

*LA EVALUACIÓN
EDUCATIVA,
SU PRÁCTICA Y
OTRAS METÁFORAS*

ICE- HORSORI

CAPITULO II

LA EVALUACIÓN DEL APRENDIZAJE DE LOS ALUMNOS

La evaluación no ha de ser un medio para conseguir que los sujetos se adapten a los medios instructivos, sino un sistema para adaptar las condiciones de instrucción a los individuos con la finalidad de maximizar su potencial de éxito.

*"The Overall Assessment Prophecy"
(Dochy y Mc Dovell, 1997)*

1. CAMBIOS FUNDAMENTALES EN LA EVALUACIÓN DE LOS APRENDIZAJES DEL ALUMNADO

En los últimos diez años se han producido un conjunto de cambios en la evaluación de los aprendizajes que han supuesto posiblemente la innovación más importante que está afectando al pensamiento actual sobre aprendizaje, escuela y enseñanza. Vamos a tratar de presentarlos brevemente.

1.1. Cambios en el enfoque del proceso de enseñanza-aprendizaje: del énfasis en la enseñanza al énfasis en el aprendizaje

Uno de los cambios fundamentales es el acontecido en el enfoque clásico respecto de la valoración del proceso de enseñanza aprendizaje. La tendencia actual se centra en orientar toda la atención en valorar los resultados de los aprendizajes de los alumnos en detrimento de los medios, en términos de recursos para la enseñanza, que se han invertido para conseguirlos. El centro de gravedad se sitúa en los "outputs" más que en los "inputs". Se cumple con ello uno de los principios básicos del nuevo paradigma organizativo de la educación:

la primacía de las finalidades (Hutmacher, 1999), según el cual se exige que las decisiones y la acción se orienten de una manera prioritaria conforme la voluntad de alcanzar los objetivos establecidos.

Entre los factores que a juicio de los expertos han conducido a este nuevo enfoque podemos citar los siguientes (Mehaffy, 2000):

— La nueva cultura del consumo. El punto de vista que debe prevalecer es el del usuario, en este caso del alumno. En esta nueva cultura existe el convencimiento de que el conocimiento es importante en la medida que es aprendido y no enseñado.

— Las nuevas medidas de petición de responsabilidades desde las administraciones. La "accountability" centra su acción en la comprobación de la asunción de los objetivos previamente establecidos. Los resultados más valorados son aquellos que se expresan en términos del rendimiento de los alumnos.

— Las demandas de la sociedad. Existe un creciente interés de la sociedad en su conjunto por comprobar si los estudiantes finalizan su formación, habiendo adquirido aquellos conocimientos, competencias y habilidades que realmente garanticen una buena inserción en niveles de formación de orden superior o en la vida social y laboral.

1.2. Cambios en los contenidos sujetos a evaluación

La naturaleza de las ejecuciones que son objeto de evaluación se modifica también sustancialmente, ampliando el campo de contenidos y asignándoles y enfatizando sus valores de forma distinta a la tradicional.

Así, a los contenidos académicos clásicos se añaden los contenidos procedimentales y actitudinales ya todos ellos se les complementa con un conjunto de capacidades, habilidades y valores de tipo transversal que también deberían ser objeto de evaluación, bástenos citar con la *Commission on Higher Education* americana (1997) los siguientes:

- Habilidad para pensar críticamente.
- Habilidad para desarrollar estrategias para la resolución de problemas.
- Capacidad efectiva para escribir y comunicarse oralmente.
- Competencia tecnológica, especialmente con bibliotecas y otros recursos de gestión de la información.

- Familiaridad con las matemáticas
- Actitudes asociadas con los valores humanos y juicios responsables.

Listado al que podríamos añadir otros sugeridos desde otras instancias, pero que en definitiva nos están indicando la obligación de afrontar la necesidad de consensuar y desarrollar nuevos objetos de enseñanza y aprendizaje para los sistemas educativos que complementen y resitúen los ya considerados clásicos.

