La evaluación formativa

©Pedro Morales Vallejo, Universidad Pontificia Comillas,
Facultad de Ciencias Humanas y Sociales, Madrid
Disponible en
http://www.upcomillas.es/personal/peter/otrosdocumentos/Evaluacionformativa.pdf

Publicado en Morales Vallejo, P. (2009). *Ser profesor: una mirada al alumno*. Guatemala: Universidad Rafael Landívar, 41-98

Índice

1. La importancia de evaluaciones y exámenes 3
1.1. Los efectos no pretendidos de los exámenes 3
1.2. El poder condicionante de los exámenes en *cómo* estudia el alumno: los *enfoques de aprendizaje* 4
1.3. Distintos puntos de partida de profesores y alumnos en relación con los exámenes 6
1.4. El valor *redentor o destructivo* de nuestros modos de evaluar 7
2. Para qué evaluamos: evaluación *sumativa* y evaluación *formativa* 7
3. La evaluación formativa *continua* 14
3.1. Niveles de formalidad en la evaluación 14
3.2. Modos sencillos de evaluación formativa 15
1) *Preguntas orales* a toda la clase 15
2) Tests objetivos muy breves (*quizzes*) 16
3) Preguntas abiertas de respuesta muy breve 17
4) Los '*one minute paper*' 17
5) Trabajos en *pequeños grupos* en la misma clase 18
6) Uso de las posibilidades de las *nuevas tecnologías* 18
3.3. Observaciones sobre las pruebas cortas y frecuentes 19
1) Cómo presentar estas evaluaciones 19
2) Momento apropiado 19
a) Al comienzo de la clase 19
b) En medio de la clase 20
c) Al final de la clase 21
3) Corrección y calificación 23

4) Coherencia de estas pruebas con los objetivos del curso 24
3.4. Efectos y consecuencias de la evaluación formativa frecuente 25
4. Problemas que podemos encontrar en la evaluación formativa frecuente 27
4.1. Problemas con nosotros mismos 27
4.2. Problemas con los alumnos 28
4.3. Problemas con otros profesores 29
5. La evaluación formativa integrada en la evaluación y exámenes convencionales ... 31
5.1. Importancia del *feedback* como estrategia de enseñanza-aprendizaje 32
5.2. Funciones del *feedback* 33
5.3. Cómo dar información específica a los alumnos sobre exámenes y trabajos 35
1) Preparar claves de corrección (*rúbricas*) 35
2) Preparar un *informe* para toda la clase 36
3) Preparar *respuestas modelo* 36
4) Revisar en clase los exámenes corregidos 37
5) Corregir los trabajos dos veces (borrador y versión definitiva). 37
6. La otra cara de la moneda. La evaluación ¿la gran ocasión perdida? 39
7. Referencias bibliográficas 42

1. La importancia de evaluaciones y exámenes

Para resaltar la importancia de evaluaciones y exámenes me voy a fijar en varios puntos en los que no siempre pensamos de manera refleja, porque por lo que respecta a los exámenes nuestra atención tiende a concentrarse en cómo vamos a examinar (y habitualmente repetiremos lo que hemos hecho otros años y no hay mucho que pensar) y en cómo vamos a calificar. Estas ideas iniciales son fundamentales para hablar después de la *evaluación formativa*.

1.1. Los efectos no pretendidos de los exámenes

Antes de clarificar lo que se entiende por *evaluación formativa*[1] es iluminador pensar en lo que *de hecho vemos que sucede* después de un examen cualquiera. Podemos recordar tanto nuestra propia experiencia cuando éramos alumnos, como lo que vemos cuando nuestros propios alumnos finalizan un examen.

Cuando los alumnos terminan un examen lo dejan en la mesa del profesor, salen de la clase, se encuentran unos con otros ¿De qué hablan? ¿Qué se preguntan? Naturalmente hablan del examen que acaban de hacer. Preguntan a sus compañeros qué resultado les ha dado en un problema, cómo han enfocado la respuesta a una pregunta. O consultan sus apuntes y textos para verificar si han respondido bien, sobre todo si tienen conciencia de no haber hecho muy bien el examen. Suele haber comentarios de este estilo:

Lo tenía bien estudiado pero no para responder a 'este tipo de pregunta',

No sabía que eso era tan importante,

Creía que lo había hecho bien pero ahora entiendo en qué me estaba equivocando, etc.

Posiblemente qué es lo más importante y cómo hay que estudiar nuestra asignatura lo hemos repetido muchas veces en clase; hemos explicado con más cuidado lo que puede ser más difícil, pero… muchos alumnos no se enteran, y mucho menos si apenas han aparecido por clase. Es cuando se enfrentan a los problemas o preguntas de un examen cuando toman conciencia de sus fallos, grandes o pequeños. Incluso hay alumnos que repasando el texto después de un examen, o preguntando a sus compañeros, *entienden lo que nunca entendieron* en las explicaciones de clase. Es posible que algunos alumnos suspendidos en ese examen podrían aprobar si les examináramos media hora más tarde, o al día siguiente, porque se han dado cuenta de sus errores e incluso, como acabo de decir, realmente *han comprendido* (y no sólo repasado para poder repetirlo de memoria) lo que hasta entonces no entendían.

La reflexión importante ahora, para nosotros profesores, es ésta: con los exámenes convencionales pretendemos unas cosas (en definitiva calificar y firmar unas actas) pero a la vez suceden otras. A los alumnos les llega una valiosa información, pero… *¿No les llega demasiado tarde?*

[1] El término *evaluación formativa* fue utilizado por primera vez por Scriven (1967) tratando de la evaluación de programas (no de alumnos); en el significado habitual de hoy día y para diferenciarlo de manera explícita de la *evaluación sumativa*, lo han popularizado Bloom, Hastings y Manaus (1971)

Aunque nos referimos ahora a los exámenes convencionales, lo mismo puede decirse de trabajos escritos hechos en casa y de otro tipo de tareas; la información que les llega a los alumnos cuando estos trabajos se devuelven corregidos y calificados suele llegar ya demasiado tarde.

La pregunta que podemos hacernos es ésta: ¿Por qué no examinar *para informar a tiempo* sobre errores, sobre cómo hay que estudiar, etc., sin esperar al final? Porque la *información eficaz* para un *aprendizaje de calidad* les llega a los alumnos cuando ven los resultados de sus respuestas y ejercicios. Cualquier tipo de evaluación envía un eficaz mensaje a los alumnos sobre qué y sobre todo *cómo* deben estudiar y *además* les fuerza a una *autoevaluación,* les enfrenta con lo que saben y con lo que no saben y les orienta de manera muy eficaz en su estudio posterior. Viendo lo que sucede después de los exámenes podríamos darle la vuelta a la ecuación y decidir que *el alumno, en vez de estudiar y aprender para examinarse, debería examinarse para aprender*. Por aquí va a ir la idea, y la práctica, de la evaluación formativa. Puede parecer complicado a primera vista, pero ya veremos que no lo es tanto.

1.2. El poder condicionante de los exámenes en *cómo* estudia el alumno: los enfoques de aprendizaje

A la hora de la verdad lo que estudia el alumno depende de la evaluación esperada. *Desde la perspectiva de nuestros alumnos la evaluación siempre define el currículum real* (Ramsden, 1992). El *currículum abierto* (temas, objetivos) es el que viene en nuestros programas, pero el *currículum oculto,* el real, el que condiciona el estudio de los alumnos, es el implícito en la evaluación tal como de hecho se da y el alumno espera.

La primera obra que puso de relieve el impacto de la evaluación esperada en cómo estudian los alumnos se titula precisamente *el currículum oculto* (*the hidden curriculum,* Snyder, 1971, una investigación hecha con alumnos del M.I.T., el prestigioso Instituto de Tecnología de Massachussets); ha sido este autor quien ha popularizado el término *curriculum oculto*. Puede suceder que el *curriculum formal* (abierto, explícito) ponga el énfasis en la originalidad, en la capacidad crítica, etc., pero el *curriculum oculto,* que es el que está implícito en los exámenes, se centre básicamente en la memorización. El mismo autor ya advirtió que muchos buenos alumnos del M.I.T. vivían con conflicto una doble necesidad: su tendencia y deseo de estudiar *en serio* y las demandas de los exámenes y que aunque hay alumnos con motivación intrínseca que no dejan que los exámenes interfieran con lo que ven como importante, la mayoría se centra, como cabría esperar, en las demandas de los exámenes.

Sobre la relación entre evaluación esperada y cómo estudia el alumno hay numerosos estudios empíricos; señalamos tres que pueden servir de ejemplo.

Scouller (1998) muestra que un enfoque superficial en el estudio está asociado a buenos resultados en un examen *tipo test,* y a peores resultados en un examen de preguntas abiertas; y al revés un enfoque profundo está asociado a buenos resultados en la prueba abierta y a peores en la prueba objetiva. La autora tiene muchas otras investigaciones semejantes.

Tian (2007) también pone de manifiesto que un enfoque profundo en el estudio está asociado a buenos resultados en *trabajos para casa* y a resultados no tan buenos en los resultados de exámenes más convencionales.

Balch (2007) presenta un ejemplo muy ilustrativo. Unos alumnos de Psicología estudian 16 conceptos de Psicología; un grupo (132 alumnos) *espera* ser examinado con una *prueba objetiva* y otro grupo (122 alumnos) *espera* ser examinado con *preguntas abiertas* de respuesta corta; de hecho todos fueron examinados con una prueba objetiva con dos tipos de preguntas, en unas hay que reconocer el ejemplo correcto y en otras hay que reconocer la definición correcta. En las preguntas de reconocer ejemplos no hay diferencias entre los dos grupos, pero en las preguntas de reconocer la definición correcta estuvieron significativamente mejor los que esperaban una respuesta abierta. La conclusión parece ser que los alumnos *estudian mejor* (al menos *de otra manera*) cuando esperan una pregunta de respuesta abierta.

Este tipo de conclusiones siempre son muy matizables porque se refieren a casos muy específicos y, por ejemplo, las pruebas tipo-test pueden ser muy memorísticas y (posiblemente menos frecuentemente) también pueden comprobar habilidades intelectuales de más calibre; lo que sí parece claro es que nuestro modo de evaluar 1) puede no ser consistente con nuestros objetivos e intenciones y 2) que ciertamente la evaluación esperada condiciona cómo estudian nuestros alumnos.

Evaluaciones y exámenes condicionan no sólo *qué* estudia el alumno, sino sobre todo *cómo* lo estudia. Desde finales de los años 70 se han publicado numerosas investigaciones y estudios sobre *cómo estudian los alumnos*, porque del *cómo* estudian depende cómo se forman o cómo se deforman.

Los distintos enfoques de aprendizaje suelen reducirse, para simplificar, a dos fácilmente comprensibles: un enfoque *superficial,* más centrado en la memorización, en tener unos apuntes y esquemas claros para un repaso memorístico final y *suficiente,* y otro enfoque *profundo,* más centrado en la comprensión, en establecer relaciones, y en definitiva, en aprender *en serio* la asignatura. Cuando los alumnos estudian sobre todo poco antes de los exámenes, no vienen mucho a clase y además esto les basta a muchos o a la mayoría para aprobar, probablemente su estudio es superficial y memorístico, suficiente para no suspender e incluso para obtener una calificación alta.

Hay que dejar claro por otra parte que el estudio de *memoria* también es importante, más en unas asignaturas que en otras; el problema está cuando la memoria ocupa el lugar de la comprensión (podemos dar correctamente una definición sin entenderla) y quizás sobre todo cuando es el modo de estudio preponderante. La verdad es que muchos profesores se quejan (nos quejamos) de que los alumnos no piensan, se aferran a sus apuntes y esquemas, estudian sobre todo de memoria, memorizan procesos sin entenderlos, etc. El que los alumnos estudien sobre todo de memoria y superficialmente no es un problema de los alumnos sino de los profesores; en definitiva el cómo estudia el alumno depende de cómo pregunta el profesor.

Los *enfoques de aprendizaje* son uno de los conceptos que están influyendo más en la investigación sobre la enseñanza y el aprendizaje en los últimos años. Los *enfoques de aprendizaje* se han convertido en un tema central porque aunque se refieren directamente a *cómo estudian los alumnos,* se relacionan también directamente con los estilos de enseñanza, con el *cómo* de la evaluación, e incluso con la concepción que

tiene el profesor de su propio rol en cuanto docente[2]. Para Torre Puente (2007b) los enfoques de aprendizaje, junto con la autorregulación y la autoeficacia, forman parte de lo que denomina *una triple alianza para un aprendizaje de calidad.*

Desde la perspectiva de la *calidad del aprendizaje* lo que interesa es potenciar en los alumnos un *enfoque profundo* en su estudio y esto no depende básicamente de los alumnos sino del profesor: *los estudiantes adoptarán un enfoque superficial si perciben que es lo que la asignatura y la evaluación requieren o que es el enfoque que en mejor situación les pone para responder a las demandas de la asignatura* (Kember, 2000).

De nuevo vemos la importancia de condicionar un estudio inteligente *pero durante el curso* y antes de que los alumnos tengan que dar cuenta de lo que saben en un examen final. Que *el examen esperado condiciona el cómo estudia el alumno* es probablemente lo más importante que se puede decir de la evaluación. De este influjo del examen esperado en cómo estudian los alumnos se desprende que si no nos gusta cómo estudian nuestros alumnos, *la manera más rápida de cambiar el estilo de estudio de los alumnos es cambiar el sistema de evaluación* (Elton y Laurillard, 1979).

El alumno estudia para aprobar la asignatura (o para obtener unas notas excelentes) y de lo primero que procura enterarse es *cómo pregunta* o cómo examina el profesor. No son los consejos del profesor, ni sus orientaciones; es sobre todo su modo de evaluar lo que va a condicionar *cómo* estudia (y no sólo *qué* estudia). También puede decidir no ir a clase, porque da lo mismo ir que no ir si al final lo que importa es llevar bien aprendidos, más o menos de memoria, unos apuntes.

1.3. Distintos puntos de partida de profesores y alumnos en relación con los exámenes.

Es muy importante tener claro lo que por otra parte ya sabemos, que los puntos de partida del profesor y de los alumnos son (pueden ser) muy distintos, tal como queda expresado en la figura 1[3].

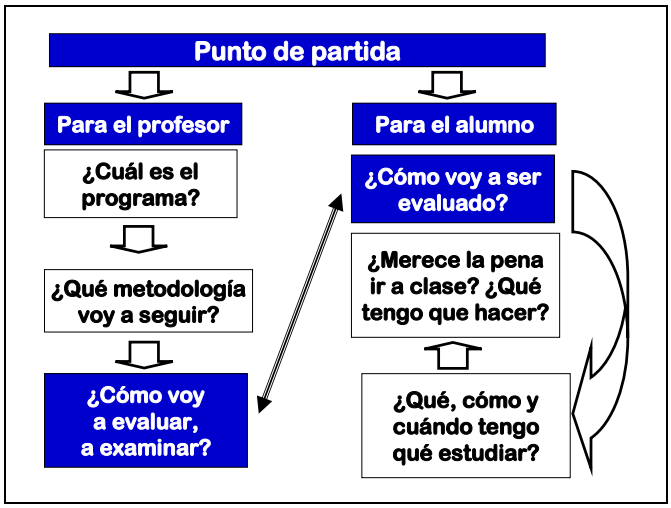

Figura 1

[2] Sobre estos enfoques hay una abundante literatura referida a la Enseñanza Superior; por ejemplo y entre otros, Ramsden (1992) y Biggs (2006).

[3] Esta figura 1 es una adaptación muy libre de la que presentan James, McInnis y Devlin, 2002.

Los profesores antes de comenzar el curso solemos pensar en nuestro programa, en nuestros temas, en cómo vamos a organizar nuestra docencia. En la evaluación en cambio solemos pensar al final del curso, cuando se acercan las fechas de los exámenes: repasamos exámenes anteriores, pensamos en lo que queremos preguntar y preparamos nuestro examen. Para el alumno la secuencia es la inversa: lo que le interesa desde el principio es saber cómo va a ser examinado; es eso y no nuestros objetivos declarados lo que va a condicionar su estudio.

La reflexión para nosotros es muy clara: hay que pensar en cómo examinamos y preguntamos, qué ejercicios y actividades vamos a proponer, desde el comienzo, procurando que la evaluación sea coherente con los objetivos del curso, con el tipo de estudio que queremos para nuestros alumnos.

1.4. El valor *redentor o destructivo* de nuestros modos de evaluar

Boud (1998) hace una interesante observación a propósito del poder condicionante de la evaluación en *cómo* estudia el alumno. La evaluación puede *redimir* unas clases más bien pobres, y al revés, puede restar eficacia a unas clases supuestamente excelentes.

a) Un *inteligente* sistema de evaluación puede tener un valor *redentor:* clases mediocres o aburridas pueden convertirse en *eficaces* estimulando un aprendizaje de calidad y quedar compensadas por un sistema de evaluación imaginativo que lleve a los alumnos a un modo de estudiar serio y comprometido.

b) También puede suceder (y sucede) todo lo contrario, unas clases inspiradas y ricas de contenido, que gustan e incluso entusiasman a los alumnos, pueden perder eficacia con un sistema de evaluación rutinario que condiciona un estudio superficial: *si evaluamos mal, nuestros alumnos aprenderán mal aunque nuestra enseñanza sea supuestamente 'excelente'* (digo *supuestamente excelentes*, porque a la hora de la verdad *la excelencia de la docencia* hay que juzgarla por *la excelencia del aprendizaje;* a clase no se va a *cautivar* o a *entretener* a los alumnos).

En un lenguaje muy informal, diría que con nuestro modo de evaluar podemos *estropear* la eficacia de unas clases excelentes, bien preparadas y que gustan a los alumnos. Y lo contrario también es verdad, un sistema de evaluación bien pensado y organizado puede compensar otras limitaciones nuestras estimulando en los alumnos un aprendizaje de calidad. A la hora de la verdad es la evaluación la que va a *contar la historia* de la calidad.

2. Para qué evaluamos: evaluación *sumativa* y evaluación *formativa*

La finalidad de la llamada *evaluación formativa* es precisamente ésta: convertir en pretendidos, e incluso con su correspondiente planificación, los efectos positivos de los exámenes convencionales que hemos recordado antes y que suelen llegar muy tarde, tal como se expresa en la figura 2.

