Imprudencia.

Por ejemplo, las tablas 8 y 9 contienen la identificación del alumno (RUT), seguida por una fila que contiene el número de cada ítem, la respuesta correcta (el número de la alternativa correcta), la respuesta elegida por el alumno, y el grado de certeza con el que contestó (0=0%; 1=20%; 2=40%; 3=60%; 4=80%; 5=100%).

Tabla 8: Datos del estudiante 1

		ESTUDIANTE 1																			
preguntas		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
correctas		4	2	3	1	2	2	1	3	2	1	2	4	2	1	2	1	3	2	4	2
respuestas		4	2	3	1	2	2	1	3	1	1	4	1	4	1	2	1	3	2	2	2
confianzas		3	4	3	3	4	4	4	4	3	4	3	4	1	3	3	4	5	4	3	4

Así, este alumno contestó correctamente la pregunta 1 con un 60% de certeza, correctamente la pregunta 2 con un 80% de certeza, incorrectamente la pregunta 9 con 60% de certeza, etc.

Y un segundo ejemplo para otro alumno, quien contestó correctamente la pregunta 1 con un 80% de confianza, incorrectamente la pregunta 2 con un 100% de confianza, etc.

Tabla 9: Datos del estudiante 2

		ESTUDIANTE 2																			
preguntas		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
correctas		4	2	3	1	2	2	1	3	2	1	2	4	2	1	2	1	3	2	4	2
respuestas		4	1	3	3	2	3	1	3	2	1	4	4	4	1	2	1	3	2	4	2
confianzas		4	5	5	4	4	5	4	5	3	5	3	5	0	5	4	0	5	5	5	5

**D3b. Focalización sobre los errores.** Luego viene la retroalimentación propiamente tal. Lo más importante en este proceso es dar a conocer a cada estudiante los indicadores de aprendizaje asociados a cada pregunta (no los ítems mismos), que él o ella contestaron incorrectamente con 80% o 100% de seguridad. Dicho de otro modo, se enfoca en aquellos aprendizajes específicos que el/la estudiante podría no haber logrado en absoluto, o incluso haber adquirido concepciones erróneas. Los indicadores generalmente se asocian con más de un ítem, de modo que esta información le permita al estudiante darse cuenta de si está fallando consistentemente en un determinado aprendizaje específico. Esto ocurre así en algunos casos, mientras que en otros las fallas se reparten entre varios indicadores.

El primer caso (errores consistentes) es un indicio de que el aprendizaje de ese tema en particular no se ha logrado en absoluto y tiene que re-abordarse completamente de nuevo. A continuación se ejemplifica un ítem para mostrar cómo se incluye su respectivo indicador de aprendizaje en la retroalimentación que recibe el estudiante (colocado al final del enunciado después de uno o más paréntesis).

Un paciente recibe una transfusión de sangre por medio de una aguja de 0.2 mm de radio y 3 cm de longitud...[...]... *calcula presión hidrostática seleccionando las variables apropiadas.*

Una forma de retroalimentar es entregar el listado de indicadores de aprendizaje asociado a cada pregunta del certamen, de modo que cada estudiante verifique los indicadores asociados a sus respuestas usando la tabla anterior. Un fragmento del listado se vería como la Tabla 10, asociados a las preguntas de una prueba de física.

Tabla 10: Indicadores de aprendizajes por ítem

1. describe y calcula el movimiento de cuerpos usando los conceptos de posición, velocidad y aceleración
2. interpreta gráficos de posición contra tiempo, rapidez contra tiempo y aceleración contra tiempo, usando las unidades apropiadas
3. plantea y resuelve vectorialmente problemas de movimiento en el plano usando aceleraciones y fuerzas
4. aplica la ley de la conservación de la energía a cambios en el estado de cuerpos, que incluyen trabajo y variación de la energía cinética e intercambio de energía potencial

O, como alternativa, se puede entregar al/la estudiante solo la información 'peligrosa', es decir, los indicadores asociados a respuestas incorrectas emitidas con 80-100% de certeza. Por ejemplo, para el alumno 1 a continuación se encontró que solo dos respuestas cayeron bajo este criterio, ambas asociadas al mismo indicador de aprendizaje (por lo que en su retroalimentación aparece dos veces):