1.3. Cambios en la lógica de la evaluación de los aprendizajes

Es en este contexto donde la evaluación de los aprendizajes cobra una especial dimensión y focaliza su atención en la comprobación de su asunción por parte de los estudiantes, pero no para únicamente proceder a emitir juicios valorativos respecto de los mismos sino para incidir en la mejora de la institución en su conjunto, especialmente en lo que hace referencia a los procesos de enseñanza-aprendizaje.

Nos encontramos en una nueva situación en la que para gestionarla y hacerla viable será preciso que su lógica de funcionamiento amplíe su campo de acción y garantice también la necesaria vinculación entre el proceso de enseñanza y el de evaluación. Así, la *Commission en Higher Education* señala como pasos a desarrollar los siguientes (CHE 1997):

- Identificar los conocimientos y las habilidades que los estudiantes deberían aprender y las competencias que deberían desarrollar para estar capacitados para su posterior aplicación en la vida real.
- Identificar los atributos personales que los estudiantes deberían adquirir y/o desarrollar en una institución educativa.
- Considerar y decidir respecto del conjunto de métodos que habrá que aplicar para realmente capturar información relevante respecto del rendimiento académico y del desarrollo personal.
- Recoger y analizar la información evaluativa requerida para juzgarla efectividad de la Institución en lo que hace referencia a la enseñanza y al aprendizaje.
- Desarrollar un sistema para comunicar la información recogida de forma que facilite la mejora de la enseñanza y del aprendizaje.

Pero no se trata simplemente de una nueva reestructuración de los posicionamientos clásicos, sino que lo que se preconiza es sobre todo un cambio en la naturaleza de los elementos que intervienen. Así, cuando hablamos de conocimientos, habilidades y competencias, es fundamental entender que hacemos referencias a aquellos que son socialmente relevantes y no basta con que lo sean desde el punto de vista puramente académico.

Cuando mencionamos el desarrollo de atributos personales, queremos señalar que no podemos dejar su activación al azar, sino que hay que preverla desde el inicio de la acción educativa en el propio marco de la institución universitaria así como su evaluación.

Al igual que el contenido de los aprendizajes se entienden desde un sentido mucho más amplio que como se ha acostumbrado a hacer desde la perspectiva tradicional, los métodos para su evaluación van a precisar de la misma transformación. Hay que ampliar la tipología de los procedimientos empleados y extender su uso a todos los ámbitos.

También será preciso nuevas formas organizativas de la institución en general y del profesorado en particular. No habrá más remedio que desarrollar el sentido colaborativo de los educadores (las nuevas aproximaciones evaluativas de los aprendizajes sobrepasan la dimensión individual del profesorado), para que estos procesos sean abordados de forma conjuntamente planificada.

También será preciso modificar los sistemas de gestión de la evaluación por parte de los centros que les obligará a explorar nuevas formas organizativas para la recogida, explotación y diseminación de la información evaluativa

Cuando nos referimos al análisis y a la comunicación de la información evaluativa, estamos preconizando cambios sustanciales. El verdadero rasgo diferencial de las sociedades postindustriales no reside tanto en la cantidad de información que poseen, sino en su capacidad de análisis y uso intensivo y extensivo de la misma.

A una institución moderna se le debe exigir capacidad organizativa para recoger la información, estructura para acumularla, uso de modernas formas de explotación de la misma por compleja que ésta sea, agilidad en los procesos de comunicación para extender con rapidez sus aplicaciones y flexibilidad y eficacia para activar los procesos de cambio.

Fundamentalmente, el colectivo más afectado por los cambios en la evaluación del aprendizaje, es el del profesorado. Las nuevas fórmulas evaluativas exigen que los profesores usen sus juicios respecto de los conocimientos de sus alumnos de forma eficiente, entiendan que deben incluirlos en la retroalimentación de su actividad docente y en atender las variables necesidades de aprendizaje de sus alumnos (Tunstall y Gips, 1996). Aprendan a compartir la toma de decisiones

en lo que concierne a los procesos de enseñanza-aprendizaje con sus colegas, con los alumnos y con los padres de los alumnos (Stiggins, 1997; Cippis, 1994).