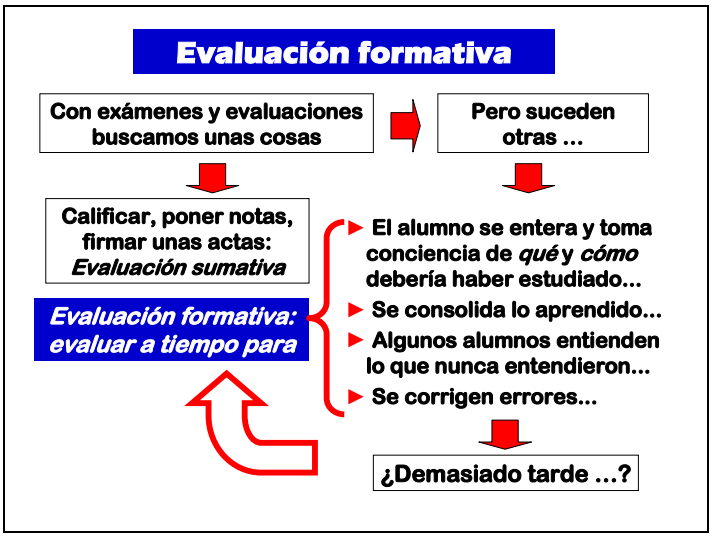

Figura 2

En esta figura 2 se expresa gráficamente lo que se busca con la evaluación formativa: que no lleguen *demasiado tarde* los buenos efectos de los exámenes finales. Si estas buenas consecuencias de los exámenes *llegan a tiempo* subirá la calidad del aprendizaje y habrá menos fracasos *sin bajar, e incluso subiendo, el nivel de exigencia* porque a una mayor información dada a los alumnos sobre su propio aprendizaje y *durante el mismo proceso de aprendizaje* se les puede ir exigiendo más.

Esta idea la iré repitiendo más de una vez: el *nivel de exigencia* (en cantidad y sobre todo calidad del aprendizaje) hay que irlo subiendo y no bajando, *pero dando facilidades*.

En la figura 3 quedan más especificadas las finalidades de la evaluación formativa que genéricamente llamamos *ayudar a aprender.*

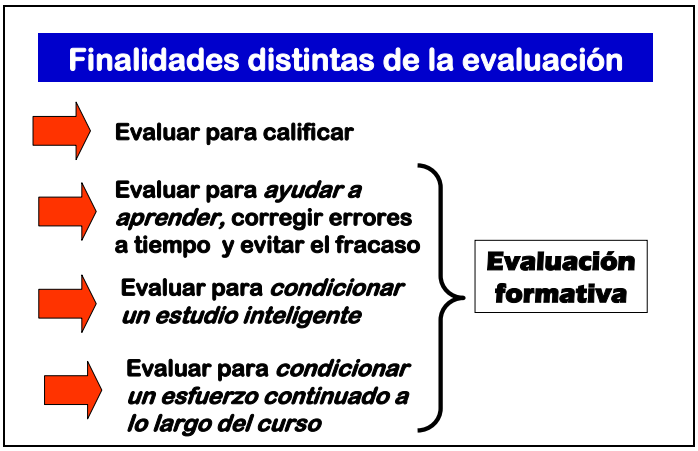

Figura 3

La cita puesta a continuación (Brown, 2006) puede parecer muy radical pero es lo que hoy día se oye desde muchas instancias en Educación Superior. Éste es ahora *el tema importante* en Educación Superior (y en otros niveles educativos) por lo que respecta a la evaluación.

> "Nuestro modo de evaluar a los alumnos en Educación Superior tiene tal impacto en el aprendizaje de los alumnos, que necesitamos repensar todo el

> proceso de programación y diseño del currículum y traer la evaluación al primer plano. Nuestros roles en cuanto profesores *deben cambiar radicalmente* de manera que podamos concentrar nuestro tiempo y energías más en la evaluación formativa y en proporcionar *feedback* a nuestros alumnos que en explicarles la materia, ya que los alumnos tienen fácil acceso a muchas fuentes de información" (Brown, 2006).

La importancia que se da ahora a la evaluación formativa en la universidad (y en todos los niveles educativos), tan nítidamente reflejada en la cita anterior, se debe en buena medida a la necesidad de condicionar en nuestros alumnos lo ya expresado en la figura 3:

a) Un *estudio inteligente,* buscando siempre la máxima calidad en el aprendizaje de nuestros alumnos,

b) Un *esfuerzo continuado a lo largo del curso*.

Para examinar si el estudio que condiciona la evaluación es *inteligente* o no lo es habría que entrar en un análisis de los tipos de pruebas, de preguntas, etc.; es un tema que no trato directamente en esta ocasión, pero que no puede quedar olvidado.

Podemos ahora concretar más las dos finalidades básicas de la evaluación; no son finalidades contradictorias pero sí es frecuente (por no decir habitual) que nos centremos en una sola de las dos.

Cuando pensamos en la evaluación solemos centrarnos en el *qué* y en el *cómo* vamos a evaluar o examinar.

El *para qué* evaluamos y examinamos lo solemos dar por supuesto: ponemos exámenes y encargamos trabajos para ver *dónde están* nuestros alumnos, ver *cuánto* saben, ponerles una *calificación* y firmar unas actas dando por finalizado el proceso.

La tradición heredada (y las normas académicas que nos gobiernan, a veces muy a nuestro pesar) hace que pensemos siempre o casi siempre en la evaluación *para calificar*. Sin embargo es en este *para qué* evaluamos donde las cosas están cambiando, tanto en el nivel universitario como en cualquier otro nivel educativo. Por lo que respecta a la Universidad ya nos van recordando con insistencia que no podemos seguir entendiendo los exámenes y evaluaciones como el *final de un proceso,* tal como ha sido, y en buena medida sigue siendo, la práctica más tradicional (citas interesantes e ilustrativas: Boud, 1998; Yorke, 2003, Brown, 2004, 2006; Smyth, 2004)[4].

Ahora se distinguen por lo tanto, y de manera muy clara, dos enfoques de la evaluación que corresponden a *dos finalidades claramente distintas* (aunque pueden integrarse en una misma evaluación). Hablo de *enfoques* porque es un término muy genérico en el que caben muchas modalidades.

Más que en los modos y técnicas de evaluación, es muy importante en primer lugar pensar y caer en la cuenta de las *finalidades* (y simultáneamente pensar en nuestras actitudes e intenciones) que puede tener la evaluación.

a) La *evaluación sumativa* cuya finalidad es, en última instancia, *calificar* a los alumnos según el nivel que hayan alcanzado. Ésta es la práctica habitual, lo que

[4] Noto incidentalmente que toda esta orientación sobre la finalidad de la evaluación que hoy día se pone muy de relieve nos viene del ámbito anglosajón.

nosotros mismos hemos vivido como alumnos: el examen señala *el final del proceso* o de una parte del proceso (en exámenes parciales y liberatorios). Examinamos al final porque tenemos que calificar a nuestros alumnos. Naturalmente esta finalidad es legítima y necesaria; obligación nuestra es certificar el nivel de aprendizaje de nuestros alumnos.

b) La *evaluación formativa* cuya finalidad no es en principio calificar sino *ayudar a aprender, condicionar un estudio inteligente* y *corregir errores a tiempo.* Esta evaluación formativa no es un punto final sino que está *integrada en el proceso de enseñanza-aprendizaje.* Por utilizar una analogía agrícola, una cosa es recoger la cosecha (*evaluación sumativa*) y otra distinta cuidar y fertilizar nuestro campo *de manera eficaz* para que nos dé una buena cosecha (*evaluación formativa*).

En el enfoque formativo de la evaluación distinguimos por lo tanto dos características relacionadas entre sí; el *para qué* y el *cuándo* de la evaluación.

1) La *finalidad* de la evaluación.

No se trata ahora de verificar resultados *finales*, sino de *ayudar* al alumno en su tarea de estudiar y aprender. Buscamos el éxito del alumno, que es también *nuestro éxito profesional en cuanto docentes,* y no el fracaso. Lo mismo que se hace en otros procesos, no se espera hasta el final para descubrir que los aparatos no funcionan o que los automóviles no pueden arrancar.

Parece algo muy elemental pero con frecuencia pasa desapercibido que el producto de nuestra *tarea profesional* en cuanto docentes es *la calidad del aprendizaje* de nuestros alumnos. También puede parecer una obviedad el decir que *el fracaso no es un indicador de éxito,* aunque en algunas culturas educacionales un determinado porcentaje de fracasos se interpreta (y a veces *casi se exige*) como un indicador de calidad y de nivel de exigencia.

Sobre el fracaso de muchos o de algunos alumnos como indicador del nivel de exigencia ha tenido un eco notable el libro de André Antibi (2005) *la constante macabra;* esta *constante macabra* es el porcentaje de suspensos, más o menos fijo, que muchos o algunos profesores mantienen como necesario e inamovible para mantener su *prestigio* de profesores exigentes. Antibi se refiere a la educación pública francesa en sus niveles primario y secundario, pero éste es un tema que nos invita a todos a una buena reflexión.

Es claro que los profesores no somos responsables *de todo*…; hay muchas variables que no controlamos y el fracaso de muchos alumnos puede tener muchas causas y concausas que no tienen que ver con nosotros, pero en cualquier caso el fracaso no es indicador de nada de lo que en principio se pueda presumir y a los profesores nos toca examinar lo que tiene que ver con nosotros. Volvemos sobre esto al final, al hablar de *la otra cara de la moneda*.

2) El *momento* de la evaluación

La evaluación formativa está *integrada en el proceso de enseñanza y aprendizaje,* es decir, no se limita a certificar con una nota un determinado nivel de aprendizaje. No va a haber *un momento,* sino todos los que sean necesarios y posibles; eso es lo que se quiere decir cuando se afirma que la evaluación formativa está *integrada* en el proceso de aprendizaje. La evaluación puede ser de hecho un excelente método de enseñanza, y

tan importante como las explicaciones del profesor. Como nos recuerda Yorke (2003) al poner de relieve la importancia de la evaluación formativa en la Universidad, *el aprendizaje depende del conocimiento de los resultados en un tiempo y en una situación en el que este conocimiento puede ser utilizado para corregir los propios errores* (Yorke, 2003).

El título de este artículo (Yorke, 2003) se podría traducir con mucha libertad de esta manera: *algo 'se está moviendo' en la concepción y en la práctica de la evaluación en educación superior.* Efectivamente *el aprendizaje del alumno* está pasando a un primer plano y este *primer plano* parece que se va a manifestar en los futuros *rankings* de las Universidades, de manera parecida a lo que ya se hace en los niveles preuniversitarios[5].

Bain (2006:47), en su excelente obra sobre *lo que hacen los mejores profesores universitarios,* recuerda una de las conclusiones de una investigación en la Universidad de Harvard: las características de las clases mejor valoradas por los alumnos[6] incluían *gran exigencia,* pero *repletas de oportunidades para revisar y mejorar* su trabajo antes de ser calificado y, por lo tanto, para aprender de sus errores en el proceso.

Podemos añadir una tercera característica de la evaluación formativa: caben *procedimientos más sencillos* que los que utilizamos en los exámenes tradicionales como exponemos más adelante.

En la figura 4 tenemos de manera muy esquemática las características de los dos enfoques que se podría completar con un tercer enfoque, el enfoque *diagnóstico* que puede ser de gran interés antes de comenzar determinadas asignaturas.

	Formativa	**Sumativa**
Funciones **Para qué …**	***Facilitar el aprendizaje*** **Informar** ***Al profesor*** **Ritmo, problemas de aprendizaje …** ***Al alumno*** **Corregir errores a tiempo,** **Cómo hay que estudiar,** **Qué es lo importante** **Cuál es el nivel de exigencia**	**Calificar** **Certificar**
Cuándo …	**Con cierta frecuencia** **Cuando sea oportuno** ***Evaluación integrada en el aprendizaje***	**Tiempos designados**
Cómo …	**Caben métodos sencillos, más informales y variados**	**Métodos habituales**
Calificación	**No, o peso menor, o tener en cuenta sólo si se ha hecho o no** **Su finalidad es 'otra'**	**Sí**

Figura 4: Enfoques en la evaluación

[5] La OECD (*Organization for Economic Cooperation and Development*) va a establecer un nuevo *ranking* de universidades (en 2010) teniendo en cuenta no el *input* (presupuesto, publicaciones e investigaciones de los profesores, medios tecnológicos, etc.) sino el *output* entendido como el aprendizaje de los alumnos (en *contenidos básicos* de las distintas disciplinas y en *habilidades transferibles*) medido con exámenes a muestras aleatorias, tal como ya se hace en los niveles obligatorios de la enseñanza (the Economist, 17th Nov. 2007).

[6] Entrevistas en profundidad, de 2 o 3 horas, a 2000 alumnos; Bain cita a Light (1990) autor de *The Harvard Assessment Seminars.*

En un simple cuadro no se puede encerrar todo lo que se puede decir sobre estos dos enfoques o finalidades, y tampoco se trata siempre y necesariamente de una dicotomía (aunque en un principio puede ser útil presentarlo así, como en la figura 4).

Los exámenes convencionales (parciales, finales; lo que convencionalmente se denomina *evaluación sumativa*), lo mismo que los trabajos para hacer en casa, prácticas, etc., que son calificados y tienen su peso la nota final pueden y deben contribuir a corregir errores (aunque no siempre se llega *a tiempo*) y a mejorar los hábitos de estudio.

Aunque en un sentido muy restringido los dos enfoques en la evaluación se distinguen claramente por el *uso prevalente* que se hace de la evaluación (evaluación formativa para informar y mejorar el aprendizaje y evaluación sumativa para calificar) en la práctica los dos usos y finalidades pueden estar presentes en la misma evaluación. Esquemas como el de la figura 4 ayudan a pensar en lo que quizás no pensamos habitualmente, aunque esta clasificación de los dos enfoques no hay que entenderla de manera muy rígida; el tema de la evaluación dista de ser una ciencia exacta.

Como no se trata de una dicotomía radical (evaluación *o* formativa *o* sumativa) y también la evaluación sumativa (pruebas y trabajos convencionales, con nota) puede tener una finalidad formativa, cabe definir la evaluación formativa de una manera más amplia, teniendo en cuenta sobre todo que no hay una definición clara y compartida de lo que es evaluación formativa (Dunn y Mulvenon, 2009)[7]:

> La evaluación formativa consiste en *un abanico de procedimientos de evaluación, formales e informales, integrados en el proceso de enseñanza-aprendizaje y orientados a modificar y mejorar el aprendizaje y comprensión de los alumnos*.

Otra definición, también muy genérica pero clarificadora es ésta:

> *La evaluación formativa es el proceso utilizado por profesores y alumnos durante el período de enseñanza-aprendizaje que aporta la información necesaria (feedback) para ir ajustando el proceso de manera que los alumnos consigan los objetivos propuestos*. (Melmer, Burmaster y James, 2008)[8].

El elemento *formal* (que hace que la evaluación sea *precisamente* formativa) está *en aportar información (feedback) durante el proceso* para mejorar los resultados finales.

La evaluación formativa no es *un añadido,* sino una parte integral del proceso, tanto como lo son las explicaciones del profesor. Cuando se habla de *proceso* lo que se quiere decir que es que no hay un tipo de prueba en concreto que *acapare* el término de evaluación formativa, sino que caben muchas estrategias distintas (que apuntamos más adelante).

[7] Crooks (2001) que toma esta *definición-explicación* de lo que es evaluación formativa de un documento del Ministerio de Educación de Nueva Zelanda

[8] Literalmente los autores se expresan así *Formative assessment is a process used by teachers and students during instruction that provides feedback to adjust ongoing teaching and learning to improve students' achievement of intended instructional outcomes.*

También es importante caer en la cuenta de que no se trata de algo que concierna exclusivamente al profesor porque tiene que implicar también a los alumnos (aunque el implicar a los alumnos también va a ser tarea del profesor).

En la figura 5 sobre modalidades de la evaluación formativa se incluyen también los exámenes finales convencionales, pero dejando claro que su eficacia como ayuda al aprendizaje dependen de la *calidad* y *oportunidad* del *feedback* o información de retorno que se dé a los alumnos.

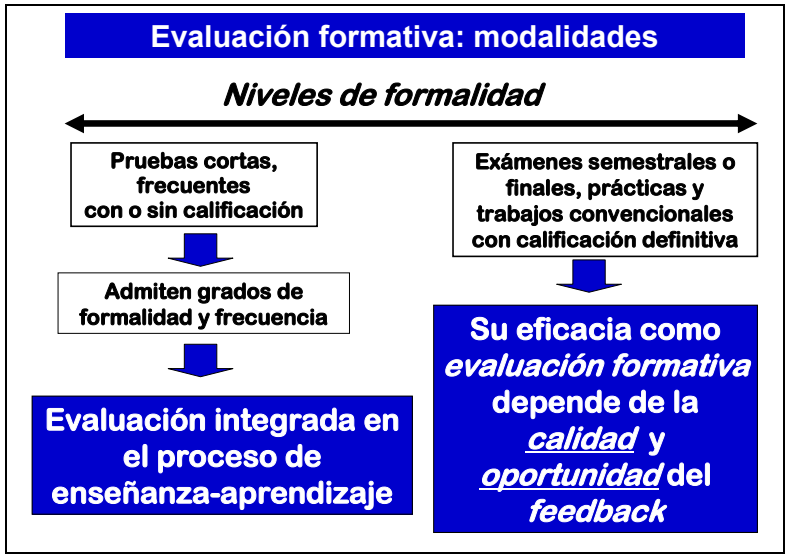

Figura 5

La evaluación formativa realmente eficaz es la integrada en el proceso de enseñanza-aprendizaje. En la figura 5 se habla de *pruebas* y *exámenes* que son términos convencionales pero quizás un tanto restrictivos en su significado habitual; habría que entender también ejercicios prácticos hechos en clase, tareas para casa, etc.