Tabla 11: Retroalimentación a estudiante con 2 respuestas "peligrosas" asociadas al mismo indicador

RUT de alumno 1

- a) describe, plantea y resuelve vectorialmente problemas de movimiento en el plano usando aceleraciones y fuerzas
- b) describe, plantea y resuelve vectorialmente problemas de movimiento en el plano usando aceleraciones y fuerzas

En cambio, en otro estudiante, las respuestas 'peligrosas' estuvieron asociadas a diferentes indicadores:

Tabla 12: retroalimentación a un estudiantes con 5 respuestas "peligrosas" asociadas a distintos indicadores

RUT de alumno 2

- a) interpreta y genera gráficos de posición contra tiempo, rapidez contra tiempo y aceleración contra tiempo, usando las unidades apropiadas
- b) describe, plantea y resuelve vectorialmente problemas de movimiento en el plano usando aceleraciones y fuerzas y los principios de Newton, y especificando el estado de equilibrio
- c) aplica la ley de la conservación de la energía a cambios en el estado de cuerpos para calcular trabajo de roce y variación de la energía cinética
- d) describe o reconoce sistemas con un movimiento armónico simple
- e) descompone las fuerzas en sus diversos componentes

#### D.4. Opiniones de los estudiantes sobre el método

Los estudiantes tienen una opinión muy positiva de este enfoque metacognitivo. Por ejemplo, cuando se les pregunta en una encuesta si les parece positivo que en el certamen se solicite especificar el grado de confianza con que se contestó cada ítem, se encuentra la siguiente distribución de respuestas (incluye al 78% de los 219 estudiantes de primer año de la carrera de medicina en la Universidad de Chile en 2010):

Tabla 13: Tasas de respuestas a la encuesta de Milton de la Fuente sobre satisfacción con el método

Muy en desacuerdo	En desacuerdo	De acuerdo	Muy de acuerdo
2%	2%	40%	56%

Cuando se les pide que expresen sus ideas al respecto, en una muestra representativa de las respuestas manifiestan que:

Tabla 14: Algunos de los comentarios de los estudiantes a la encuesta de Milton de la Fuente

- Me parece que es una herramienta útil para ver qué materias son las más entendidas y cuáles no.
- Me parece súper bien que hagan eso porque ayuda a los alumnos a autoevaluarse.
- Encuentro que es una muy buena medida para que los profesores vean realmente cuán adquiridos tenemos los conocimientos.
- Resulta algo muy bueno, ojalá se siga haciendo y de esta manera se puedan ver las falencias que se tienen en el aprendizaje y mejorarlo.
- Creo que es un método muy útil para evaluar si realmente se aprendió.

#### D.5. Conclusiones sobre esta acción-investigación

Milton de la Fuente ha establecido un método para retroalimentar a los estudiantes respecto a sus aprendizajes luego de las pruebas de alternativas, tanto en general con respecto a ítems especialmente difíciles, como individualmente. En principio, esta retroalimentación debería contribuir a mejorar los aprendizajes, fomentando que los estudiantes no pasen por la prueba con ideas muy erradas respecto a los aprendizajes que deberían lograr. Los estudiantes se pronuncian a favor del sistema.

#### E. Ejemplos analizados con el modelo CAIRO

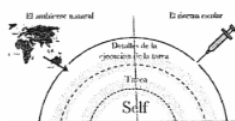
En la Tabla 15 se analizan cuatro Dispositivos de Evaluación de los Aprendizajes (DEAS) según el modelo CAIRO (Figura 2). Estos cuatro DEAS han sido analizados de acuerdo con sus componentes (parte derecha del modelo CAIRO), y no según sus impactos y niveles de profundidad (parte izquierda del modelo CAIRO). Los dispositivos son:

- Spa (bolígrafos azules y rojos): ver Capítulo 21, sección C.3.
- TEM (Test Espectral Metacognitivo): ver Capítulo 9.
- Med UCH (sistema automático de retroalimentación): ver Capítulo 21, sección D.
- PdP (Pruebas de Progreso): ver Capítulo 18.