Todo ello les obliga, en palabras de Hargreaves (1999), a reconstruir su pensamiento en lo que hace referencia al significado del aprendizaje y al papel de la evaluación y a modificar un conjunto de hábitos organizativos

anclados en el pasado.

Es justamente en estas características donde va a residir el verdadero reto de una innovación en profundidad de los procesos de evaluación de los aprendizajes de los alumnos y evidentemente va a ser donde vamos a encontrar las mayores dificultades.

2. LOS RETOS DE LA RECONCEPTUALIZACIÓN DE LA EVALUACIÓN DE LOS APRENDIZAJES

Los cambios que se plantean suponen una reconceptualización de la evaluación y de sus procesos y antes de continuar con la exposición, proponemos tratar de anticipar las implicaciones que su introducción supondrá y analizar con Hargreaves (1999) sus consecuencias desde las perspectivas tecnológica, política, cultural y social.

2.1. Perspectiva tecnológica

Esta perspectiva asume que cualquier innovación es una tecnología con soluciones predecibles y que pueden ser transferidas de una situación a otra. Aplicado al campo de la evaluación de los aprendizajes, la perspectiva tecnológica se focaliza en aspectos tales como la organización, la estructura, la estrategia y las habilidades para el desarrollo de nuevas técnicas de evaluación.

Por tanto, desde esta perspectiva asumimos que un proceso alternativo de evaluación de los aprendizajes como el planteado, constituiría una tecnología compleja que requeriría formación y práctica en desarrollo de sistemas de medición fiables y válidos y contruidos de manera que sean capaces de capturar la complejidad de la información sobre la ejecución de los aprendizajes por parte de los estudiantes (*performance-based assesment*).

Las consecuencias previsibles desde esta perspectiva estarán en las que presenten los profesores que se manifestarán como inexpertos en el tema y faltos de tiempo tanto para su formación como para desarrollar la instrumentación o tecnología apropiada para este enfoque evaluativo.

2.2. La perspectiva política

La perspectiva política en innovación educativa implica el ejercicio y la negociación del poder y de la autoridad y la competencia de intereses entre diferentes grupos (House, 1981). Si nos situamos en el ámbito específico de la evaluación hemos de reconocer que tecla evaluación implica actos de poder y que los modelos alternativos propuestos suponen ceder y compartir parte del mismo.

Torrance y Pryor (1995) *identifican* des aproximaciones evaluativas conceptualmente distintas, que ellos denominan como la evaluación convergente y la divergente.

En la evaluación convergente, lo realmente importante es determinar le que el alumno sabe, comprende o puede hacer respecto de alguna cosa previamente determinada. El poder de la toma de decisiones se sitúa fundamentalmente en el área del profesor.

En la divergente, en cambio, el énfasis se centra en lo que aprende y en cómo lo hace. Los alumnos deben aceptar parte de la responsabilidad en su propio proceso y los profesores en la creación de las condiciones para que éste ocurra, Este ejercicio supone compartir parte del poder, como puede ser el hecho de que los criterios evaluativos deben ser previamente conocidos y a menudo desarrollados de forma colaborativa con ellos.

Huelga señalar que en los nuevos planteamientos, desde el momento que los procesos evaluativos son compartidos y diseñados conjuntamente por los profesores, el compartir el poder no sólo hace referencia a los estudiantes sino que supone una clara redistribución del mismo entre el propio equipo de profesores.

Llevado hasta sus últimas consecuencias, habría que incluir a los agentes políticos, administrativos y sociales como partícipes en los nuevos diseños evaluativos, en la medida que otorgan substantividad y relevancia a ciertos tipos de conocimientos o habilidades y condicionan con ello los nuevos procesos de aprendizaje y de evaluación.