Lo que es importante ahora es caer en la cuenta de que *a finalidades distintas 'pueden' corresponder tiempos y maneras de evaluar distintas*. Aunque, y esto hay que dejarlo muy claro, toda evaluación que de alguna manera es *final* (como un trabajo o un examen parcial o final, con una nota) *debería ser* también una eficaz ayuda al aprendizaje, aunque de hecho no siempre lo es. El carácter formativo de la evaluación sumativa habitual dependerá del *feedback* o información de retorno que demos a nuestros alumnos, cómo corregimos y comunicamos resultados, etc. Sobre este punto tratamos más adelante porque, si nos lo proponemos, toda evaluación sumativa puede (y ciertamente *debería*) ser también formativa. En lo que quizás no pensamos es en que puede haber evaluaciones *solamente* formativas (es decir, sin ningún tipo de calificación porque su finalidad es *otra*).

Esta concepción, y sobre todo la práctica de una evaluación formativa basada en pruebas frecuentes y *muy integradas en el mismo proceso de aprendizaje,* puede ponernos a la defensiva si pensamos en el trabajo extra que puede suponer para los profesores una evaluación más frecuente, aunque no se trata de poner y corregir exámenes convencionales todos los días. Hay modos y maneras, algunos muy simples, de llevar a cabo esta evaluación formativa de manera continua (los vemos más adelante), pero el primer paso es convencernos no ya de su utilidad, sino de su necesidad si de verdad buscamos el éxito de nuestros alumnos y *nuestro propio éxito*

como profesionales de la docencia. Por eso es tan importante *traer a la conciencia* los efectos en los alumnos de exámenes y evaluaciones comentados al principio. No son los métodos e instrumentos habituales de evaluación los que en muchos casos deberían cambiar, sino la filosofía subyacente de la evaluación y los objetivos (el *para qué*) de su uso y aplicación: *los alumnos deberían aprender a través de la evaluación en vez de aprender para ser evaluados* (Mowl, 1996).

El objetivo último de un profesor universitario no es examinar, sino hacer de sus alumnos buenos profesionales en el área de su competencia; de manera análoga se podría decir lo mismo de cualquier nivel educativo. *El producto final en Educación Superior (y en cualquier nivel educativo) es un aprendizaje de calidad, no un examen*.

3. La evaluación formativa *continua*

Cuando se habla de evaluación *continua* o de *evaluación integrada en el aprendizaje* está claro que no se quiere decir exámenes todos los días, pero sí con la suficiente frecuencia como para que sean *eficaces* en el aprendizaje, al menos por lo que se refiere a las modalidades más sencillas de evaluación formativa. En muchos casos el término *examen,* por sus connotaciones, puede ser inadecuado; no toda evaluación es necesariamente un examen en el sentido habitual del término.

En la figura 5 se hace la distinción entre pruebas *cortas y frecuentes* (*con* o *sin* calificación) y exámenes finales o al menos parciales (para entendernos, *exámenes convencionales*). Dando por hecho que esta distinción no es ni tiene que ser nítida, vamos a tratar en primer lugar de la *evaluación frecuente,* con énfasis en su dimensión de *ayuda durante el proceso de aprendizaje* y que admite modalidades muy sencillas. Pero antes conviene recordar otras formas de evaluación más informales que conviene *traer a la conciencia*.

3.1. Niveles de formalidad en la evaluación

En la figura 5 al distinguir entre pruebas más cortas e informales y exámenes más convencionales estamos dejando fuera de nuestra atención otras formas de evaluación que conviene tener presentes y que quedan explicitadas en la figura 6. Lo que incluimos en esta figura son otras formas de evaluación que no solemos conceptualizar como tales, pero que pueden ser muy importantes.

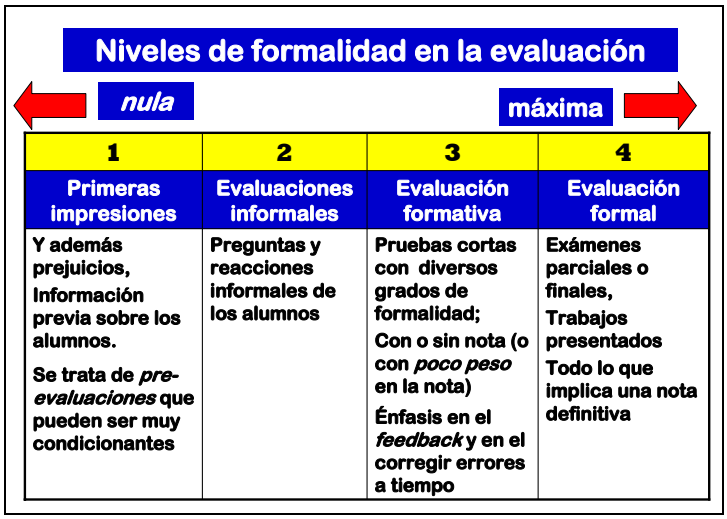

Figura 6

En el máximo nivel de *informalidad* tenemos lo que ya sabemos (o creemos que sabemos) de los alumnos, tanto de alumnos individuales como de grupos enteros. Primeras impresiones, prejuicios hacia determinados tipos o grupos de alumnos, la información que nos llega por *tradición oral*, etc. No trato aquí de estas formas de evaluación (es evaluación en la medida en que *se trata de o nos llevan a juicios valorativos*) pero pueden ser muy condicionantes para la relación profesor-alumno, nivel de exigencia, etc. Realmente éste es un *tema distinto*[9] pero conviene recordarlo porque en la vida lo formal y lo informal son inseparables.

El segundo nivel de formalidad tiene ya más que ver con el tema específico de la evaluación formativa tal como la estamos entendiendo. Las *preguntas espontáneas* que nos hacen los alumnos dentro y fuera de clase, sus *reacciones verbales y no verbales*…; con mucha frecuencia, y sin necesidad de pruebas de ningún tipo ya sabemos lo que nuestros alumnos están o no están entendiendo, qué confunden con qué, qué ejercicio adicional están necesitando, qué objetivos de aprendizaje hay que recordarles, etc. La relación habitual e informal con los alumnos, sobre todo lo que *vemos* durante las clases, es una fuente importante de información a la que debemos estar atentos. Esta atención va a depender de nuestras actitudes personales, de nuestro interés personal por el aprendizaje de nuestros alumnos.

3.2. Modos sencillos de evaluación formativa

En la figura 5 al explicitar de manera esquemática el *cómo* de la *evaluación formativa* vemos que *caben métodos sencillos, más informales y variados*. También queda indicado en la figura 5 (y en la figura 6); hay grados en formalidad y frecuencia. Conviene pensar en lo que podríamos hacer y no hacemos porque nuestra atención, por lo que respecta a la evaluación, está centrada en los exámenes convencionales de un tipo u otro (parciales, finales, trabajos que tienen que entregar).

Son muchas las maneras de llevar a cabo la evaluación formativa de manera casi continua y sin mucho trabajo adicional para el profesor. Pueden verse enumeradas diversas estrategias en Boston (2002) y Jenkins (2003) (ambos disponibles en Internet) y a nosotros se nos pueden ocurrir otras.

La evaluación formativa puede adquirir múltiples modalidades; entre otras el profesor puede estructurar situaciones que permitan la *autoevaluación personal y grupal* como nos recuerda Yorke (2003) en un excelente artículo sobre evaluación formativa en la universidad. La evaluación formativa puede ir desde lo muy informal hasta lo muy formal; puede incluso ser muy ocasional y sin embargo dar un apoyo esencial al aprendizaje. Una buena referencia es también Angelo y Cross (*Classroom assessment techniques. A handbook for college teachers*, 1993), que describen numerosos métodos de evaluación formativa, algunos muy originales, utilizados por profesores en la Universidad.

Enumeramos algunas *estrategias sencillas* de evaluación formativa.

1) *Preguntas orales* a toda la clase.

Es lo más sencillo aunque quizás no siempre lo más eficaz (es más eficaz que cada alumno se responsabilice de su propia respuesta por escrito, aunque unas cosas no

[9] El tema de la *relación profesor-alumno en el aula* lo trato con más extensión en otro lugar (Morales, 2002) y en otro documento sobre *el profesor educador* (Morales, 2009).

quitan otras). El lanzar una pregunta a toda la clase para que respondan los que quieran puede parecer muy infantil, pero puede funcionar muy bien en cualquier nivel[10].

Estas preguntas orales hechas en clase, para que sean a la vez un buen método de aprendizaje y generen en los alumnos la reflexión deseada deben cumplir algunas condiciones.

a) Tienen que ser *buenas preguntas,* que incidan en lo que realmente el profesor juzga que es importante conocer y sobre todo *entender*. Las preguntas orales que hacemos en clase suelen ser muy espontáneas y responden a nuestras intuiciones más inmediatas; cabe también pensar de antemano e incluso anotar qué preguntas puede ser conveniente hacer en clase, sin limitarnos a preguntas que se pueden responder correctamente de memoria.

b) Hay que hacer estas preguntas siendo conscientes *desde dónde se hacen y para qué se hacen:* para corregir errores a tiempo, para dejar claro lo que es importante, para estimular la reflexión, etc., es decir, para facilitar un estudio y un aprendizaje de calidad. Esta *conciencia* de la finalidad de estas preguntas ayuda a seleccionarlas bien.

c) Los alumnos deben darse cuenta de la finalidad de estas preguntas para que se las tomen más en serio; hay que *explicitar* de manera clara por qué se hacen estas preguntas.

Las preguntas orales a toda la clase pueden ser un buen recurso didáctico, siempre lo han sido, pero, sin dejarlas de lado y si se trata de una evaluación formativa, posiblemente es preferible que *cada alumno* tenga que dar algún tipo de respuesta escrita y con su firma, de manera que cada uno se perciba como *sujeto responsable* de su propia respuesta.

2) Tests objetivos muy breves (*quizzes*).

Estos tests son fáciles de preparar; pueden incluso ser del tipo *verdadero-falso*. El formato *verdadero-falso* puede ser discutible en un examen final (se puede responder correctamente al 50% de las preguntas respondiendo al azar), pero como *sondeo rápido* a la clase puede ser cómodo y útil. Y con este formato caben preguntas certeras que obligan a reflexionar.

Menciono el *verdadero-falso* porque nos puede parecer más sencillo, pero pueden ser preguntas convencionales de elección múltiple. Si *adivinan* la respuesta correcta, da lo mismo; no hay nota; se trata en definitiva de *ejercicios de autoevaluación* que dan pié a aclaraciones y explicaciones adicionales, facilitan que los alumnos intervengan y pregunten sus dudas, etc. Una vez preparadas sirven de un año para otro con pequeñas modificaciones.

Cabe por supuesto *más formalidad* en estas pruebas relativamente cortas y frecuentes (*evaluación continua*), y pueden ser calificadas, pero si la finalidad fundamental que las motiva es *corregir errores a tiempo, enseñar a estudiar,* etc., parece que no deberían tener *mucho peso* en la nota final. Posiblemente es preferible

[10] Aquí es oportuno recordar a Platón como el maestro que a través de preguntas y respuestas (orales, naturalmente) hace pensar a sus discípulos y estimula su capacidad crítica. En Torre Puente (2007a, en el anexo) se dan unas *pautas para regular la conducta docente de hacer preguntas en clase.*

que no haya calificación para evitar la tensión y otros problemas derivados de la evaluación frecuente, como pueden ser las quejas de otros profesores (los alumnos pueden dejar de ir a otras clases cuando se acerca una de estas pruebas) y dificultades con la administración cuando estas pruebas tienen (o *parece* que tienen) la *calidad* de examen (en otro momento tratamos de los problemas que nos puede crear la evaluación frecuente, con los alumnos y con otros profesores). La razón fundamental al menos para cuestionar o plantearse el que estas pequeñas pruebas sean calificadas, es que, como venimos diciendo, su finalidad es *otra*[11].

3) Preguntas abiertas de respuesta muy breve

Puede tratarse de una sola pregunta, de respuesta muy breve en la que se pide a los alumnos que resuman la idea principal explicada en clase, que definan un concepto, que indiquen el principio aplicable en un caso o problema, a qué criterio de clasificación corresponde un ejemplo, dar la propia opinión sobre un caso controvertido, etc. El mero hecho de tenerse que *definir por escrito* obliga al alumno a pensar y expresarse con claridad.

Estas preguntas pueden hacerse al comienzo de la clase, en medio de la clase (sobre lo que se acaba de explicar) o al final de la clase.

Este tipo de *mini-pruebas* (abiertas o cerradas) admiten muchas variantes. Se pueden recoger (firmadas), corregir (sin poner nota) y comentar los resultados en la clase siguiente. En este caso, si no hay calificación, sí se puede anotar quiénes respondieron, y éste es un dato (el mero hecho de haber respondido, no el haber respondido bien o mal) que, a juicio del profesor, se puede tener en cuenta al final (refleja asistencia a clase, interés, participación).

En general pruebas cortas y frecuentes son más eficaces que pruebas más largas e infrecuentes; el nuevo aprendizaje se debería evaluar sin esperar más de una semana (recomendaciones de Black y Wiliam, 1998). Todo esto es fácilmente posible sin pensamos en formas rápidas y eficaces de evaluación que por otra parte no suponen mucho tiempo o esfuerzo extra para el profesor.

4) Los '*one minute paper*'

Las breves pruebas pensadas para el final de la clase y conocidas como *one minute paper* pueden ser de hecho un recurso fácil y eficaz de evaluación formativa. Más que de *pruebas* en sentido propio, en su forma más tradicional suele tratarse de dos breves preguntas de este estilo: *qué es lo que más te ha interesado en esta clase* y *qué es lo que todavía te ha quedado confuso* y necesita una mayor aclaración. Algunas de las sugerencias anteriores pueden caer en esta categoría, que admite muchas modalidades[12]. No tiene mucho sentido calificar respuestas que más que *ciencia* reflejan *sentimientos;* pueden ser incluso anónimas; en cualquier caso lo que sí se podría tener en cuenta es el mero hecho de haber respondido.

[11] *There is some evidence that students gain the most learning value from assessment when feedback is provided without marks or grades. Where marks are provided, they often seem to predominate in students' thinking, and to be seen as the real purpose of the assessment* (Crooks, 2001).

[12] Más información en *El 'one minute paper'* (Morales, 2008); procedimientos semejantes pueden verse en Angelo y Cross (1993) (*Classroom assessment techniques. A handbook for college teachers*).

5) Trabajos en *pequeños grupos* en la misma clase

Estas tareas grupales (en parejas, o grupos de tres tal como están sentados en clase) pueden ser de corta duración; los alumnos pueden comentar un tema, concretar sus dificultades de comprensión, *responder juntos* a unas pocas preguntas, resolver un problema o caso, *corregir en común* sus propios ejercicios o revisar en común las correcciones hechas por el profesor, etc. Hay muchas maneras de organizar breves sesiones de estudio colaborativo (o de discusión y reflexión), como puede ser también invitar a los alumnos a pensar en la respuesta a una pregunta, o a escoger (y justificar) una respuesta entre las varias ofrecidas por el profesor.

Boston (2002) propone algunas actividades muy sencillas, aptas para todos los niveles, que cumplen con la finalidad de la evaluación formativa (corregir errores, ayudar a aprender) y que además hacen la clase más activa, por ejemplo:

- Invitar a los alumnos a discutir en parejas o pequeños grupos sobre un tema y que después un representante de cada grupo exponga lo que se ha dicho en su grupo.
- Presentar posibles respuestas a una pregunta y que los alumnos voten cuál es la respuesta correcta.
- Pedir a los alumnos que escriban la respuesta a una pregunta y luego pedir a algunos que la lean en voz alta.

Los ejercicios y el estudio *colaborativo,* en pequeños grupos, son muy eficaces; ya va habiendo una abundante investigación sobre este tema, referida también a Educación Superior. Hay incluso tesis doctorales sobre la eficacia de la corrección de trabajos en pequeños grupos en la Universidad (Olalla, 2005) y publicaciones que recogen diversas experiencias en distintas carreras universitarias (Watts y García Carbonell, 2006, sobre *evaluación compartida,* disponible en Internet)[13]. Una referencia importante (entre muchas otras) sobre la eficacia del estudio colaborativo es Johnson, Johnson y Stanne (2000), de la Universidad de Minnesota, que presentan una síntesis y comentario de 164 estudios sobre *aprendizaje cooperativo,* muchos referidos a la Universidad (también disponible en Internet).

6) Uso de las posibilidades de las *nuevas tecnologías*

En estos tiempos no podemos olvidar las múltiples posibilidades que nos ofrecen las *nuevas tecnologías*. Hay buenas experiencias de evaluación formativa por ordenador que admiten múltiples modalidades: preguntas de autoevaluación de elección múltiple con *feedback* inmediato, estudio de casos, *weblogs,* FAQ (*frequently asked questions*), etc. (Jenkins, 2004, presenta diversas posibilidades). Existen ya numerosas experiencias de las que se puede aprender (por ejemplo Dopper y Sjoer, 2004, sobre evaluación formativa *online* en ingeniería).

Los alumnos pueden incluso colgar en la página Web de la asignatura o departamento trabajos breves que los demás alumnos pueden ir viendo, añadir comentarios, etc. (por ejemplo, Feldman, 2005)[14]. Breves pruebas objetivas puestas en

[13] No sobra el caer en la cuenta de que nuestro propio trabajo docente y nuestras experiencias con la evaluación pueden ser también una oportunidad para investigar y publicar.

[14] En el programa de la asignatura (*Being a New Teacher;* University of Massachusetts School of Education), se especifica que los alumnos deben hacer un breve trabajo *semanal* (dos páginas comentando una lectura), colgarlo en

la Web pueden tener la corrección programada y el profesor puede conocer en todo momento los resultados que va obteniendo cada alumno, con el consiguiente ahorro de tiempo[15].

También es verdad que a propósito de las nuevas tecnologías no conviene olvidar lo que ya hace años afirmaba Clark (1983): *la tecnología educativa influye en la calidad de la educación lo mismo que la camioneta que transporta alimentos influye en la nutrición; depende de la calidad de los alimentos que transporta*.

3.3. Observaciones sobre las pruebas cortas y frecuentes

Con un poco de imaginación y revisando experiencias de otros no es nada complicado *integrar en la misma clase* tareas de evaluación que pueden ser distintas en asignaturas distintas. Añadimos algunas observaciones sobre el cómo y cuándo de estas pruebas.