Los números que aparecen en algunas celdas hacen referencia a notas que se describen debajo de la Tabla 15.

Tabla 15: Cuatro DEAS analizados con la sección "Componentes" del modelo CAIRO

(ver Capítulo 21.A)



		La retroalimentación	Spa	TEM	UCH	PdP
	anticipada por	R-alim-A1...consignas específicas	1	1	2	2
		R-alim-A2...grados de certeza	X	X	X	

		sobre.....y consiste en...				
	las preguntas	R-alim-P1...ayuda (información adicional)	3			
		R-alim-P2...la respuesta correcta		X	X	X
		R-alim-P3...comentarios sobre la RC	4	4		
		R-alim-P4...comentarios sobre las RI	4	4		

		la(s) respuesta(s)				
		R-alim-R1...el puntaje total (o el % de éxito) en la prueba	X	X	X	X
		R-alim-R2...binaria (correcta / incorrecta) para cada respuesta	X	X	X	X
		R-alim-R3...sobre conjuntos de respuestas (mismo contenido)				X
		R-alim-R4...sobre conjuntos de respuestas (mismo proceso)			X	
		R-alim-R5...la evolución personal del estudiante				X
		R-alim-R6...la posición en el grupo (referencia normativa)				X
		R-alim-R7...los resultados futuros (predicciones)				X

		la metacognición y la regulación				
		R-alim-M1...la posibilidad de contestar de nuevo	X			
		R-alim-M2...consejos para remediar			X	
		R-alim-M3...una invitación a la metacognición		X		
		R-alim-M4...la obligación de metacognición		5		6
		R-alim-M5...una invitación a interacciones sociales metacognitivas	X	X	X	X
		R-alim-M6...la obligación de interacciones sociales metacognitivas				7

## Notas

Acerca de las consignas:

- Los estudiantes saben que un debate (inmediato) vendrá después de la prueba, y que en esta instancia podrán defender sus posiciones (respuestas).
- Los estudiantes saben que recibirán una retroalimentación detallada por computador.

Acerca de las preguntas:

- Entre las dos series de respuestas (con bolígrafo de tinta azul y después de tinta roja), la atención de los estudiantes es focalizada sobre un proceso mental particular.
- Estos comentarios van a ocurrir durante el debate.

Acerca de la metacognición y la regulación:

- Obligación de redactar un informe metacognitivo “en frío”.
- Obligación de constituir un portafolio con elementos metacognitivos.
- Encuentros periódicos (obligatorios) con un tutor.

## F. Conclusiones

Hemos dedicado mucha atención a las variadas formas de retroalimentación a los estudiantes acerca de sus resultados en las evaluaciones, porque esta operación no solo pertenece al ámbito de la evaluación, sino que también al de la enseñanza y el aprendizaje. Esperamos que el ingenio que se refleja en los ejemplos descritos en el capítulo fortalezcan las intenciones de los lectores de innovar en este asunto central en la educación.