Las dificultades que previsiblemente pueden surgir, se centran en lo reacios que son los sujetos de cualquier colectivo profesional a ceder parte de aquello que considera de su personal incumbencia y la poca cultura organizativa que existe en la realidad de los centros educativos frente al trabajo en equipo.

2.3. La perspectiva cultural

La perspectiva cultural analiza cómo las innovaciones son interpretadas e integradas en el contexto social y

cultural de una institución (Fullan, 1993; Hargreaves, 1994).

Desde la perspectiva cultural, el verdadero reto de las nuevas propuestas evaluativas se sitúa en la necesidad de que los profesores reconceptualicen y reculturalicen sus *juicios* respecto de la naturaleza y el propósito de la evaluación de los aprendizajes.

La innovación en la evaluación de los aprendizajes supone la *aplicación* coordinada de un conjunto importante de distintas estrategias. Una evaluación auténtica es por naturaleza multi-diréccional, directa, su uso ha de ser intensivo de forma que produzca retroalimentación continua y rápida.

Los criterios evaluativos no deben ser ni ocultos ni misteriosos sino transparentes. La actitud de profesores y alumnos no puede ser pasiva sino que todos deben estar activados y perfectamente ensamblados en un proyecto que lo perciben como común. La evaluación desde esta perspectiva cultural implica reevaluación y autoevaluación hasta alcanzar los logros de aprendizaje previstos.

Los objetivos de la nueva evaluación no se sitúan en el control sino en la mejora, en la optimización de la capacidad de aprendizaje de los alumnos, de enseñanza de los profesores y de creación de entornos favorecedores de ambos por parte de la institución.

Es obvio que las dificultades desde esta perspectiva surgirán de la dificultad de cambio en el pensamiento tanto de profesores como de alumnos. Nuestra sociedad está acostumbrada a otras metáforas, y modificar hábitos sobre prejuicios tan enquistados en las rutinas supondrá indudablemente un importante esfuerzo.

2.4. La perspectiva social

Las sociedades modernas avanzadas se han ido construyendo sobre un conjunto de elementos que les confieren unas características muy determinadas y que afectan sustancialmente la innovación educativa. Desde la perspectiva social se posada la necesidad de analizar en qué medida las sociedades modernas avanzadas afectan la innovación educativa.

Así y en primer término, hemos de aceptar que nuestra sociedad constituye una realidad compleja, diversa e incierta. Hechos que aunque hay que percibir como elementos enriquecedores, suponen una enorme complicación en el momento de pretender establecer y desarrollar la misión de la institución y consecuentemente orientar su innovación y establecer los objetivos y la evaluación de su nivel de logro.

En segundo lugar y como un hecho transversal, hay que considerar el nivel de impacto que ha producido la presencia de las nuevas tecnologías, las actitudes que ha despertado frente a ellas y las expectativas que ha generado. Así, en lo que afecta a los procesos de innovación educativa, la sociedad se pregunta sobre los tipos de habilidades básicas que tendrían que desarrollar necesariamente nuestros alumnos (y que consecuentemente habrían de constituir objeto de formación en la educación); cómo se deberían reorganizar los procesos de enseñanza y aprendizaje: de qué manera estarían mejor representados los aprendizajes si aplicarnos en su diseño los avances tecnológicos o cómo se podría integrar la evaluación en el contexto de las nuevas tecnologías, entre otras cuestiones.

Finalmente, hay que considerar también las distancias que se han abierto entre la realidad que vive el sujeto en su ambiente habitual y el de los centros educativos, de tal forma que se admite que cada vez es más habitual el sentimiento por parte del estudiante de sentirse ajeno a todo lo que le rodea en el ámbito escolar (sólo hay que constatar la creciente dificultad por conseguir la participación del estudiantado). La propia diversidad familiar, social y cultural en la que el sujeto se mueve y está inmerso, genera tales diferencias entre ellos que después no es sencillo construir en el espacio escolar un discurso común que sea percibido como relevante para todos.