1) Cómo presentar estas evaluaciones

No es necesario fotocopiar una hoja con estas preguntas aunque a veces será lo más conveniente. Caben procedimientos económicos y sencillos, como pueden ser distribuir papeletas; los alumnos ponen su nombre y la fecha y responden a la pregunta o preguntas que se les presentan en alguna de estas formas:

a) Se dictan las preguntas que también se pueden ir escribiendo en la pizarra. Esto es muy práctico cuando las preguntas casi se han improvisado o se le ocurren al profesor durante la misma clase.

b) Se presentan con un retroproyector o en una presentación Power Point. En este caso obviamente hay que prepararlas de antemano, pero el trabajo de un día sirve para otras ocasiones.

2) Momento apropiado

Estas breves pruebas, ejercicios o preguntas caben en diversos momentos, al comienzo de la clase, hacia la mitad de la clase o al final. Cada momento tiene sus peculiaridades y efectos y se pueden poner en un momento u otro según los objetivos del profesor. Conviene caer en la cuenta de que *según para qué* no es lo mismo un momento que otro.

a) ***Al comienzo de la clase***

Cuando *se comienza* la clase con alguna pregunta que hay que responder por escrito la finalidad suele ser alguna de estas (o más de una):

a) Verificar si han leído lo previamente asignado antes de comenzar un nuevo tema o al menos estimular que realmente hagan estas lecturas previas,

b) Encauzar el comienzo de la explicación

c) Estimular una discusión posterior.

---

la Web de la asignatura, leer lo que han puesto los demás y responder brevemente a uno o dos compañeros; estas lecturas serán centrales en la discusión de la clase siguiente (Feldman, 2005).

[15] Daniel y Broida (2004) comparan los resultados en los exámenes finales de tres grupos según hayan respondido a estas pruebas (*quizzes*) en clase, en la Web o no las hayan tenido. Las medias mayores en los exámenes corresponden al grupo que respondió en clase. También enumeran las *estrategias* que utilizan los alumnos (según ellos mismos) para responder correctamente en la Web y que hay que tener en cuenta al menos si estas pruebas van a ser calificadas.

Puede parecer extraño comenzar la clase con un breve ejercicio escrito, pero no faltan ejemplos que muestran su eficacia (Connor-Green, 2000; Leeming, 2002; Narloch, 2006; Padilla-Walker, 2006).

b) *En medio de la clase*

Cuando se hacen estas preguntas (con breve respuesta escrita) *en medio de la clase,* interrumpiendo la lección magistral, se verifica sobre la marcha si los alumnos van entendiendo y se crea un espacio de reflexión.

Dos ejemplos de investigaciones sobre estas preguntas hechas una vez comenzada la clase, y que incidentalmente también son ejemplos de un tipo de investigación útil y *publicable* muy asequible a cualquier profesor que se lo proponga.

a) Dravick, Weisberg, Paul, y Bubier (2007) presentan un interesante experimento sobre la eficacia de responder a una pregunta *por escrito* (respuesta abierta, del tipo *one minute paper,* y *sin calificación*). La muestra es de 978 alumnos (cuatro profesores trabajando en equipo, varias secciones, uniendo dos semestres). Las clases son asignadas aleatoriamente a dos condiciones; en una condición los alumnos tienen que escribir durante cinco minutos (*writing condition*) expresando su opinión sobre un tema controvertido; en la otra condición los alumnos no tienen que escribir sino *pensar* (también durante cinco minutos; *thought condition*) sobre el mismo tema. En ambos casos el profesor ha introducido un tema previamente. En ambos grupos sigue una discusión de toda la clase durante 10 minutos (la duración de la actividad está en torno a los 15 minutos).

En exámenes posteriores (elección múltiple, exámenes parciales) los alumnos en la *writing condition* obtuvieron significativamente mejores resultados que los alumnos en la *thought condition* (sin diferencias entre los sexos y controlando asistencia a clase y el nivel previo). Conclusión: cinco minutos *escribiendo* a la semana (un total de 45 minutos en todo el semestre) son más eficaces que cinco minutos *pensando*

b) En la investigación de Butler, Phillmann y Smart (2001) se verifica el efecto de las preguntas hechas en clase cuando después se hacen las mismas preguntas en el examen (no exactamente las mismas pero sí referidas al mismo contenido). En dos clases (y en 12 ocasiones) una vez explicado un concepto nuevo, los alumnos responden por escrito a una pregunta de respuesta abierta sobre lo que se acaba de explicar; en el examen se encontrarán preguntas semejantes, pero con otro formato (tipo test). Las preguntas respondidas en clase no son las mismas para los dos grupos pero los exámenes sí son iguales. Cada grupo supera al otro en aquellas preguntas que corresponden a las que respondieron en clase.

La conclusión de este tipo de estudios (en términos generales) es tan obvia como poco tenida en cuenta; si queremos que los alumnos entiendan de verdad algo importante no hay que esperar al día del examen para preguntarlo y verificar si de verdad se han enterado[16].

Una utilidad específica de breves ejercicios hechos hacia la mitad de la clase es despertar de nuevo la atención de los alumnos; puede ser un excelente procedimiento

[16] En este tipo de experimentos didácticos es habitual también preguntar a los alumnos (mediante un cuestionario sencillo) su parecer sobe su utilidad y cualquier otro tipo de efecto pretendido o posible (asistir más a clase, consultar más la biblioteca, etc.)

para dinamizar la lección magistral (Denman, 2005, *How To Create Memorable Lectures*).

La figura 7 (adaptada de Biggs, 2006) muestra lo que es experiencia común: que a los 10 o 20 minutos baja drásticamente la atención de los alumnos (como la nuestra en conferencias y actos académicos)[17]. Es muy difícil conseguir que *todos* los alumnos mantengan la atención *todo* el tiempo. Mucho tiempo valioso de clase se pierde sencillamente porque los alumnos, incluso los que aparentemente están *atentos*, tienen la mente *en otro lugar*.

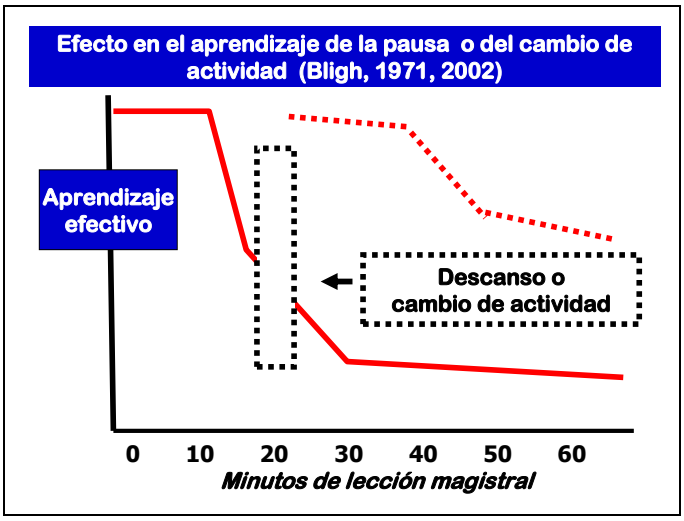

Figura 7

Una breve actividad, como es responder a alguna pregunta por escrito, devuelve a los alumnos a la clase y se prologa su capacidad de seguir atendiendo Estas estadísticas no pretenden ser exactas ni universales pero muestran lo que suele suceder. Si además esta actividad cumple con la finalidad de la evaluación formativa (verificar comprensión, corregir errores a tiempo, etc.) la clase se convierte en funcionalmente mucho más útil.

c) *Al final de la clase*

Estas breves pruebas de evaluación suelen dejarse para el *final* de la clase (es lo habitual en el caso de los *one minute papers*); es cuando parece que tienen más sentido si se trata de alguna manera de una evaluación de lo aprendido o de una reflexión sobre lo explicado en esa clase.

Sobre estas breves preguntas hechas al final de la clase hacemos dos observaciones.

1. Con frecuencia los minutos finales de la clase son minutos perdidos; los alumnos empiezan a recoger sus cosas, sólo piensan en irse y les cuesta más atender. Un breve ejercicio al final de la clase es una buena manera de aprovechar esos últimos minutos.

[17] Biggs (2006) toma esta información (y la de la siguiente figura) de Bligh (1972, 2002), autor de reconocida autoridad en el tema de la *lección magistral* (la obra citada cuenta con ediciones muy posteriores a la fecha indicada que corresponde a la citada por Biggs). Estas estadísticas lo son de manera muy aproximada (es una adaptación personal sin tener a la vista los datos originales) pero muestran con claridad *lo que suele suceder*.

2. Sobre todo cuando estas preguntas o breve actividad están centradas en un repaso sobre lo que se supone que han aprendido y entendido en esa misma clase, se garantiza mejor la retención del aprendizaje.

La figura 8 (adaptada de Biggs, 2006) muestra de manera muy gráfica que la retención de lo aprendido o escuchado en una clase cae rápidamente a los 10 días; basta un breve repaso o actividad final para que la retención se prolongue apreciablemente. A veces los profesores *apuramos el tiempo* para explicar rápidamente lo que nos queda del tema del día cuando lo más rentable para el aprendizaje del alumno es crear un espacio de repaso y reflexión final.

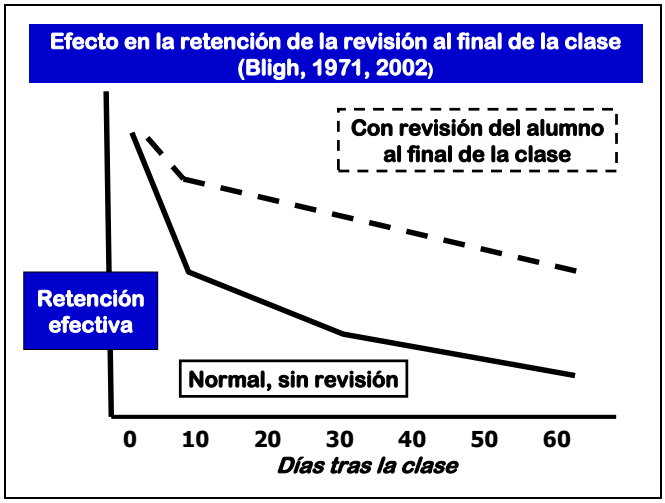

Figura 8

Las preguntas o breves ejercicios hechos al final de la clase solemos asociarlos a los populares *one minute paper*, pero caben otras posibilidades que podemos calificar como más académicas.

Un ejemplo muy ilustrativo lo tenemos en Isaksson (2008) en un curso de arqueología de dos horas de clase y diez semanas de duración (en Suecia). Al final de la clase los alumnos tienen que escribir un ensayo (que el autor denomina *ten minutes essay* aunque podría durar hasta 20 minutos) *de una extensión de una sola página,* que sí es calificado y además con un *alto nivel de exigencia* para poder obtener la calificación máxima. Al comienzo se les entrega uno de estos ensayos bien hecho como modelo. Es un buen ejemplo de a la vez evaluación *sumativa* (con nota) y *formativa* como se ve por el porcentaje progresivamente mayor de los que obtienen la máxima calificación a lo largo de las 10 semanas.

Merece la pena reproducir (de manera aproximada, figura 9) los resultados que presenta el autor.

Se observa de manara clara cómo según van pasando las semanas, son más los que obtienen la máxima calificación (en torno al 90% al terminar el período; la correlación entre semanas transcurridas y número de notas máximas es de .874). Es claro que los alumnos van aprendiendo a hacer este tipo de trabajos de manera rápida y eficaz gracias a esta evaluación frecuente (en cada clase) e integrada en el mismo proceso.

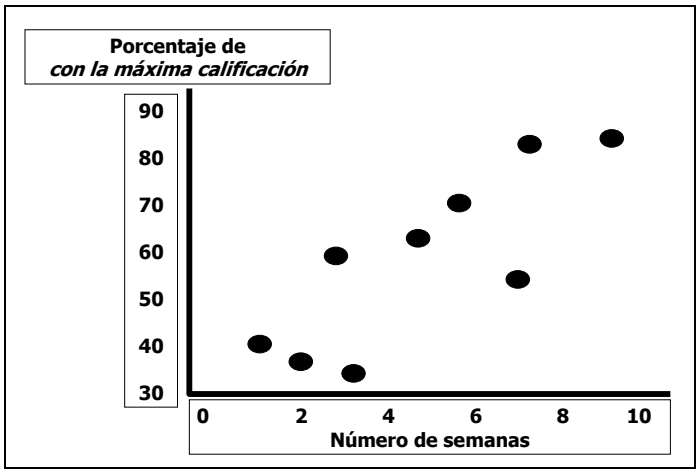

Figura 9

Una observación interesante que hace el autor es que estos ejercicios puestos al final de la clase son especialmente indicados para clases vespertinas y relativamente largas (como sucede en este caso) en las que alumnos están más cansados y se limitan a escuchar al profesor y a estudiar antes del examen.

3) Corrección y calificación

Es importante que *el miedo a la corrección* (por el tiempo que nos puede llevar) no nos detraiga de estas prácticas que pueden ser muy rentables para el aprendizaje de los alumnos, por eso conviene ver procedimientos rápidos y sencillos; tan sencillos como no corregir.

La corrección y calificación no debe ser mayor problema. Según los casos y la importancia de estas breves evaluaciones, se pueden recoger las respuestas, anotar quién las hizo, devolverlas en la clase siguiente sin corrección y comentar en la clase siguiente cuáles son las respuestas correctas, admitir preguntas aclaratorias, etc.; puede tratarse básicamente de un ejercicio de auotevaluación. Otros prefieren dar la respuesta correcta nada más responder y los alumnos se autocorrigen. Aun así es conveniente que los alumnos entreguen sus evaluaciones ya corregidas al profesor para tener constancia de quiénes las hicieron, y se devuelven en otra clase. En estos casos no tiene sentido que *cuenten para nota,* salvo el mero hecho de haber hecho estas evaluaciones.

Cabe por supuesto recoger las respuestas y corregirlas de manera convencional. En cualquier caso hay maneras de minimizar el trabajo y tiempo extra que pueden suponer para el profesor sin perder su eficacia didáctica.

Por lo que respecta a la calificación o *peso* en la nota de estas pruebas cortas y frecuentes, concebidas fundamentalmente como una *ayuda* en el aprendizaje, se ve todo tipo de prácticas; cada profesor hace lo que cree más conveniente o le va dictando la experiencia (experiencia que con frecuencia es *evaluada y publicada*)[18].

Cuando van a tener un cierto valor en la nota (pueden incluso constituir el examen distribuido a lo largo del curso y con gran aceptación por parte de los alumnos,

[18] En estas *evaluaciones de la evaluación* (o de cualquier innovación didáctica) suele tenerse en cuenta la opinión de los alumnos (como *utilidad percibida* y otras cuestiones relacionadas con el procedimiento específico utilizado), además de los análisis cuantitativos adecuados (como relación con notas y con otras variables del alumno, comparación con un grupo de control, etc.).

Leeming, 2002) no suelen tenerse en cuenta las dos o tres peor hechas o pueden dejar de hacerlas (Leeming, 2002, Connor-Greene, 2000).

Si estas breves pruebas *no van a ser calificadas* (o *sólo* se va a tener en cuenta si se hicieron o no se hicieron) y no hay por lo tanto *ningún riesgo* para el alumno:

1) Se elimina la *tensión* inevitable que acompaña a los exámenes.
2) Las *limitaciones en la corrección* por falta de tiempo (que puede ser muy rápida, o se revisa solamente una selección) no tienen mayor importancia (al menos son limitaciones que no tienen consecuencias en las calificaciones).
3) Se pueden hacer *preguntas difíciles, desafiantes,* para estimular el pensamiento y la creatividad de los alumnos; pueden hacerlo mal al principio, pero también experimentarán en ocasiones sucesivas que lo difícil se convierte en fácil.
4) Se pueden hacer *preguntas pretendidamente fáciles,* que todos o la mayoría van a responder bien (aunque a veces hay sorpresas), para crear *situaciones de éxito* (un tanto artificiales) y que los alumnos *vean que pueden,* que la asignatura no es tan difícil. Puede ser una buena estrategia para reforzar la *autoeficacia* de los alumnos (su conciencia de que son capaces).

Este reforzar *la autoeficacia y el sentimiento de la propia capacidad,* y consecuentemente mejorar la motivación, es muy importante. Todo lo relacionado con la evaluación brinda al profesor excelentes oportunidades para que sus alumnos se vayan sintiendo progresivamente más capaces. En definitiva *la autoeficacia es aprendida,* y se aprende sobre todo percibiendo el propio éxito. A este respecto es muy clara esta cita de Bandura (2002), que es sin duda el autor de referencia tratando sobre la autoeficacia.

> "Los maestros que quieren reforzar la *autoeficacia* de sus alumnos *de manera eficaz* tienen que hacer algo más que hacer comentarios positivos o dar homilías inspiracionales. Además de cultivar las creencias de sus alumnos en su propia capacidad *tienen que estructurar actividades de manera que experimenten el éxito* y evitar el colocarles en situaciones en las que previsiblemente van a tener repetidas experiencias de fracaso" (Bandura, 2002).

4) Coherencia de estas pruebas con los objetivos del curso

Dos observaciones importantes sobre estas evaluaciones formativas.

a) Estas evaluaciones, aunque sean muy sencillas, deben ser *coherentes con la evaluación sumativa* más formal que vendrá en otro momento, para verificar los resultados finales y pretendidos del aprendizaje.

Por supuesto se facilita el aprendizaje y disminuyen los fracasos, pero eso es precisamente lo que se pretende. *Hay asignaturas artificialmente difíciles por falta de información a tiempo,* información sobre errores comunes, sobre cómo corregirlos y sobre cómo estudiar.

b) El hecho de ir preparando preguntas de evaluación durante el curso nos obligará a los profesores a *concretar más los resultados esperados* del aprendizaje (habilidades, competencias que deben adquirir, capacidad crítica, cómo deben manifestar la comprensión, etc.) sin esperar a los exámenes finales.

Es difícil, por ejemplo, *enseñar a pensar* si no tenemos claro *cómo se manifiesta el pensar* (o *interpretar,* etc.) en el contexto de una determinada materia y que luego plasmaremos en preguntas y ejercicios. Esto hay que irlo calibrando y concretando al preparar nuestras evaluaciones (*tipos* de preguntas, *tipos* de ejercicios). Dicho de otra manera, además de pensar *qué* preguntamos, tenemos que pensar en el *cómo* preguntamos, de manera que la respuesta del alumno *manifieste el tipo de resultado de aprendizaje deseado* (no es lo mismo *recordar y repetir,* que *interpretar, comparar, juzgar,* etc. *¿qué verbos obligamos a conjugar a nuestros alumnos?*). El pensar en *preguntas-tipo* cuyas respuestas manifiestan el tipo de aprendizaje (comprensión, análisis, etc.) que busca el profesor, puede ser más eficaz que expresar esos mismos resultados en forma de objetivos específicos.