## Referencias

- BANGERT-DROWNS, R., KULIK, C., KULIK, J. Y MORGAN, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, 61 (2), pp. 213-238.
- BUTLER, D. Y WINNE, P. (1995). Feedback and Self-Regulated Learning: A theoretical synthesis. *Review of Educational Research*, 65 (3), pp. 245-281.
- CORBETT, A.T., Y ANDERSON, J. R. (2001). Locus of feedback control in computer-based tutoring: Impact on learning rate, achievement and attitudes. In *Proceedings of ACM CHI 2001 conference on human factors in computing systems* (pp. 245-252). New York: ACM Press.
- CROSS, L. Y FRARY, R. (1977). An empirical test of Lord's theoretical results regarding formula scoring of multiple choice tests. *Journal of Educational Measurement*, vol. 14, 313-321.
- JAMES, W. (1890). *The principles of Psychology*, Vol. 2, Chapter 18, Macmillan, London.
- KEALY, W. Y RITZHAUPT, A. (2010). Assessment certitude as a feedback strategy for learners' constructed responses. *Journal of Educational Computing Research*, 43 (1), p. 25-45.
- KLUGER, A.N. Y DE NISI, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254-284.
- KULHAVY, R.W. Y STOCK, W.A. (1989). "Feedback in written instruction: The place of response certitude". *Educational Psychology Review* 1(4): 279-308.
- LAGUESSE, C. (2012). Quelle perception des élèves de l'enseignement secondaire supérieur général et de leurs enseignants quant à l'utilisation de Tests Spectraux Métacognitifs. Master thesis. Université de Liège.
- LECLERCQ, D., GEORGES, F., GILLES, J.L., PIRSON, M. Y SIMON, F. (2003). Les feedbacks des check-up MOHICAN aux étudiants et aux enseignants. Chapitre 2 de Leclercq, D. (Ed) (2003). *Diagnostic cognitif et métacognitif au seuil de l'université. Le projet MOHICAN mené par les 9 universités de la Communauté Française Wallonie Bruxelles*. Liège : Editions de l'université de Liège, pp. 51-66.
- LECLERCQ, D., GILLES, J.L. Y DUPONT, C. (2003). L'information des feedbacks pour les professeurs. L'opération VOCA PROFS à l'ULg ou Estimation par les professeurs des résultats des étudiants au check-up de Vocabulaire. Chapitre 6 de Leclercq, D. (Ed) (2003). *Diagnostic cognitif et métacognitif au seuil de l'université. Le projet MOHICAN mené par les 9 universités de la Communauté Française Wallonie Bruxelles*. Liège : Editions de l'Université de Liège, pp. 117-125.
- LECLERCQ, D., LAGUESSE, C. Y HENROTAY, P. (2012). Metacomprehension and cognitive vigilance. A multi-teachers experiment in last year of secondary school. Symposium of the EARLI SIG "Metacognition". Milano.

- MASON, B.J. Y BRUNING, R. (2001). Providing Feedback in Computer-based Instruction: What the Research Tells Us, <http://dwb.unl.edu/Edit/MB/MasonBruning.html> (consultado el 10 de enero de 2014).
- PIAGET, J. (1937). *La construction du réel chez l'enfant*. Neuchâtel: Delachaux ey Niestlé. (Translation: *The construction of reality in the child*. London: Routledge y Kegan Paul, 1968).
- SKINNER, B.F. (1958). Reinforcement Today. *American Psychologist*, 13, 3.
- THORNDIKE, E.L. (1913). *Educational Psychology*. Vol. I: *The original nature of man*. New-York: Columbia University, Teachers College.

# 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	<b>ATOME (Alineamiento en un Tablero de Objetivos, Métodos y Evaluaciones.</b> 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.	<b>D.Leclercq &amp; Álvaro Cabrera</b> p. 23-34
2	<b>Los componentes de un dispositivo de evaluación de los aprendizajes (DEA)</b> 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.	<b>D. Leclercq</b> p. 35-50
3	<b>El prisma de las características de un Dispositivo de Evaluación de los Aprendizajes (DEA)</b> 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.	<b>D. Leclercq</b> p. 51-82
4	<b>ETIC PRAD: Ocho criterios de validez de un Dispositivo de Evaluación de los Aprendizajes (DEA)</b> 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).	<b>D. Leclercq</b> p. 83-92
5	<b>Autodescribir y evaluar el Dispositivo de Evaluación de los Aprendizajes (DEA) de un curso</b> 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.	<b>D. Leclercq &amp; Álvaro Cabrera</b> p. 93-102

6	<p><b>La calificación subjetiva de los desempeños complejos: Criterios y rubricas</b>                  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><b>D. Leclercq &amp; Álvaro Cabrera</b>  p. 103-128</p>
7	<p><b>Evaluar la capacidad de resolver problemas</b>                  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><b>D. Leclercq, S. Delcomminette</b>  (HERS) &amp; A. Cabrera p. 129-152</p>
8	<p><b>ECOE: Exámenes Clínicos Objetivos y Estructurados</b>                  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><b>G. Philippe (ULg), D. Leclercq &amp; J-P. Bourguignon (ULg)</b> p. 153-170</p>
9	<p><b>Meta cognición y Tests Espectrales Metacognitivos (TEMs)</b>                  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><b>D. Leclercq &amp; Álvaro Cabrera</b> p. 171-196</p>
10	<p><b>Evaluar los Aprendizajes en la Pedagogía Por Proyectos (PPP)</b>                  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><b>Álvaro Cabrera</b>  p. 197-220</p>
11	<p><b>Evaluar la contribución de cada participante a un trabajo grupal</b>                  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><b>D. Leclercq, P. Gillet (ULg), M. Erpicum (ULg) &amp; A. Cabrera</b> p. 221-242</p>
12	<p><b>Los Portfolios: Hacia una evaluación más integrada y coherente con el concepto de desempeño complejo</b>                  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><b>M. Poumay (ULg) &amp; Chr. Maillard (ULg)</b> p. 243-260</p>