Muchos profesores, en todos los niveles educativos, admiten la dificultad de contactar con sus alumnos llegándolos a percibir como difíciles de conocer y de predecir (Bigum y Fitzclarence, 1994). Una actitud frecuentemente expresada por parte del profesorado ante el alumnado es la de desconcierto, que evidentemente constituirá uno de los elementos con los que habrá que contar en la construcción de nuevas formas evaluativas.

Las características generales aquí señaladas cuestionan la posibilidad de establecer modelos únicos de evaluación, de aquí la importancia de introducir la necesidad de que la evaluación se sustente sobre una multiplicidad de fuentes y evidencias; debe admitir diversos puntos de vista, compartir juicios y ha de ser permanentemente redefinible. En general es fundamental el aceptar como su carácter más propio la pluralidad.

3. MODELOS DE EVALUACIÓN DE LOS APRENDIZAJES

Finalizábamos el apartado anterior señalando que el carácter más propio de la evaluación actual es su pluralidad: pluralidad de los enfoques, de los contenidos, de los métodos, de los instrumentos, de los agentes, de las audiencias, etc.; vamos a tratar, desde esta aceptación plural, de presentar de forma comparativa los enfoques tradicionales de la evaluación de los aprendizajes frente a las alternativas emergentes.

La realidad es *que* a lo largo de la centuria, que justamente ahora acaba, el enfoque tradicional basado casi en exclusiva en la recogida de información evaluativa mediante test contruidos por el propio profesor o

estandarizados ha sido y continúa siendo el modelo preponderante.

Esta perspectiva tal como se plantea, nos proporciona un tipo de información cuantificada que es fundamentalmente útil para comparar las realizaciones de unos grupos de estudiantes frente a otros y no facilita, en realidad, otros tipos de información también relevante que nos permitan una gestión más eficaz de los juicios extraídos sobre el conocimiento de nuestros alumnos.

Mediante la aplicación de las estrategias y de los instrumentos del enfoque tradicional, no sabemos a ciencia cierta cómo están aprendiendo nuestros estudiantes y cómo les afecta la aplicación del currículum; aún es más, el tipo de capacidades que ponen en juego no siempre están relacionados con el tipo de conocimientos o habilidades que tienen verdadera significación en el mundo real.

Por todo ello surgen movimientos que enfatizan el uso de métodos que faciliten la observación directa del trabajo de los alumnos y de sus habilidades. Este nuevo enfoque es conocido como “evaluación alternativa”.

Como acostumbra a pasar cuando aparece un nuevo término, rápidamente aparecen variaciones en su significado y distintas argumentaciones según sea éste. Así, es bastante habitual utilizar como sinónimos los términos evaluación alternativa, auténtica y de ejecución (*performance assessment*).

Evidentemente, cualquier método que aplique instrumentos de medición del rendimiento distintos de las pruebas de elección múltiple o de construcción basada en el desarrollo de objetivos predeterminados, se puede considerar propiamente como una fórmula alternativa al sistema tradicional. Y prácticamente todos los sistemas alternativos comprometen al alumno en la realización de algún tipo de ejecución.

Sin embargo, Meyer (1992), afirma que la evaluación denominada como de ejecución y la denominada como auténtica no son la misma cosa. Para que una evaluación alternativa pueda ser calificada de auténtica es preciso que los alumnos estén comprometidos en la ejecución de tareas pertenecientes a la vida real o de un gran parecido, y este tipo de prueba no puede ser ideado por el profesor. Por su lado, esta condición no es en absoluto necesaria en una prueba de ejecución.