Stefani (2004) nos recuerda que los profesores solemos tener un programa de la asignatura muy detallado, sabemos muy bien qué queremos enseñar, y tenemos conciencia de que hacer un buen programa forma parte de nuestra idea de *profesionalidad docente*, de la misma manera que nos consideramos buenos profesores porque dominamos la asignatura que enseñamos. Tenemos que pensar que, *además* de tener ciencia y un buen programa, en esta *profesionalidad docente* entra también el saber especificar los *resultados* que esperamos en los alumnos, tal como ya podemos verlos *en nuestras preguntas y ejercicios* de evaluación formativa. Para Yorke (2001) el saber llevar a cabo la evaluación formativa es también un componente clave de la *profesionalidad docente* (*scholarship of teaching*) e incluso ya se habla de la profesionalidad específica del profesor *en cuanto evaluador del aprendizaje* en Enseñanza Superior (*scholarship of assessment,* Rust, 2007).

3.4. Efectos y consecuencias de la evaluación formativa frecuente

La evaluación integrada en el proceso de enseñanza-aprendizaje tiene unas consecuencias muy positivas. Resumimos lo que se puede esperar de estas evaluaciones formativas.

a) Estas evaluaciones dan a los alumnos la oportunidad de ir poniendo en práctica las habilidades que van adquiriendo, en una situación sin riesgos (o con un riesgo menor), e ir consolidando lo aprendido. Además si a los muy pocos días de clase, o ya en la segunda clase, se encuentran con una de estas evaluaciones, se dan cuenta rápidamente de que *aquí no jugamos a nada, esto va en serio y ya desde el primer momento.*

b) El *modo de preguntar* orienta a los alumnos sobre *cómo deben estudiar;* se facilita la *autorregulación* en el estudio: los alumnos *aprenden a estudiar* y a organizarse. Si sentimos que no es necesario que los alumnos *aprendan a estudia*r en el ámbito específico de una materia determinada, es posible que se queden en un estudio puramente memorístico (y lo que es peor, que eso sea suficiente para obtener una buena calificación).

En este sentido estas evaluaciones son especialmente importantes en el *primer curso,* cuando al comenzar sus estudios universitarios el alumno necesita una

orientación más específica; es entonces cuando tienen que situarse y entender lo que significa estudiar en la Universidad (o en el colegio y sobre todo siempre que se comienza un *nuevo ciclo*). También puede ser muy importante *al comienzo* de cualquier curso para que el profesor pueda dar a los alumnos una *orientación eficaz y a tiempo* sobre cómo tienen los alumnos que estudiar *su* asignatura.

c) Las preguntas muestran al alumno lo que realmente es *importante* en cada tema o asignatura. A veces lo importante no es el tema, sino el modo de preguntarlo (y consecuentemente de estudiarlo).

d) Una consecuencia de esta *evaluación frecuente integrada en el aprendizaje* es que los alumnos se dan cuenta del *nivel* que se espera de ellos e incluso a mayor información sobre el aprendizaje y con mayores oportunidades de mejorarlo, *se puede elevar el nivel de exigencia*[19]. La evaluación frecuente formativa supone en definitiva *dar facilidades* a los alumnos y este darles facilidades nos permite ir elevando el nivel de exigencia, más en la calidad y complejidad de las preguntas que les hacemos y tareas que les encargamos que en la cantidad de materia que les exigimos.

e) La evaluación formativa afecta positivamente a la *motivación* de los alumnos, y en general mejora el clima de la clase. El alumno percibe que el profesor *está con él*, que busca su éxito y que *el éxito es posible*, que *quiere* ayudarle a corregir sus errores cuando aún es posible hacerlo, que en lo que de él dependa, no está por suspender a nadie.

Tratando de la motivación de los alumnos no se puede dejar de recordar que la estrategia motivadora de muchos o de algunos profesores está frecuentemente asociada, quizás inconscientemente, a provocar o manipular la ansiedad, no la esperanza y la posibilidad del éxito: la incertidumbre sobre el examen final, la prueba no anunciada, el miedo al suspenso, incluso el anuncio de que son muy pocos los que pueden aprobar la asignatura, etc.

Algunos profesores parecen creer (o funcionan como si creyeran) que aumentando los niveles de ansiedad aumenta la motivación y los alumnos se esfuerzan más. Ciertamente los exámenes son muy motivadores (*¿cuándo estudian sobre todo los alumnos? antes de los exámenes*). Pero los exámenes formales y definitivos no son tan frecuentes (afortunadamente) como para que puedan afectar a la motivación de la mayoría de los alumnos para el estudio en el día a día. Realmente un buen uso de la evaluación formativa y frecuente *para* facilitar el aprendizaje (*Assessment FOR learning,* Stiggins, 2005) puede ser un factor motivador para un estudio más constante y sin provocar el *stress* asociado a los exámenes.

f) Cuando en clase hay evaluaciones formativas, los alumnos perciben que las clases son *funcionalmente necesarias.* Muchos alumnos tienden a ser muy pragmáticos en sus decisiones, y van a clase solamente cuando van a recibir una información que ellos perciben como útil (Dolnicar, 2005); muchos faltan a clase simplemente porque para aprobar el curso da lo mismo ir que no ir. La información básica necesaria ya se la transmiten los compañeros que sí van a clase. La evaluación formativa integrada en la marcha de las clases cambia radicalmente el panorama de lo que sucede en clase y de la utilidad de las mismas clases. La experiencia dice que con estos sistemas de evaluación

[19] *There is a body of firm evidence that formative assessment is an essential component of classroom work and that its development 'can raise standards of achievement'* (Black y Wiliam, 1998).

integrada en el mismo proceso de enseñanza-aprendizaje por una parte disminuye el número de fracasos y por otra es más difícil aprobar una asignatura si se falta a clase porque también ha ido subiendo el nivel de dificultad y exigencia.

g) Estas estrategias de evaluación no sólo nos informan sobre qué y cómo aprenden nuestros alumnos sino que también nos informan en qué medida nuestra enseñanza, lo que hacemos y *hacemos hacer* a nuestros alumnos es eficaz, si el *ritmo* es el adecuado, etc. Toda evaluación, formativa y sumativa, afecta tanto al evaluado como al evaluador, que puede *ir ajustando* mejor su enseñanza en beneficio de sus alumnos (Yorke, 2003). De alguna manera toda evaluación de los alumnos es (debe ser) también un ejercicio de autoevaluación del profesor y de cómo lleva la clase.

Por otra parte la eficacia de la evaluación formativa (es decir, frecuente, con énfasis en el *feedback* y corrección de errores) para aumentar y mejorar de manera apreciable el rendimiento de los alumnos tiene desde hace tiempo un gran apoyo en la literatura de carácter experimental (Black y Wiliam, 1998)[20]. Esta ayuda al aprendizaje de la evaluación formativa es muy clara con los alumnos que van peor o que por alguna razón están *en situación de riesgo,* pero sus beneficios se extienden a todos los alumnos.

4. Problemas que podemos encontrar en la evaluación formativa frecuente

Al llevar a la práctica este tipo de evaluación formativa y *frecuente* podemos encontrarnos con varios tipos de dificultades; unas tienen más que ver con nosotros mismos, otras con los alumnos y otras, que pueden ser las más molestas, pueden tener que ver con otros profesores e incluso con las autoridades académicas.

4.1. Problemas con nosotros mismos

Por parte nuestra podemos al menos ver tres dificultades:

a) El *trabajo extra* que puede suponer el *preparar* estas evaluaciones,

b) El trabajo e incomodidad de la *corrección,*

c) Quizás nos hacemos esta pregunta: *si pregunto 'todo' durante el curso ¿qué me queda para el examen de verdad?*

Comentamos estas tres dificultades potenciales.

a) Por lo que respecta al trabajo que puede suponer el preparar estas evaluaciones, ya hemos visto que caben modalidades muy sencillas, y además podemos encontrar buenos ejemplos. Tampoco hay que comenzar haciendo todo lo que sería posible o deseable hacer. Las preguntas o ejercicios de un año sirven para el siguiente y poco a poco vamos acumulando material y experiencia. Suele decirse y con mucha razón, que *diez años de experiencia no es lo mismo que un año de experiencia repetido diez veces*.

[20] Los autores (Black y Wiliam, 1998, revisan 250 investigaciones publicadas (seleccionadas por su calidad de las 580 fuentes encontradas). Cuando se compara en los mismos tests el rendimiento de los alumnos que han tenido de manera sistemática evaluación formativa con los que no la han tenido, la diferencia en términos de la *magnitud del efecto* [un procedimiento específico de cuantifica diferencias entre grupos] está entre .40 y .70 [mayor de lo que es habitual encontrar en estos estudios]; esto quiere decir que el *alumno medio* que ha tenido evaluación formativa, y que en su propio grupo supera al 50% de sus compañeros, está por encima de entre el 65% y 76% de los alumnos equivalentes que no han tenido evaluación formativa. En investigación educacional un *tamaño del efecto* de .30 o poco más se considera ya relevante. En esta *investigación de investigaciones* se incluyen grupos de alumnos muy diversos y ha sido objeto de críticas metodológicas muy matizadas que complican el generalizar los resultados a cualquier otra situación (Dunn y. Mulvenon, 2009), pero en términos generales apoyan de manera clara la utilidad de la evaluación formativa.

La experiencia y los años pueden ser muy rentables si sabemos aprovechar el trabajo pasado y vamos mejorando y ampliando lo ya hecho en cursos anteriores. Más o menos es también lo que solemos hacer con los exámenes convencionales. Además como estas *evaluaciones más informales* no forman parte de nuestras obligaciones formales, podemos sentirnos más libres para experimentar y seguir con nuestro propio ritmo.

b) La corrección es un problema menor sobre todo si no hay calificaciones. Por lo general estas breves evaluaciones se corrigen con mucha facilidad y rapidez, casi de un *golpe de vista.* Ya hemos indicado que se pueden corregir en la misma clase o en la clase siguiente, comentando algunas de las respuestas. Se pueden hacer cosas distintas en días distintos. Algunos profesores hacen algún análisis sencillo, como una distribución de frecuencias, una media, cuántos han respondido correctamente a cada pregunta, etc., de manera que haya una referencia que permita situarse a cada alumno en el contexto de su propia clase.

c) Si vamos preguntando todo durante el curso ¿Qué preguntamos en el examen? La respuesta es muy sencilla: lo mismo, pero *de otra manera,* con otro caso, con otro ejemplo, etc. Además tampoco se pregunta *todo* durante el curso y lo importante suele ser *cómo* se pregunta más que el *qué* se pregunta.

Sobre preguntar *lo mismo* que se ha ido preguntando durante el curso en un examen final caben varias posibilidades cuya conveniencia debe juzgar el profesor. El que *algunas* preguntas sean las mismas (por ejemplo en una prueba objetiva) es una opción, aunque en muchas preguntas que no son de memoria lo que sí es importante es mantener el mismo tipo de *proceso mental* para responder correctamente, que es lo que realmente se desea evaluar. Si *sentimos* que si preguntamos mucho durante el curso no *nos queda nada* para el examen… posiblemente hacemos demasiadas preguntas de memoria.

Un estudio experimental sobre preguntar en el examen *lo mismo* (aunque sólo en parte) es el de Landrum (2007) que tiene su interés porque caben réplicas parecidas. Los alumnos (N = 253) responden a una prueba objetiva *semanal* (15 semanas, cada semana un tema distinto) de 20 preguntas (con tres respuestas); los alumnos reciben a la semana siguiente sus resultados y las respuestas correctas de cada pregunta. Si faltan a clase alguna semana, pueden responder a la prueba (modificada) en otro momento.

En el examen final (de toda la materia, 150 preguntas) la mitad de las preguntas son las mismas ya preguntadas (y los alumnos lo saben desde el principio), y la mitad de esta mitad tienen alterado el orden de las alternativas de respuesta. La finalidad del autor es doble, reducir la ansiedad en el examen y mejorar el aprendizaje. Los alumnos con mejores promedios semanales hicieron mejores exámenes finales (como era de esperar), pero los que comenzaron peor (y tenían más que ganar) fueron quienes más mejoraron su aprendizaje.

4.2. Problemas con los alumnos

Por parte de los alumnos un problema que podemos encontrar, sobre todo en algunas culturas académicas, es su resistencia para hacer algo con lo que no van a ganar ningún punto, *que no vale para nota*. No va a ser siempre éste el caso, pues puede haber evaluaciones frecuentes que sí van a ser calificadas o de alguna manera *premiadas*.

Es muy importante hacer conscientes a los alumnos de que la evaluación formativa es *una parte integral de su proceso de aprendizaje,* que tiene más de

actividad didáctica que de examen, y que no deben identificarla con la evaluación o exámenes convencionales que desembocan en una calificación. Vienen a clase a aprender, no a ganar puntos.

A veces, con más razón si no hay *nota*, puede bastar un cambio de nombre y hablar de prácticas o ejercicios más que de evaluación propiamente dicha, que es un término que tiene unas connotaciones difíciles de evitar.

Siempre cabe preguntar a los alumnos *¿Cuándo quieren Vds. saber lo que no saben y deberían saber para aprobar, ahora o el día del examen?* La respuesta previsible es *ahora,* y será el momento de empezar a cambiar e inaugurar otras tradiciones. La experiencia da que los alumnos terminan por esperar y agradecer estas evaluaciones formativas, que eliminan la ambigüedad, les dan seguridad, les enseñan a estudiar, manifiestan nuestra buena voluntad como profesores que buscamos en buen resultado *en todos* y hacen que disminuyan los fracasos finales.

Si el profesor cree que estas pruebas cortas, orientadas a corregir errores, enseñar a estudiar, etc., sí deben ser calificadas (las situaciones pueden ser muy distintas y no estamos ante una *ciencia exacta*) lo que sí parece conveniente es que en cualquier caso y como ya se ha repetido varias veces, tengan un *peso menor* en la calificación final, e incluso se puede tener en cuenta solamente el haberlas hecho independientemente de que estén bien o mal; también se pueden tener en cuenta para subir la nota.

4.3. Problemas con otros profesores

La evaluación formativa supone concebir la evaluación integrada en el aprendizaje para corregir errores a tiempo y a la vez poder aumentar el nivel de exigencia y evitar, en la medida en que nos es posible a los profesores, el fracaso de los alumnos. Ya hemos visto que esto se puede hacer de muchas maneras (*se puede hacer y se hace* como queda claro en las múltiples y variadas citas aducidas). El ir por este camino supone una concepción muy clara de nuestro rol como profesores (*ayudar al alumno a aprender*) que puede chocar con el ambiente en el que nos toca movernos y las *creencias* prevalentes en muchos de nuestros colegas y, lo que puede ser más incómodo, en nuestras autoridades académicas. Todo esto nos puede llevar a problemas de convivencia académica a los que conviene dar un somero repaso.

En algunas culturas académicas un determinado porcentaje de fracasos (suspensos) es un indicador importante del nivel de exigencia. Este indicador se convierte en una *norma implícita* que puede ser muy condicionante, porque el profesor que no suspende al porcentaje *normal,* o que aprueba (o da una buena nota) a alumnos que los demás profesores suspenden o califican claramente peor, puede ser *mirado con sospecha* porque puede haber varias razones *para sospechar*. Lo que se puede sospechar es que unos mejores resultados (en notas) no se deben a que los alumnos saben más y han aprendido mejor, sino a que el profesor (por ejemplo)…

a) Baja el nivel de exigencia (la *eficacia docente* es juzgada como *benevolencia*, sin más),

b) Se quita de encima problemas con los alumnos,

c) Elimina o minimiza el trabajo extra que suponen los exámenes en segundas convocatorias,

d) Consigue unas evaluaciones más benévolas de los alumnos, etc.

Que unas buenas notas (en términos comparativos a lo que suele suceder en nuestro entorno más inmediato) implican un menor nivel de exigencia y un deseo de captar la benevolencia de los alumnos es algo que hemos oído con frecuencia. O que confundimos la enseñanza universitaria con el preescolar. Podemos oír comentarios que nos pueden parecer incluso ofensivos.

Estas explicaciones de un menor número de fracasos pueden ser correctas pero también pueden ser injustamente incorrectas y no ser otra cosa que una *valoración sesgada* de una realidad concreta: puede haber menos suspensos y mejores notas en general porque los alumnos sencillamente han aprendido más y mejor, y esto es mérito del profesor correspondiente. Su trabajo le ha costado.

Un profesor con una *política didáctica de evaluación frecuente,* orientada a corregir errores a tiempo, a enseñar a estudiar, *a dar segundas oportunidades,* etc., y a hacer ver a los alumnos que son capaces y que el éxito es posible:

a) Puede conseguir que los alumnos *estudien mejor* (*aprenden a estudiar* la asignatura),

b) Puede conseguir que los alumnos estudien no solamente *mejor*, sino de manera *más constante,* sin esperar a la víspera de los exámenes,

c) Puede estimular la *autoconfianza y motivación de todos,* pero quizás sobre todo de aquellos alumnos que con mayor probabilidad podrían fracasar (por diversas razones; porque tienen una menor capacidad aunque tengan una capacidad *suficiente,* porque tienen malos hábitos de estudio, porque en otras circunstancias, sin evaluaciones frecuentes, apenas hubieran ido a clase, etc.).

d) Puede *subir (y no bajar) el nivel de exigencia* porque a mayor información sobre el propio aprendizaje es muy factible ir exigiendo más sin que suba el nivel de fracasos,

e) En definitiva con una práctica de evaluación frecuente, aunque sea muy senilla, puede conseguir que disminuyan los fracasos y que aumente tanto el nivel de exigencia como el nivel y calidad de aprendizaje de sus alumnos.

Todo esto (buenos efectos en el aprendizaje) no son meras hipótesis; son consecuencias constatadas de muchas maneras en múltiples estudios empíricos y publicados en buenas revistas[21].