13	<p><b>Las Preguntas de Selección Múltiples (PSM): del currículo escondido a la vigilancia cognitiva</b>                  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><b>D. Leclercq &amp; Álvaro Cabrera</b>                   p. 261-286</p>
14	<p><b>Reglas de redacción de las Preguntas de Selección Múltiples y la habilidad para responder pruebas</b>                  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><b>D. Leclercq</b>                   p. 287-300</p>
15	<p><b>Evaluar procesos cognitivos según la Taxonomía de Bloom</b>                  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><b>D. Leclercq</b>                   p. 301-328</p>
16	<p><b>Auto-evaluación con grados de certeza: un microscopio para la evaluación de los aprendizajes</b>                  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><b>D. Leclercq</b>                   p. 329-356</p>
17	<p><b>Grados de certeza y docimología: como calificar</b>                  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><b>D. Leclercq</b>                   p. 357-386</p>
18	<p><b>PdP: Pruebas de Progreso</b>                  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><b>D. Leclercq, A. Cabrera &amp; C. Van der Vleuten (U. Maastricht)</b>                   p. 387-408</p>
19	<p><b>TCS : El Test de concordancia de Script</b>                  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><b>V. Massart (ULg), A. Collard (ULg) D. Giet (ULg)</b>                   p. 409-418</p>

20	<p><b>Concebir Dispositivos de Evaluación de los Aprendizajes (DEA) al nivel de un programa</b>                  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><b>D. Leclercq,                  C. Van der Vleuten                  &amp; A. Cabrera</b>                  p. 419-430</p>
21	<p><b>Retroinformaciones (Feedbacks )</b>                  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><b>D. Leclercq,                  M. de la Fuente                  (UCH) &amp; A.                  Cabrera</b>                  p. 431-454</p>
22	<p><b>Los roles de un SMART: Servicio Metodológico de Apoyo a la Realización de Tests</b>                  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><b>D. Leclercq &amp;                  P. Detroz (ULg)</b>                  p. 455-476</p>
23	<p><b>Índices cuantitativos en Docimología</b>                  Consiste en un catálogo de conceptos útiles para tratar cuantitativamente los datos resultando de evaluaciones estandarizadas como</p> <ul style="list-style-type: none"> <li>-los tipos de categorías (nominales, ordinales, métricas).</li> <li>-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>).</li> <li>-las presentaciones gráficas de distribuciones.</li> <li>-índices de comparación o de progreso: la amplitud del efecto (AE), la ganancia relativa (GR).</li> <li>-la fiabilidad de la nota (<i>reliability</i>) al total de la prueba y el alfa de Cronbach.</li> <li>-el umbral de éxito, fijado a priori o a posteriori.</li> <li>-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</li> <li>-el análisis automática de una prueba</li> <li>-el valor heurístico de los nubes de puntos.</li> </ul>	<p><b>D. Leclercq,                  R. Roco (Chile)                  &amp;                  A. Cabrera</b>                  p. 477-543</p>
24	<p><b>Index de los autores</b>                  426 autores citados.</p>	<p><b>D. Leclercq &amp;                  A. Cabrera</b>                  p. 545-549</p>
25	<p><b>Index de los conceptos</b>                  Se puede bajar gratuitamente via <a href="http://hdl.handle.net/2268/180060">http://hdl.handle.net/2268/180060</a></p>	<p><b>D. Leclercq &amp;                  A. Cabrera</b></p>