A modo de ejemplo de una evaluación de ejecución que no es auténtica, podemos tomar el siguiente caso (Meyer, 1992) en el que se compromete a los estudiantes en la realización de un ejercicio de evaluación de su capacidad para redactar. Los alumnos son informados que deberán escribir una historia corta, que será usada para evaluar sus habilidades como escritores. El profesor presenta a los alumnos los elementos fundamentales para la realización del escrito. Con estas indicaciones se les concede tres sesiones de una hora a lo largo de tres días para realizarla. El primer día es dedicado a generar el tópico que servirá de base para la historia, el segundo para crear un primer borrador y el tercero y último para preparar el definitivo.

Ciertamente, los alumnos se ven involucrados en el proceso de escribir, pero pocos escritores en la vida real, escriben historias cortas bajo estos espacios de tiempo tan rígidos o siguiendo las limitadas instrucciones del profesor.

En definitiva, la evaluación basada en ejecuciones y la auténtica son dos modalidades dentro del denominado enfoque alternativo, en el cual, la evaluación de ejecución requiere que los estudiantes, en un momento determinado, demuestren, construyan, desarrollen un producto o solución a partir de unas definidas condiciones y estándares (Rodríguez, 2000), y la auténtica se basa en la realización de tareas reales.

Por su lado, Khattri y Sweet (1996) señalan que adoptar la evaluación de ejecución implica los siguientes pasos a realizar por los estudiantes:

- Estructurar las tareas objeto de evaluación
- Aplicar una información previa.
- Construir respuestas.
- Explicar el proceso que les ha llevado a una determinada respuesta.

Es evidente que las tareas reales incluyen todas las condiciones contenidas en la evaluación basada en ejecuciones, pero añade la realidad como elemento diferenciador, y es obvio que la evaluación alternativa debería tender a la auténtica; sin embargo, es de todo ingenuo pensar que seremos capaces de generar tantos escenarios de realidad como necesidades evaluativas tengamos.

Aún así, los educadores pueden tender a la ideación de aproximaciones con el mayor carácter de autenticidad posible. Frazee y Rudnitski (1995) nos señalan como ejemplos de aproximaciones a la autenticidad: la resolución de problemas, el diseño y la ejecución de experimentos sobre problemas reales, la creación de vídeos, trabajos de campo, desarrollo de demostraciones, la construcción de modelos, la promoción de debates, la publicación de diarios escolares, simulaciones por ordenador, etc., (por supuesto no se trata de una lista exhaustiva). La clave está en recordar que la evaluación auténtica es aquella que evalúa las ejecuciones de los estudiantes bajo situaciones estratégicas que conectan a los estudiantes con las situaciones y las condiciones de la vida real.

Podríamos resumir con Wolf y Reardon (1996), las características más básicas de los dos grandes enfoques que hemos confrontado: el tradicional y el alternativo.

3.1. Características de la evaluación tradicional

- Muestras: pruebas de elección múltiple, de emparejamiento, de verdadero o falso, de ordenamiento, de

completar, etc.

- Juicio evaluativo basado en la recogida de información objetiva e interpretación de las puntuaciones.
- Focalizada fundamentalmente sobre la puntuación del alumno en comparación con la puntuación de los otros alumnos.
- Permite al evaluador presentar el conocimiento del alumno me diante una puntuación.
- La evaluación tiende a ser generalizable.
- Provee la información evaluativa de tal forma que inhibe la acción curricular o instruccional.
- Coloca la evaluación bajo la decisión del profesor u otra fuerza externa.

3.2. Características de la evaluación alternativa

- Muestras: experimentos de los alumnos, debates, portafolios, productos de los estudiantes.
- Juicio evaluativo basado en la observación, en la subjetividad y en el juicio profesional.
- Focaliza la evaluación de manera individualizada sobre el alumno a la luz de sus propios aprendizajes.
- Habilita al evaluador a crear una historia evaluativa respecto del individuo o del grupo.
- La evaluación tiende a ser idiosincrásica.
- Provee información evaluativa de manera que facilita la acción curricular.
- Permite a los estudiantes participar en su propia evaluación.