También puede suceder en ocasiones que este muy probable mejor aprendizaje no sea tan visible si al calificar se mantiene más o menos un porcentaje *normal* (o a lo sumo ligeramente inferior a lo considerado normal) de suspensos (para no quebrantar las *normas implícitas*); pueden incluso quedar suspendidos alumnos que quizás *con esa misma ciencia* (potenciada por el sistema de evaluación) y sin esa presión ambiental o normas implícitas hubieran aprobado en otras situaciones.

Si parece claro que la evaluación frecuente puede tener otras consecuencias no previstas que habrá que examinar y resolver:

a) Conflictos con otros profesores pues los alumnos *comparan y comentan,* ven y viven las diferencias entre unos profesores y otros, o asisten más a unas clases

[21] Las revistas citadas son una mínima selección de las que pueden encontrarse sobre didáctica y temas afines en Educación Superior.

que a otras, etc. (ya sabemos que van más a clase si esperan algún tipo de evaluación).

b) Conflictos reales o aparentes con las *normas* sobre evaluaciones y exámenes (que es verdad que en ocasiones pueden ir en contra de la eficacia didáctica y de la calidad del aprendizaje).

c) Conflictos con la autoridad competente que recibe diversos tipos de quejas, que no es partidaria de innovaciones didácticas cuando *las cosas van bien* (al menos es así como se percibe y que en cualquier caso no se plantea que podrían ir *mejor*), que recibe otro tipo de informaciones y presiones y no quiere problemas, etc.

Este tipo de consecuencias y conflictos puede llevar:

a) A inventar, imaginar o sobrevalorar las *razones para sospechar* enunciadas antes y que podrían explicar unas mejores calificaciones.

b) A minimizar la calidad y eficacia de muchas de estas pruebas de evaluación formativa, que pueden ser juzgadas, *a priori* y sin examinarlas en su contexto, como muy infantiles y también desde el desconocimiento de su uso en excelentes universidades (*the Minute Paper is now used in more than 400 classes at Harvard***,** Cross, 1998)

c) A minimizar también la utilidad real de la evaluación continua o frecuente, que puede llevarse a cabo de maneras muy distintas y con niveles muy distintos de formalidad; ya hemos citado numerosos ejemplos.

Por lo que respecta a los efectos claramente positivos en la motivación y calidad del aprendizaje de los alumnos de la evaluación frecuente, de las *segundas oportunidades* antes de la calificación y de todas las prácticas relacionadas con la evaluación formativa, lo que está claro es que tienen una *confirmación empírica abrumadora* y fácilmente verificable en todos los niveles educativos incluido, por supuesto, el universitario.

5. La evaluación formativa integrada en la evaluación y exámenes convencionales

Me he referido expresamente a la evaluación formativa, que en principio puede incluso no ser calificada, y sobre todo *muy integrada* en el proceso de aprendizaje. Pensando ahora tanto en la evaluación específicamente formativa (para informar y corregir errores *a tiempo*) como en la evaluación más convencional (sumativa y final, aunque se trate de exámenes parciales), la que hacemos todos para calificar a nuestros alumnos, también es importante poner de relieve la importancia de una buena *información de retorno* de cara a corregir errores y encauzar el futuro aprendizaje. La eficacia de la evaluación formativa, y de toda evaluación, depende en buena medida de la información que reciba después el alumno.

Es importante pensar en primer lugar en la importancia del *feedback* y en sus funciones y sugerir después diversos procedimientos para dar esta información a los alumnos después de los exámenes y trabajos convencionales (es más sencillo cuando se trata de evaluaciones más breves de carácter sobre todo formativo).

5.1. Importancia del *feedback* como estrategia de enseñanza-aprendizaje

En la figura 5 se incluyen también los exámenes formales (exámenes parciales liberatorios, finales, trabajos y tareas que tienen su nota) como una modalidad de evaluación formativa, pero no lo es necesariamente. Sobre todo los exámenes parciales, cuando todavía queda curso por delante, sí deberían tener también de manera clara la finalidad de la evaluación formativa. No se trata simplemente de comunicar resultados (una nota, número de respuestas correctas, etc.). Para que cualquier tipo de evaluación sea eficaz para corregir los propios errores (o ignorancias, o malos hábitos de estudio) es muy importante que el alumno reciba algún tipo de *feedback más específico*. Es una queja de muchos alumnos: reciben de vuelta un examen o un ejercicio corregido con marcas rojas, o comentarios breves y genéricos pero siguen sin enterarse de por qué está mal su respuesta, en qué se han equivocado, cómo pueden mejorar. No se trata de que el alumno se entere de lo que tiene mal; además debe enterarse de por qué está mal o cómo puede mejorar.

La eficacia del *feedback* en el aprendizaje está bien confirmada en muchos estudios de tipo experimental; menciono un solo ejemplo sencillo pero muy indicativo. Un cuestionario muy utilizado para evaluar la calidad de la enseñanza es el *Course Experience Questionnaire* (Ramsden, 1992); el ítem que distingue con más claridad los mejores y peores cursos es precisamente el referido al *feedback* proporcionado por los profesores (*teaching staff here normally give helpful feedback on how you are doing*; (Gibbs y Simpson, 2004; estos autores mencionan varios meta-análisis, o síntesis de muchos estudios, que no dejan dudas sobre la eficacia del *feedback* en el aprendizaje).

Sobre la importancia del *feedback* en la tarea de aprender, podemos recordar un ejemplo no fácilmente imitable pero sí clarificador. Si nos piden que demos dos ejemplos de buenas universidades conocidas internacionalmente, seguramente nos vendrán a la mente Oxford y Cambridge. Son dos universidades que siguen estando entre las primeras en los *rankings* que se publican ocasionalmente o están disponibles en Internet[22].

Tradicionalmente los alumnos de estas dos universidades han tenido que escribir un ensayo o hacer un trabajo semanal que entregan a su tutor del que reciben un *feedback* detallado, oral e inmediato. En el caso de las ciencias estas tareas pueden ser problemas o informes sobre el trabajo de laboratorio. El profesor, más que dar una clase semanal al modo tradicional, es un tutor que comenta y discute con el alumno su propio trabajo. La *enseñanza,* entendida al modo tradicional, casi se reduce a estos encuentros semanales, puramente de evaluación formativa y sin calificaciones; la evaluación sumativa con una calificación dependía frecuentemente de un solo examen final al cabo de tres años (Gibbs y Simpson, 2004). Esta práctica no es fácilmente (ni difícilmente) imitable, pero pone de relieve la importancia del *feedback,* a poder ser inmediato, en la calidad del aprendizaje. Realmente en esas universidades el comentar con los alumnos sus propios trabajos ha sido *el método de enseñanza* fundamental.

Si nos ponemos como ejemplo a las Universidades de Oxford y Cambridge, posiblemente lo único que conseguiremos es desanimarnos, pero la idea importante que se desprende del éxito y prestigio de estas universidades es que tareas frecuentes con un *feedback* detallado y personalizado por parte del profesor son clave para un aprendizaje

[22] Por ejemplo el del Institute of Higher Education, Shanghai Jiao Tong University, 2005 (dirección electrónica en las referencias bibliográficas).

de calidad. Si buscamos (y *pregonamos* en nuestros documentos) que lo que más nos importa es la *excelencia* tendremos que tomarnos en serio todo lo relacionado con esta *información de retorno* dada *a tiempo* a los alumnos.

Cualquier tipo de evaluación *sumativa* debe tener una doble finalidad; una es la que se refleja en una nota, y otra contribuir al aprendizaje del alumno. Esta función no la da, o la da muy pobremente, una simple nota o un juicio muy general (del tipo *bien, regular,* etc.) o comentarios muy breves. Hace falta una *información de retorno* mucho más específica, de manera que el alumno se dé cuenta de en qué está bien y en qué está mal y cómo puede mejorar. Esto es fácil decirlo pero si pensamos en el tiempo extra que puede requerir una información de retorno más personalizada, la tentación es seguir como siempre y no plantearnos ninguna novedad en este terreno. Ciertamente no se pueden minimizar los problemas derivados de nuestro escaso tiempo, sobre todo si tenemos clases numerosas y otras ocupaciones, por eso en un apartado posterior apuntamos algunas sugerencias, muy distintas entre sí, que pueden facilitar nuestra tarea.

En buena medida lo que hagamos dependerá de cómo nos vemos como profesores (*lo mío es explicar la asignatura y después examinar,* o más bien *mi tarea es sacar de aquí profesionales competentes*…).

5.2. Funciones del *feedback*

Tanto en la evaluación formativa como en la sumativa, las funciones que cumple una buena *información de retorno al alumno* son extremadamente útiles para la formación de los alumnos (puede verse una buena sistematización en Nicol y Macfarlane, 2004, 2006). Es importante que pensemos en todo lo que puede proporcionar esta información porque es lo que hace realmente útil cualquier evaluación si pretendemos algo más que *poner notas*.

1) Facilita la *autoevaluación* del alumno, la reflexión sobre su propio aprendizaje. El alumno suele estar pendiente de la evaluación del profesor pero también debe desarrollar la actitud y la habilidad necesarias para autoevaluarse. Este punto (la *autoevaluación*) es muy importante y hay que hacérselo ver así a los alumnos. La excesiva dependencia de la evaluación y comentarios del profesor puede inhibir el desarrollo de la capacidad del alumno para autoevaluarse correctamente. Vive pendiente del juicio del profesor pero a la larga lo que le hará progresar, y no sólo en los estudios, es su capacidad de autocrítica.

2) Facilita la *comunicación entre profesores y alumnos y entre los propios alumnos,* que pueden aclararse muchas cosas entre sí. Con frecuencia los alumnos se explican entre sí puntos difíciles con más claridad que el profesor, en su propio lenguaje. Es más, se puede dar a los alumnos la oportunidad de comentar en pequeños grupos y en clase sus propios ejercicios; el profesor queda como una referencia y autoridad externa (ya hemos mencionado que la *corrección colaborativa,* y en general el aprendizaje colaborativo, es un tema de mucha actualidad en didáctica universitaria sobre el que hay numerosas investigaciones).

3) Clarifica los *criterios de evaluación,* queda más claro qué se espera de los alumnos, cuáles son exactamente los objetivos de aprendizaje, cuál es el nivel de exigencia. Estas cosas ya suelen decirse al comenzar el curso y durante las clases pero

es claro que muchos alumnos no se enteran; la información *eficaz* les viene cuando repasan sus propias respuestas y los comentarios del profesor.

4) Facilita el *cambio* en el alumno en su modo de estudiar y de hacer sus trabajos. En este sentido el recibir un *feedback* adecuado es especialmente importante en el *primer año* de Universidad, cuando el alumno necesita una mayor orientación; por otra parte los efectos de un *feedback negativo* en los alumnos que están comenzando sus estudios pueden ser *devastador*es[23].

5) El *feedback* debe llevar al alumno a mejorar, pero este *debe mejorar* no se cumple de manera automática por el mero hecho de recibir unos comentarios sobre su trabajo. La *eficacia positiva* del *feedback* depende de dos factores:

a) El *feedback* es eficaz en la medida en que esta información es clara y específica. Los juicios muy generales ayudan poco.

b) Quizás sobre todo el *feedback* es eficaz también en la medida en que puede *repetir* su trabajo o hacer otro semejante. Este *poder repeti*r es más aplicable a los trabajos hechos en casa

Si el alumno entrega un trabajo, recibe su nota y pasa a otra tarea, se pierde una buena oportunidad (o la *única* oportunidad) de mejorar. Naturalmente esta *repetición* no es siempre fácilmente posible, pero no sobra considerar su posibilidad (la *doble corrección* de trabajos, que aunque parezca extraño no es una práctica *tan inusual,* la veremos en el apartado siguiente).

Cuando el alumno recibe una *nota* y además unos comentarios, es muy normal que se fije en la nota, que es lo que realmente le interesa, y se olvide de los comentarios, a no ser que perciba que le pueden ser útiles en una segunda oportunidad. Si no hay *todavía* una nota definitiva, los comentarios ganan en eficacia. La nota, sin más, lleva al alumno a compararse con otros (y también en muchos casos a criticar al profesor) más que a reflexionar sobre cómo ha trabajado y cómo puede mejorar. Se puede intentar hacer mejor las cosas cuando hay una *segunda oportunidad*.

6) El *feedback* puede ser un *factor motivado*r que puede contribuir a aumentar la *autoeficacia* del alumno, la percepción que tiene de su propia capacidad y de sus posibilidades de éxito. He dicho *puede* porque también puede suceder todo lo contrario. Ciertamente una mera nota, sobre todo si es una mala calificación, no contribuye a la autoestima y motivación del alumno. Aquí entran en juego la habilidad del profesor y también sus propias actitudes. El profesor puede destacar *éxitos parciales,* hacer ver que los propios errores son una excelente oportunidad para mejorar y aprender, dejar claro qué y cómo se debe mejorar. *Hay que saber convertir el error en una oportunidad.*

La motivación y la autoeficacia se potencian posiblemente con más facilidad en evaluaciones que son solamente formativas, sin ningún tipo de nota, pero también en cualquier otra situación en la que se devuelven y comentan exámenes y trabajos.

7) Con un *feedback* adecuado y oportuno se solucionan problemas y limitaciones de los alumnos que pueden tener muy fácil solución.

[23] Poulus y Mahony (2008), en una investigación sobre la eficacia del *feedback* desde la perspectiva de los alumnos, es un estudio hecho con cuatro grupos focales de la Facultad de Ciencias de la Salud de la Universidad de Sydney.

Por lo que respecta a los trabajos escritos (*essays*) las dificultades que encuentran los alumnos suelen provenir de tres tipos de limitaciones que pueden solucionarse (en parte al menos) si damos con un eficaz procedimiento de *feedback* (Hartley y Chesworh, 2000).

a) *Déficits personales* (errores de ortografía, puntuación, etc.),

b) *Desconocimiento de las normas,* saber qué se espera de ellos, cómo presentar el trabajo, cómo citar, si están o no permitidas notas a pie de página, etc. (incluso distintos profesores y departamentos pueden tener normas distintas),

c) *Limitaciones institucionales,* como no devolver los trabajos en un tiempo razonable y con un *feedback* suficientemente específico (estas limitaciones suelen tener que ver con el número de alumnos), no tener disponibles las lecturas recomendadas, etc.

Estos autores (Hartley y Chesworh, 2000) presentan un cuestionario (analizado en una muestra de 102 alumnos de psicología) de 30 preguntas (con respuestas *sí, a veces, no*) para evaluar este tipo de problemas que los alumnos pueden encontrar en los trabajos escritos; podría ser útil hacer un estudio semejante en nuestro medio.

5.3. Cómo dar información específica a los alumnos sobre exámenes y trabajos

Damos por hecho que los *comentarios personales* son importantes y posiblemente es el procedimiento más habitual, pero no siempre es posible dar a cada alumno una información muy específica y útil, sobre todo en clases numerosas. Si los comentarios se limitan a una nota o a señalar errores, su utilidad para mejorar futuros trabajos puede ser muy escasa o simplemente nula; la atención del alumno se centra en la nota y se olvida de todo lo demás. En cualquier caso los comentarios personales del profesor son compatibles con otros modos de dar información a toda la clase.

Apuntamos algunos procedimientos que pueden facilitar esta tarea; en muchos casos se trata de dar la misma información a toda la clase de manera que sirva de referencia para que cada alumno examine su propio trabajo y vea cómo puede mejorar en otra ocasión. Algunas de estas sugerencias pueden ser especialmente útiles para dar información más detallada sobre exámenes convencionales (la 2, y la 4) y otras pueden ser más útiles cuando se trata de trabajos escritos (la 2, 3 y 5).

1) Preparar claves de corrección (*rúbricas*)

Estas claves son especialmente útiles para evaluar todo tipo de trabajos, por lo general hechos en casa, individualmente o en grupo. Estas *claves de corrección*[24] son a la vez muy informativas. Se preparan de antemano especificando de manera detallada las características o aspectos sobre los que se quiere informar al alumno y que son también los criterios de calificación. Si los alumnos conocen estas claves de antemano, son a la vez las *normas* para que hagan un buen trabajo (normas que en cualquier caso deberían estar claras).

Estas *claves* facilitan la lectura y corrección y son muy adecuadas para todo tipo de trabajos, tareas en grupo, prácticas, etc. Se devuelven cumplimentadas a cada alumno que recibe así una información muy específica sobre sus puntos fuertes y débiles.

[24] *Rubrics* en inglés y de ahí posiblemente se ha popularizado la palabra *rúbrica* con el significado de *clave de corrección.*

Podríamos decir que, con estos sistemas, nuestras correcciones y recomendaciones están escritas de antemano y a cada alumno se le señala lo que le corresponda[25].

2) Preparar un *informe* para toda la clase

Otra manera de dar información útil para los alumnos sin limitarnos a una nota o a unos comentarios breves y genéricos en el margen, consiste en preparar *un buen informe para toda la clase* sobre los resultados de un examen o de los trabajos que han entregado. No podemos dirigirnos a cada alumno en particular, pero sí es fácil hacerlo a toda la clase.

En ese informe caben todos los comentarios y orientaciones pertinentes, y se puede hacer llegar a los alumnos de diversas maneras, a) se puede fotocopiar y entregar a todos, b) se puede poner en una página Web y ya es accesible para todos los alumnos, o c) se puede mandar a todos por correo electrónico.

Otra manera de dar una buena información a toda la clase es preparar una presentación en Power Point.

Nada más corregir una serie de exámenes o trabajos de los alumnos, cualquier profesor tiene muchas cosas que decir de manera casi improvisada:

a) Fallos generalizados y cómo corregirlos; puntos del texto o apuntes que conviene repasar, etc.

b) Qué distingue a los mejores exámenes o trabajos de los no tan buenos,

c) Incluso cabe una re-explicación de algún punto o remitirse a determinados lugares de un texto para clarificar algunas ideas.

Para muchos ese informe llegará tarde por lo que respecta a ese examen o trabajo concreto, pero puede quedar una buena orientación para el futuro, y quedará patente al menos el interés del profesor por el buen aprendizaje de sus alumnos.

3) Preparar *respuestas modelo*

Otro modo de informar a los alumnos es preparar y entregarles una *respuesta modelo* con la que pueden comparar sus propias respuestas o trabajos.