4. EL PROCESO EVALUATIVO

Independientemente de los distintos enfoques y en líneas generales, el proceso evaluativo de los aprendizajes de los alumnos, atendiendo a la mayoría de los manuales al uso, se articula en las siguientes fases (Rodríguez, 2000): establecimiento de los objetivos de la evaluación asignación de las tareas a realizar por el alumnado, fijación de los criterios de realización de las mismas, explicitación de los estándares o niveles de logro, tomar las muestras de las ejecuciones de cada alumno/a, valoración de dichas ejecuciones, proveer el adecuado “*feedback*” al alumno/a y tomar la decisión, fases que pasamos a comentar a continuación.

4.1. Establecimiento de los objetivos de evaluación

El dominio de contenidos cuyo logro por parte de los alumnos se pretende comprobar, es necesario establecerlo al inicio del proceso evaluativo. Evidentemente, la amplitud y la rigidez con que serán fijados dependerá en gran manera del enfoque, pero al margen de ellas, todo proceso evaluativo parte de algún tipo de concreción respecto de los mismos.

4.2. Asignación de las tareas a realizar por el alumnado

Posiblemente este es uno de los momentos más cruciales de los procesos de evaluación de los aprendizajes, el momento en que el dominio a evaluar debe ser concretado en el tipo de tarea sobre la que se recogerá información evaluativa. Obviamente aquí las diferencias según sea el posicionamiento metodológico serán profundas y distantes. ¿Autenticidad o artificialidad en la naturaleza de las tareas? ¿Proactividad o reactividad en las propuestas concretas? Éstos y otros muchos dilemas deberán ser resueltos en el momento de plantearse las tareas a realizar para evaluar los aprendizajes que supuestamente subyacen en ellas.

4.3. Fijación de los criterios de realización de las mismas

Las tareas ofrecen muchas variantes de realización, más cuanto mas cercanas a la realidad estén. Es preciso establecer criterios sobre su realización de forma que permitan al observador concretar la focalización donde situar la recogida de información evaluativa.

4.4. Explicitación de los estándares o niveles de logro

Es necesario establecer los niveles que permiten afirmar que la tarea se ha realizado convincentemente de acuerdo a los niveles de logro preestablecidos por definición o consenso. No es sencillo esta tarea, que exige, por un lado, un cierto grado de objetivación de realidades complejas y, por otro, experiencias previas que ayuden a fijar dónde residen los niveles de logro aceptables en contextos normativizados.

4.5. Tomar las muestras de las ejecuciones de los alumnos

Es el momento de seleccionar los procedimientos o estrategias para capturar la información. La recogida de muestras de las ejecuciones de los alumnos es otra de las fases fundamentales del proceso. De la calidad de la información recogida dependerá la substantividad de la toma de decisiones. Junto con la fase de determinación de los contenidos a evaluar suponen a nuestro juicio las etapas más comprometidas del proceso.

4.6. Valoración de dichas ejecuciones

Constituye el momento propiamente evaluativo: el de la asignación de juicios a valorar respecto de las ejecuciones realizadas. Exige que previamente se hayan establecido los referentes correspondientes, ya que un juicio exige necesariamente siempre algún tipo de comparación. Como comentaremos con más detalle más adelante, las dos formas más al uso son los juicios de referencia normativa, más propios del enfoque tradicional y los de referencia criterial más usados en la evaluación alternativa.

4.7. Retroalimentar adecuadamente al alumno

La información evaluativa recogida debe ser usada extensiva e intensivamente, muy especialmente para retroalimentar los procesos formativos de los alumnos. Una adecuada retroalimentación está en la base del llamado aprendizaje autónomo; sólo desde la reflexión evaluativa es posible capacitar al alumno para que tome conciencia de su situación en el continuum trazado por el proceso de aprendizaje.

Ubicarse en ese contexto resulta primordial para graduar el esfuerzo, incrementar la atención sobre los elementos que lo precisan y dirigir la acción de forma inteligente y apropiada. El incremento de la consciencia en los procesos de aprendizaje constituye uno de los más importantes retos educativos a los que la evaluación puede proporcionar soluciones.