Huxham (2007) presenta una investigación en la que da a los alumnos dos tipos de *feedback*, comentarios personales y una *respuesta modelo;* los alumnos prefieren (en este caso) los comentarios personales pero son las *respuestas modelo* las que más contribuyeron a hacer mejor los exámenes y tareas subsiguientes (la muestra es de 155 alumnos de dos cursos de biología; sus preferencias y opiniones sobre los dos tipos de *feedback* se recogieron con un cuestionario).

Una *respuesta modelo* o *trabajo modelo* también se puede entregar *antes* de que los alumnos hagan un determinado trabajo o una serie de trabajos, para que les sirva de

[25] Ejemplos de rúbricas muy claros para trabajos escritos pueden verse en Smyth (2004) y Kuisma (1999). También para trabajos escritos Rawson y otros (2005) (casos clínicos en veterinaria) presentan un listado de 7 características *(precisión, organización, justificación de lo que se dice,* etc., evaluadas de 1 a 3) y Kogan y Shea (2005) (ensayos breves en medicina) presentan otro listado de 14 características (valorados de 1 a 4.). Sobre *rúbricas* y en el contexto de la *evaluación de competencias*, puede verse Blanco (2008), que incluye bibliografía adicional; las rúbricas se utilizan sobre todo para evaluar *competencias*. Otras referencias de interés sobre rúbricas son Andrade (2000), Moskal (2000), Mertler (2001). En Internet se encuentran con facilidad ejemplos de rúbricas de todo tipo (búsqueda con palabras clave *rubrics* y *authentic assessment*).

guía y orientación. Esta práctica (entregar respuestas o trabajos modelo) no es tan inusual. Ya hemos mencionado a Isaksson (2008) a propósito del ensayo que tenían que escribir los alumnos al final de la clase. Otro ejemplo lo tenemos en Kogan y Shea (2005); a los alumnos se les entrega una *respuesta modelo* antes de hacer seis análisis de casos en medicina a lo largo de un semestre. *En definitiva se trata de poner los medios para que los alumnos trabajen bien desde el principio.*

4) Revisar en clase los exámenes corregidos

Otra buena estrategia para *integrar la evaluación sumativa y formativa* puede ser:

1º Devolver a los alumnos los exámenes ya corregidos junto con el *feedback* del profesor (tanto si se da cada alumno en particular en su propio examen como si se da a todos conjuntamente con algún tipo de informe).

2º En otra clase dar un tiempo para repasar el examen y clarificar lo que sea necesario. Si se va a repasar el examen en clase (*pregunta por pregunta*) ya no es tan importante el haber hecho previamente comentarios personales en cada examen, con un ahorro de tiempo que puede ser considerable.

Esta práctica es útil con cualquier tipo de examen o ejercicio (excepto en los exámenes finales, cuando ya no quedan más clases).

Un ejemplo sencillo de la eficacia de revisar los exámenes en clase es el que presenta Winniger (2005). En un grupo (*grupo experimental*) se hizo en clase la revisión del examen y en otro (*grupo de control*) no se hizo la revisión del examen (se trata del mismo examen de 50 preguntas objetivas). Una semana más tarde se pone el mismo examen a los dos grupos; es un *examen sorpresa*, sin previo aviso, pero sólo válido para subir nota, por lo que los alumnos aceptan sin problema la *sorpresa*. En el primer examen los resultados de los dos grupos son prácticamente idénticos; en la repetición del examen el grupo experimental (que tuvo revisión del examen) aventaja al grupo de control de manera muy apreciable.

Esta actividad (de revisión y aclaración) puede en algún momento ser evaluada por los mismos alumnos; en este caso se hizo con un sencillo cuestionario con preguntas de este estilo *me gustaría que se hiciera lo mismo* [comentar los exámenes en clase, pregunta por pregunta] *en otras clases, es una manera útil de emplear el tiempo de clase, me ayuda a entender lo que no tenía claro, es una buena oportunidad para hacer preguntas*.

Esta evaluación de los alumnos (sus juicios e impresiones) siempre es de interés; el mero hecho de responder ayuda a los alumnos a tomar conciencia del interés del profesor en su aprendizaje y a fijarse más en las correcciones y recomendaciones que acompañan a los exámenes.

5) Corregir los trabajos *dos veces* (borrador y versión definitiva)

Por lo que respecta a trabajos hechos en casa, lo que hacen algunos profesores es una doble corrección; la primera vez corrigen un borrador para informar sobre fallos y sugerir cómo hacerlo mejor, y la segunda y definitiva para calificar. Los alumnos entregan en primer lugar, y en la fecha prefijada, un borrador o *versión provisional* de su trabajo en el que debe aparecer un esquema general y completo de lo que están haciendo y cómo lo van desarrollando. En algunos casos, si la tarea está ya bien hecha, esta primera versión puede ser ya la definitiva.

Un modo de hacerlo es clasificar los trabajos, *tal como* se entregan en la primera versión provisional, en cuatro categorías que corresponden a otras tantas calificaciones aproximadas y previsibles como pueden ser A: de 8 a 10, B: 6 ó 7, C: 5 ó 4 y D: suspenso claro. El profesor debe pensar antes en los criterios para asignar a cada trabajo a una categoría u otra; con unos criterios claros esta clasificación puede hacerse muy rápidamente. Los alumnos reciben esta calificación aproximada y provisional, y en la clase siguiente se explican (o se vuelven a explicar) los criterios de calificación (que naturalmente corresponden a las normas dadas previamente) de manera que puedan mejorar sus trabajos y se califica después la versión definitiva del trabajo[26].

La información *más precisa* sobre fallos, comentarios, modos de mejorar, etc., *correspondientes a cada categoría de la clave o rúbrica de corrección* se puede dar, con un importante ahorro de tiempo para el profesor, a toda la clase (informe oral, fotocopia, en una página Web, una presentación en Power Point). Los alumnos perfilan después su trabajo y es calificado cuando lo entregan por segunda vez.

Lo que se puede esperar con este sistema de doble corrección es que los alumnos aprendan a trabajar, a redactar mejor, que corrijan sus fallos y mejore su calificación; *la posibilidad de mejorar la nota es siempre una buena motivación*.

El entregar una versión previa puede ser obligatorio o puede ser libre; cuando es libre una razón aducida por los alumnos para no entregar una primera versión suele ser la falta de tiempo; lo que sí está comprobado experimentalmente es que aunque la motivación de los estudiantes sea simplemente mejorar la nota, los que de hecho someten una primera versión de sus trabajos a la corrección y *feedback* del profesor, también, y como era de esperar, mejoran sus habilidades de redacción y organización general del trabajo.

Ejemplos específicos sobre la eficacia de esta doble corrección en Covic y Jones (2008 y Ellery (2008) En el caso de Covic y Jones (2008) con 54 alumnos de Psicología; el trabajo es una redacción o ensayo; los alumnos son calificados ya en la primera versión y además reciben un *feedback* detallado junto con la clave de evaluación (según siete aspectos; introducción, organización, estilo, etc.); la segunda versión deben entregarla dos semanas después y obtienen la calificación definitiva.

En la experiencia de Ellery (2008) la autora después de corregir solamente el 10% de los trabajos y comentar con los alumnos deficiencias importantes encontradas en estos trabajos, da a los alumnos la oportunidad (para los que quieran) de rehacer sus trabajos; los que de hecho optan por una repetición la calificación será la del trabajo que los mismos alumnos prefieran; el resultado es una disminución clara de notas bajas y un aumento muy notable de notas altas.

Este corregir *dos veces* el mismo trabajo puede ser más eficaz que encargar dos trabajos. Algunas investigaciones muestran que (como cabría esperar) el *feedback* es realmente eficaz si de alguna manera hay un *paso ulterior* en el que los alumnos *tienen que* rehacer el trabajo teniendo en cuenta la información dada por el profesor en una primera versión (Rust, 2002). Si solamente hay una primera versión que es la definitiva y es calificada, no es nada seguro que la información que pueda dar el profesor sobre el trabajo se tenga en cuanta en trabajos futuros.

[26] Ignacio Cervera (Universidad Pontificia Comillas), comunicación personal.

La primera corrección del borrador se presta a establecer un diálogo con la clase sobre los criterios de evaluación (*'assessment dialogues',* Carless, 2006); este diálogo es más eficaz cuando todavía no se ha entregado la versión definitiva.

El estudio de Carless (2006) merece la pena comentarlo con cierto detalle. Se trata de una investigación sobre cómo entienden el *feedback* los alumnos y profesores (investigación hecha con una muestra de 460 profesores y 1740 alumnos de ocho universidades; cuestionario y entrevistas grupales). La investigación se refiere a la corrección de *una primera versión* o borrador de un trabajo escrito. Este estudio pone de manifiesto la distinta percepción que profesores y alumnos pueden tener de la eficacia del *feedback,* y la mayor utilidad del *feedback* en referencia a estos dos puntos:

1º En una primera corrección más que en la versión final (los alumnos tienen la oportunidad de aprovecharlo inmediatamente)

2º En aspectos generales extrapolables a otros trabajos (por ejemplo sobre cómo citar) más que en detalles muy específicos; esta segunda conclusión la vemos mencionada en otros estudios; si el *feedback* se centra en lo muy específico se detrae la atención de estrategias que pueden ayudar a mejorar cómo se trabaja (Poulos y Mahony, 2008).

Cuando se informa en clase a los alumnos sobre la evaluación (normas, criterios de calificación, etc.) tenemos asegurada la atención de los alumnos. Este proceder obliga por otra parte al profesor a ser más explícito, a dejar claras *las reglas del juego,* a esclarecer puntos importantes. Con frecuencia profesores y alumnos tienen percepciones distintas de los criterios de calificación; aunque el profesor los haya explicado suficientemente bien (a su juicio) los alumnos no siempre los entienden e interpretan correctamente. Por otra parte los alumnos tienen también la oportunidad de expresar qué tipo de *feedback* les resulta útil, sobre todo para trabajos futuros.

Ya hemos mencionado en otro lugar a Bain (2006:47) comentando los resultados de los *Harvard Assessment Seminars:* los profesores mejor evaluados por los alumnos mantienen un *alto nivel de exigencia*, pero a la vez *dan múltiples oportunidades* para que los alumnos puedan rehacer sus trabajos antes de ser calificados. Lo mismo comenta Weber (1990); según los alumnos, otra excelente característica (*enormous benefit*) de los mejores cursos (en Harvard) era la oportunidad de entregar una versión preliminar de sus trabajos y poder incorporar la información y críticas del *feedback* en la versión definitiva (Weber, 1990). Lo que queda claro es que *el feedback no es eficaz hasta que el alumno presenta un trabajo en el que las correcciones previas han sido entendidas y tenidas en cuenta* (Taras (2002).

6. La otra cara de la moneda. La evaluación ¿la gran ocasión perdida?

Con frecuencia los profesores hacemos a los alumnos los únicos responsables de su fracaso o de un menor rendimiento del que cabría esperar. El catálogo de razones para culpar a los alumnos puede ser muy amplio:

Les falta motivación,
No dedican tiempo suficiente al estudio,
Faltan mucho a clase,

No se implican a fondo con sus estudios; no estudian *en serio,* estudian sólo *para aprobar*
Nos vienen con déficits importantes de preparación previa,
Puede que incluso pensemos que les falta capacidad para determinados estudios.

Todo esto puede ser verdad, pero Wilson y Scalise (2006) nos recuerdan oportunamente que también debemos preguntarnos *si las prácticas educacionales al uso no estarán contribuyendo a esta falta de preparación y escaso rendimiento* y esto lo dicen en un artículo (*assessing to improve learning in higher education*) que trata precisamente de un programa de *evaluación formativa* en la Universidad de Berkeley, integrando la evaluación (*embedded assessment*) en el mismo proceso de aprendizaje.

Si estamos hablando de la evaluación *para facilitar el aprendizaje,* pensemos también que de hecho la evaluación se puede utilizar de manera muy eficaz para obstaculizar, hacer más difícil o impedir un buen aprendizaje, incluso (como nos recuerda Boud, 1998) cuando las clases impartidas por el profesor pueden calificarse como inspiradas y excelentes. Recordemos que Snyder (1971), en la investigación ya mencionada sobre el *currículum oculto*, cuenta que algunos excelentes alumnos del M.I.T. vivían con conflicto personal una doble necesidad: por una parte aprobar la asignatura (lógico) y por otra… enterarse en serio de la asignatura.

Una mirada rápida a la *otra cara de la moneda* nos puede ayudar a captar las obvias ventajas de la evaluación formativa. Si nosotros, como profesores, queremos influir eficazmente en un cambio a mejor en Educación Superior tenemos que enfrentarnos con los estilos y prácticas de la evaluación que tienden de hecho a obstaculizar el aprendizaje. Con la evaluación podemos estar funcionando sin darnos mucha cuenta en contra de nuestros propósitos declarados (*aprendizaje de calidad, excelencia,* etc.). Probablemente es en el área de la evaluación donde en Educación Superior hay una mayor ignorancia y una práctica más deficiente (Boud, 1995).

Las consecuencias de una práctica deficiente de evaluar a los alumnos repercute de manera muy directa en la calidad del aprendizaje. Los alumnos pueden salir bien librados, no sin cierta dificultad, de una enseñanza de poca calidad, pero de lo que no pueden librarse por definición (si quieren aprobar) es de los efectos de un mal sistema de evaluación (Boud, 1995). Incluso puede suceder que el *cómo* evaluamos vaya en contra de nuestros propios objetivos declarados y estemos fomentando un estudio superficial cuando pretendemos que los alumnos estudien más en serio.

Es posible que tengamos que reconsiderar *cómo, cuándo* y *para qué* evaluamos e intentar *desaprender* modos y maneras que en definitiva son prácticas heredadas, no sometidas a crítica y que pueden ser manifiestamente mejorables. Lo que con frecuencia se hace para evaluar en Educación Superior (tipos de exámenes, de preguntas, de ejercicios, *uso* de la evaluación, modos y criterios para calificar, etc.) en buena medida se basa en malos o discutibles hábitos heredados del pasado y en la falta de un examen crítico sobre lo que hacemos y que difícilmente se puede justificar con criterios educacionales (Boud, 1998, con un título muy expresivo: *assessment and learning: unlearning bad habits of assessment*).

Una manera de conseguir que la evaluación (y exámenes en general) no ayude al aprendizaje es mantenerse en una cierta ambigüedad y dejar que los alumnos intuyan los criterios con los que van a ser evaluados, o qué es lo que tiene una especial importancia; los alumnos brillantes no tendrán muchos problemas pero otros sí los tendrán,

aseguraremos un razonable nivel de fracasos y podremos mantener nuestro prestigio de profesores exigentes. Esto no nos lo decimos así, pero ya hemos recordado al comienzo que puede estar operando eficazmente (y de manera más o menos inconsciente) en algunas culturas educacionales (*aunque ningún fabricante de nada dice que es el mejor porque el 30, o el 50 o el 80 % de su producción es defectuosa*). Podemos olvidar o simplemente no caer en la cuenta de que *el alumno tiene obligación de estudiar, pero no de adivinar qué y cómo debe estudiar*. El misterio pertenece a otros ámbitos, no al de la docencia y del aprendizaje. Otra manera eficaz de poner obstáculos al aprendizaje es ceñirnos a exámenes finales y con una nota ya inamovible.

Hay otras muchas estrategias de evaluación que realmente constituyen obstáculos para *el buen aprendizaje de todos* (puede verse un buen catálogo en Brown, 2004). El problema puede estar en que son modos de actuar espontáneos, rutinarios, heredados (hacemos con nuestros alumnos lo mismo que nuestros profesores hicieron con nosotros), no sometidos a crítica y que aceptamos como normales. Definitivamente hoy día se piensa ya en la evaluación como algo *integrado* en el proceso de enseñanza-aprendizaje y no solamente como el *último episodio* del proceso. Enseñanza y aprendizaje son actividades complementarias y la evaluación es un componente esencial de ambas para mejorar la calidad del aprendizaje y minimizar el número de fracasos sin bajar (o incluso subiendo) el nivel de exigencia.

En tiempos de cambio (*y estos lo son*) no sobra una reflexión tranquila sobre nuestro rol como profesores (*facilitar el aprendizaje*) y sobre nuestras prácticas relacionadas con evaluaciones y exámenes. La evaluación formativa, en cualquiera de sus diversas modalidades, puede ser precisamente un buen contrapeso o antídoto contra prácticas discutibles, en buena medida heredadas, internalizadas y vividas como normales. Es ya hora de ir pensando en *otra normalidad* en nuestro quehacer docente y poner más énfasis en la evaluación formativa para mejorar la calidad del aprendizaje de nuestros alumnos. Solemos quejarnos de falta de tiempo para explicar toda la materia en clase, pero si se trata de conseguir que nuestros alumnos aprendan y aprendan bien, podemos recordar de nuevo la cita puesta en las primeras páginas: *nuestros roles en cuanto profesores deben cambiar radicalmente de manera que podamos concentrar nuestro tiempo y energías más en la evaluación formativa y en proporcionar feedback a nuestros alumnos que en explicarles la materia, ya que los alumnos tienen fácil acceso a muchas fuentes de información* (Brown, 2006).

No se puede negar el cambio bien visible (*si se quiere ver*) que se está produciendo hoy día en cómo se percibe y utiliza (o se intenta utilizar) la evaluación. El acento no está evaluar para calificar sino en evaluar para ayudar a aprender. No deja de llamar la atención el número de congresos o jornadas internacionales dedicadas en los últimos años a la evaluación en Educación Superior; evaluación no *del* aprendizaje (*assessment of learning*) sino como *actividad* de aprendizaje (*assessment as learning*) y *para contribuir* al aprendizaje (*assessment for learning*)[27].

[27] Mencionamos algunas de estas jornadas o congresos fácilmente localizables en Internet: *Enhancing Teaching and Learning Through Assessment* (Universidad Politécnica de Hong Kong, 2005); *Assessment design for learner responsibility* (tres universidades escocesas, Mayo 2007, organizada por el proyecto REAP o *RE-enginering Assessment Practices*), *Evaluation and Learning* (American Evaluation Association, Baltimore, November 2007); *Undergraduate Assessment Symposium* (*to enhance the full undergraduate educational experience through assessment),* North Carolina State University*,* Cary, North Caroline, April, 2009); *Assessment for a Creative World* (The International Association for Educational Assessment Conference, Brisbane Australia, September 2009)

Evaluaciones y exámenes suponen mucho tiempo y muchas energías, tanto para los profesores como para los alumnos y su potencial para mejorar la calidad del aprendizaje es enorme. Si casi la única función de los exámenes es en último término el poner una nota a nuestros alumnos, exámenes y evaluaciones pueden ser *la gran ocasión perdida*.

7. Referencias bibliográficas

ANDRADE, HEIDI GOODRICH (2000), Using Rubrics to Promote Thinking and Learning, *Educational Leadership,* Volume 57 Number 5 http://www.smallschoolsproject.org/pdfs/coho103/using_rubrics.pdf

ANGELO, THOMAS A. and CROSS, K. PATRICIA (1993). *Classroom assessment techniques. A handbook for college teachers.* San Francisco. Jossey-Bass Publishers.

ANTIBI, ANDRÉ (2005). *La constante macabra o cómo se ha desmotivado a muchos estudiante.* Madrid: El Rompecabezas.

BAIN, KEN (2006). *Lo que hacen los mejores profesores universitarios*. Valencia: Universitat de València.

BALCH, WILLIAM R. (2007). Effects of Test Expectation on Multiple-Choice Performance and Sujective Ratings. *Teaching of Psychology*, Vol. 34 (4) 219-225.

BANDURA, ALBERT. (2002). *Self-efficacy. The exercise of control*. New York: W.H. Freeman and Company, 5ª ed.

BIGGS, JOHN (2006). *Calidad del aprendizaje universitario*. Madrid: Narcea.

BLACK, PAUL and WILIAM, DYLAN (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80 (2): 139-148. (Available online: http://www.pdkintl.org/kappan/kbla9810.htm) (consultado 26, 04, 2008).

BLANCO BLANCO, ÁNGELES (2008). Las rúbricas: un instrumento útil para la evaluación de competencias. En Prieto Navarro, Leonor (Coord.) (2008). *La enseñanza centrada en el aprendizaje: estrategias útiles para el profesorado*. Barcelona: Octaedro, 171-178

BLIGH, D. A. (1972). *What's the Use of Lectures*, Penguin Books. y San Francisco: Jossey Bass, 2000.

BLOOM, BENJAMIN S.; HASTINGS, J. THOMAS and MADAUS, GEORGE F. (1971), *Handbook of Formative and Summative Evaluation of Student Learning*. New York: McGraw-Hill (en español, *Evaluación del aprendizaje,* Buenos Aires, Troquel, 1974).

BOSTON, CAROL (2002). The concept of formative assessment. *Practical Assessment, Research & Evaluation*, 8(9). http://PAREonline.net/getvn.asp?v=8&n=9 (Consultado 6, Junio, 2009).

BOUD, DAVID (1995). Assessment and learning: contradictory or complementary? En Knight, P. (Ed.) *Assessment for Learning in Higher Education*. London: Kogan Page, 35-48. http://www.education.uts.edu.au/ostaff/staff/publications/db_9_boud_seda_95.pdf (consultado 17, Dic., 2007).

BOUD, DAVID (1998). *Assessment and learning – unlearning bad habits of assessment.* Presentation to the Conference 'Effective Assessment at University', University of Queensland, 4-5 November 1998. Disponible en http://www.tedi.uq.edu.au/Conferences/A_conf/papers/Boud.html (Consultado 26, Oct. 2006)

BROWN, SALLY (posted February 2006). *Assessment is the most important thing we do for HE students.* London, The Open University, disponible en http://stadium.open.ac.uk/perspectives/assessment (consultado 12, Oct., 2006).

BROWN, SALLY, (2004). Assessment for Learning. *Learning and Teaching in Higher Education*. Issue 1, 2004-05 pp. 81-89 http://www.glos.ac.uk/adu/clt/lathe/issue1/index.cfm (consultado 12, Oct., 2006).

BUTLER, ADAM; PHILLMANN, KAYAH-BAH and SMART, LONA (2001). Active Learning Within a Lecture: Assessing de Impact of Short, In-class Writing Excercises. *Teaching of Psychology*. Vol. 28, nº 4, 257-259.

CARLESS, DAVID (2006). Differing perceptions in the feedback process. *Studies in Higher Education,* Vol. 31 (2), 219-233.

CLARK, RICHARD E. (1983). Reconsidering Research on Learning from Media. *Review of Educational Research*, 53 (4), 445-459

CONNOR-GREENE, PATRICIA A. (2000). Assessing and Promoting Student Learning: Blurring the Line Between Teaching and Testing. *Teaching of Psychology*. Vol. 27, nº 2, 84-88.

COVIC, TANYA and JONES, MAIRWEN K. (2008). Is the essay resubmission option a formative or a summative assessment and does it matter as long as grades improve? *Assessment & Evaluation in Higher Education***,** Vol. 33 Issue 1, p75-85.

CROOKS, TERRY (2001). *The Validity of Formative Assessments*. Paper presented to the British Educational Research Association Annual Conference, University of Leeds, 13-15 September 2001 http://www.leeds.ac.uk/educol/documents/00001862.htm (consultado 25, 04, 2008)**.**

CROSS, K. PATRICIA (1996). Classroom Research: Implementing the Scholarship of Teaching (Presented al the AACP Annual Teachers Seminar, Reno NV. July 14, 1996) *American Journal of Pharmaceutical Education* Vol. 60, Winter 1996 , 402-407 http://www.ajpe.org/legacy/pdfs/aj6004402.pdf (consultado 8, 06, 2009).

DANIEL, DAVID B. and BROIDA, JOHN (2004). Using Web-Based Quizzing to Improve Exam Performance: Lessons Learned. *Teaching of Psychology,* Vol. 31, nº 3, 207-208.

DENMAN, MARIATTE (2005) How To Create Memorable Lectures, *Newsletter*, The Center for Teaching and Learning, Stanford University, Vol. 14, nº 1, 2005 http://ctl.stanford.edu/Newsletter/ (consultado 7, 06, 2009).

DOLNICAR, SARA (2005). Should We Still Lecture or Just Post Examination Questions on the Web?: The nature of the shift towards pragmatism in undergraduate lecture attendance. *Quality in Higher Education***,** Jul2005, Vol. 11 Issue 2, pp.103-115.

DOPPER, SOFIA M. and SJOER, ELLEN (2004). Implementing formative assessment in engineering education: the use of the online assessment system Etude. *European Journal of Engineering Education,* Vol. 29, Nº 2, 259-266.

DRAVICK, DEBORAH, A.G.; WEISBERG, ROBERT, PAUL, LUCY and BUBIER, JENNIFER L. (2007). Keeping it Short and Sweet: Brief, Ungraded Writing Assignments Facilitate Learning. *Teaching of Psychology,* Vol. 34, nº 3, 172-175.

DUNN, KAREE E. and MULVENON, SEAN W. (2009). A Critical Review of Research on Formative Assessments: The Limited Scientific Evidence of the Impact of Formative Assessments in Education. *Practical Assessment Research & Evaluation*, 14(7). Available online: http://pareonline.net/getvn.asp?v=14&n=7 (consultado 7 de Junio, 2009).

ELLERY, KAREN (2008) Assessment for learning: a case study using feedback effectively in an essay-style test. *Assessment & Evaluation in Higher Education*, V. 33, 4, p421-429

ELTON, L. R. B. and LAURILLARD, D. M. (1979). Trends in Research on Student Learning. *Studies in Higher Education,* 4, 87-102.

FELDMAN, ALLAN (2005). Education 695B: *Being a New Teacher,* University of Massachusetts, Secondary Teacher Education Program, http://people.umass.edu/afeldman/beingnewteacher/index.html (consultado 8, 06, 2009).

GIBBS, GRAHAM. and SIMPSON, CLAIRE (2004). Conditions Under Which Assessment Supports Students' Learning. *Learning and Teaching in Higher Education* Issue 1, 2004-05, pp. 3-31. http://www.glos.ac.uk/adu/clt/lathe/issue1/index.cfm (consultado 12, Oct., 2006).

HARTLEY, JAMES and CHESWORTH, KATHRYN (2000). Qualitative and Quantitative Methods in Research on Essay Writing. *Journal of Further and Higher Education*, Vol. 24, nº 1, 15-24.

HUXHAM, MARK (2007). Fast and effective feedback: are model answers the answer? *Assessment & Evaluation in Higher Education*, Dec2007, Vol. 32 Issue 6, p601-611.

INSTITUTE OF HIGHER EDUCATION, SHANGHAI JIAO TONG UNIVERSITY (2005). *Academic Ranking of World Universities*. http://ed.sjtu.edu.cn/ranking.htm (consultado 12, Oct., 2006).

ISAKSSON, SVEN (2008). Assess as you go: the effect of continuous assessment on student learning during a short course in archaeology. *Assessment and Evaluation in Higher Education,* Vol. 33 (1), 1-7.

JAMES, RICHARD; MCINNIS, CRAIG and DEVLIN, MARCIA (2002). *Assessing Learning in Australian Universities, Ideas, strategies and resources for quality in student assessment.* Centre for the Study of Higher Education, The University of Melbourne, Victoria, Australia. http://www.cshe.unimelb.edu.au/ (consultado 5, Junio, 09)

JENKINS, MARTIN (2004). Unfulfilled Promise: formative assessment using computer-aided assessment. *Learning and Teaching in Higher Education,* Issue 1, 2004-05 Pp.67-80 http://www.glos.ac.uk/adu/clt/lathe/issue1/index.cfm (consultado 12, Oct., 2006).

JOHNSON, DAVID W.; JOHNSON, ROGER T. and STANNE, MARY BETH (2000). *Cooperative Learning Methods: A Meta-Analysis*. University of Minnesota http://www.co-operation.org/pages/cl-methods.html (consultado 21, Oct., 2006)

KEMBER, DAVID (2000). Misconceptions about learning approaches, motivation and study practices of Asian students. *Higher Education,* 40, 99-121.

KOGAN, JENNIFER R. and SHEA, JUDY A. (2005). [University of Pennsylvania Health System Philadelphia, Pennsylvania, USA] Psychometric Characteristics of a Write-Up Assessment Form in a Medicine Core Clerkship. *Teaching and Learning in Medicine,* 17 (2), 101–106

KUISMA, RAJA (1999). Criteria Referenced Marking of Written Assignments. *Assessment and Evaluation in Higher Education,* Vol. 24 (1), 27-39.

LANDRUM, R. ERIC (2007). Introductory Psychology Student Performance: Weekly Quizzes Followed by a Cumulative Final Exam. *Teaching of Psychology*, Vol. 34, nº 3, 177-180.

LEEMING, FRANK C. (2002). The Exam-A-Day Procedure Improves Performance in Psychology Classes. *Teaching of Psychology*, vol. 29, nº 3, 210-212.

LIGHT, RICHARD (1990). *The Harvard Assessment Seminars*. Cambridge, Mass: Harvard University (capítulo 3, *Strengthening Colleges and Universities*, disponible en http://www.educause.edu/ir/library/pdf/ffp0604.pdf y http://www.educause.edu/ir/library/pdf/ff0603S.pdf (consultado 14, 04, 2008).

MELMER, R., BURMASTER, E., & JAMES, T. K. (2008). *Attributes of effective formative assessment.* Washington, DC: Council of Chief State School Officers. Retrieved October 7, 2008, from http://www.ccsso.org/publications/details.cfm?PublicationID=362

MERTLER, CRAIG A. (2001). Designing scoring rubrics for your classroom. *Practical Assessment, Research & Evaluation*, 7(25). Retrieved December 17, 2006 from http://PAREonline.net/getvn.asp?v=7&n=25

MORALES VALLEJO PEDRO (2009). *El profesor educador*. http://www.upcomillas.es/personal/peter/otrosdocumentos/RelacionEducativa.pdf

MORALES VALLEJO PEDRO. *El 'one minute paper'* http://www.upcomillas.es/personal/peter/otrosdocumentos/OneMinutePaper.pdf, (última revisión 26, 03, 2008)

MORALES VALLEJO PEDRO (2002) *La relación profesor-alumno en el aula,* 3ª edición. Madrid: PPC

MOSKAL, BARBARA M. (2000). Scoring rubrics: what, when and how? *Practical Assessment, Research & Evaluation*, 7(3). Retrieved December 17, 2006 from http://PAREonline.net/getvn.asp?v=7&n=3

MOWL, GRAHAM (1996). *Innovative Student Assessment*. En DeLiberations on Learning and Teaching in Higher Education, http://www.city.londonmet.ac.uk/deliberations/assessment/mowl_fr.html (London Guildhall University http://www.city.londonmet.ac.uk/ed/) (consultado 12, Oct., 2006).

NARLOCH, RODGER (2006). Benefits of Prelecture Quizzes. *Teaching of Psychology*. Vol. 33, nº 2, 109-112.

NICOL, DAVID and MACFARLANE-DICK, DEBRA (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*. Vol. 31, issue 2, 199-218.

NICOL, DAVID and MACFARLANE-DICK, DEBRA (2004). *Rethinking Formative Assessment in HE: a theoretical model and seven principles of good feedback practice.* Disponible en http://www.heacademy.ac.uk/assessment/ASS051D_SENLEF_model.doc (consultado 20 Oct., 2006) (y en http://www.enhancementthemes.ac.uk/documents/events/20040604/NicolMacfarlaneDickpaperrevised.doc (revisado 14, 04, 2008).

OLALLA MARAÑÓN, Mª ANTONIA (2005). *La corrección cooperativa como instrumento didáctico en la enseñanza de la expresión escrita a alumnos universitarios de español como segunda lengua*. Madrid: Universidad Pontificia Comillas, tesis doctoral.

PADILLA-WALKER, LAURA M. (2006). The Impact of Daily Extra Credit Quizzes on Exam Performance. *Teaching of Psychology,* Vol.33, nº 4, 236-239.

POULUS, ANN and MAHONY, MARY JANE (2008). Effectiveness of feedback: the students' perspective. *Assessment and Evaluation in Higher Education,* Vol. 32, nº 7, 143-154.

RAMSDEN, PAUL, (1992). *Learning to Teach in Higher Education*. London: Routledge.

RAWSON, RICHARD E.; QUINLAN, KATHLEEN M.; COOPER, BARRY J.; FEWTRELL, CLARE and MATLOW, JENNIFER R. (2005). Writing-Skills Development in the Health Professions. *Teaching and Learning in Medicine,* 17 (3), 233–239.

RUST, CHRIST (2007). Towards a scholarship of assessment. *Assessment and Evaluation in Higher Education,* Vol. 32, nº 2, 229-237.

RUST, CHRIST (2002). The impact of assessment on student learning. *Active Learning in Higher Education,* 3 (2) 145-158.

SCOULLER, KAREN M. (1998). The influence of assessment method on students' learning approaches: Multiple choice question examination versus assignment essay. *Higher Education*, 35 (4): 453-472.

SCRIVEN, MICHAEL (1967) 'The Methodology of Evaluation,' in Ralph W. Tyler et al., eds., *Perspectives in Evaluation,* American Educational Research Association Monograph Series on Curriculum Evaluation, No. 1, Chicago, Ill.: Rand McNally, pp. 39-83

SHUELL, THOMAS J. (1986). Cognitive Conceptions of Learning. *Review of Educational Research,* 56 (4), 411-436.

SMYTH, KAREN (2004). The benefits of students learning about critical evaluation rather than being summatively judged. *Assessment and Evaluation in Higher Education,* Vol. 29, nº 3, 269-378.

SNYDER, B. R. (1971). *The Hidden Curriculum.* Cambridge, MA: M.I.T. Press.

STEFANI, LORRAINE (2004). Assessment of Student Learning: promoting a scholarly approach. *Learning and Teaching in Higher Education,* Issue 1, 2004-05, pp. 51-66 http://www.glos.ac.uk/adu/clt/lathe/issue1/index.cfm (consultado 20, Oct., 2006).

STIGGINS, R. J. (2005). From formative assessment to assessment FOR learning: A path to success in standards-based schools. *Phi Delta Kappan, 87* (4). http://www.pdkintl.org/kappan/k_v87/k0512sti.htm (consultado 7 de Junio, 2009)

TARAS, MADDALENA (2002). Using Assessment for Learning and Learning from Assessment. *Assessment & Evaluation in Higher Education*, Vol. 27 Issue 6, p501-510

THE ECONOMIST, *University Rankings, Measuring mortarboards*. November 17th 2007, p. 64

TIAN, XIAOWEN (2007). Do assessment methods matter? A sensitivity test. *Assessment & Evaluation in Higher Education*, Vol. 32 Issue 4, 387-401.

TORRE PUENTE, JUAN CARLOS (2007a). Estrategias para potenciar la autoeficacia y la autorregulación académica en los estudiantes universitarios. En Prieto Navarro, Leonor (Coord.). *La enseñanza centrada en el aprendizaje: estrategias útiles para el profesorado*. Barcelona: Octaedro

TORRE PUENTE, JUAN CARLOS (2007b). *Una triple alianza para un aprendizaje universitario de calidad*. Madrid: Universidad Pontificia Comillas.

WATTS, FRANCES y GARCÍA CARBONELL, AMPARO (Eds.) (2006). *La evaluación compartida: investigación multidisciplinar.* Valencia: Universidad Politécnica de Valencia. Disponible en http://www.upv.es/gie/Publicaciones.html (consultado 21, Oct., 2006).

WEBER, JEROME C. (1990). *The Harvard Assessment Seminars: A Review*. IDP Newsletter Archive: University of Oklahoma Program for Instructional Innovation. (Disponible en http://www.ou.edu/pii/newsletters/archive/c5.html consultado 14, 04, 2008).

WILSON, MARK and SCALISE, KATHLEEN (2006). Assessing to improve learning in higher education: The BEAR Assessment System. *Higher Education.* 52: 635-663.

WINNINGER, STEVEN R. (2005). Using Your Tests to Teach: Formative Summative Assessment. *Teaching of Psychology*, vol. 32 (3), 164-166.

YORKE, MANTZ (2003). Formative assessment in higher education: Moves towards theory and enhancement of pedagogic practice. *Higher Education*, 45: 477-501.

YORKE, MANTZ (2001). Formative Assessment and its Relevance to Retention. *Higher Education Research and Development*, Volume 20, Number 2, pp. 115-126.