4.8. Toma de decisiones

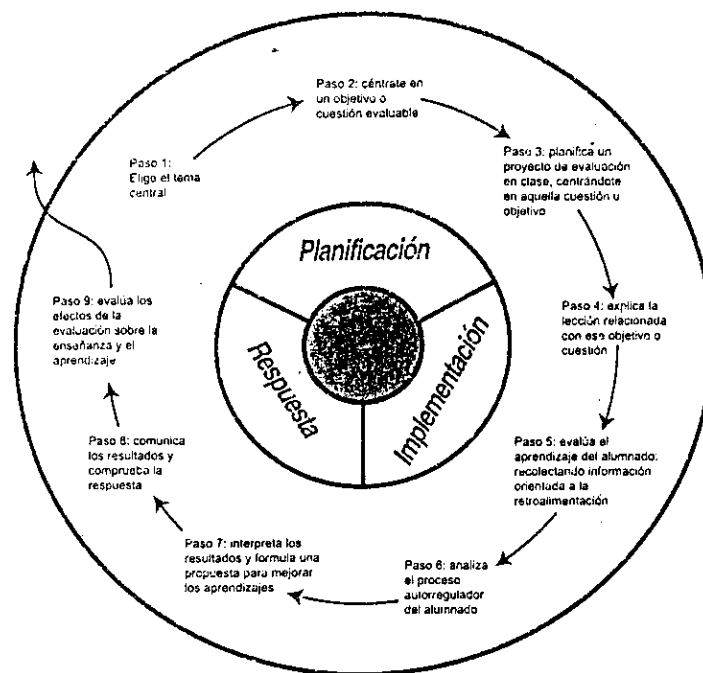
La retroalimentación comentada es una de las consecuencias derivadas del proceso evaluativo, pero evidentemente la información evaluativa deberá situarse en el contexto estratégico de la institución para analizar y desarrollar el cúmulo de decisiones que se podrán derivar de ella bien sea directa o indirectamente.

Los modelos de toma de decisiones en las organizaciones complejas modernas, acostumbran a ser de corte matricial. La información fluye en las redes establecidas de forma que alimenta los diferentes ámbitos de decisión y su estructura facilita no tan sólo los análisis de decisión directa sino también los que son resultado del entramado de interacciones existentes.

Posiblemente, la información evaluativa constituye uno de los flujos de mayor importancia en el funcionamiento de las instituciones educativas, pero desgraciadamente no encuentra, en la mayoría de los casos, culturas organizativas lo suficientemente desarrolladas para que pueda aplicarse en toda su potencialidad.

Ante el riesgo que ello implica: toma de decisiones ingenuas, superficiales, parciales, evidentes, que no tienen en cuenta los niveles de rechazo, etc. o, lo que es peor, que suponen una mejora puntual o inmediata y resultan negativas a la larga o para el conjunto, es necesario extremar la prudencia y se impone el tratar de reflexionar previamente respecto del modelo de calidad de formación que se persigue y del tipo de organización en el que se inserta la medida, de forma que se eliminen al menos sus posibles efectos negativos.

A modo de ilustración presentamos el ciclo evaluativo de aprendizajes en el aula diseñado por Thomas y Cross (1993):



5. PROCEDIMIENTOS TRADICIONALES DE EVALUACIÓN DE LOS APRENDIZAJES

La sustitución del término contenido por el de logro (enfazando con ello la conversión del contenido en aprendizaje) y la ampliación del dominio de este último, añadiendo a los logros académicos y profesionales, los denominados logros personales (competencias genéricas y transferibles a múltiples situaciones (Rodríguez, 2000)), ha colocado en una evidente crisis a los procedimientos clásicos más orientados a la evaluación de enfoque tradicional e incapaces de capturar este nuevo tipo de información evaluativa.

Sin embargo, desde nuestra perspectiva, evaluar el nivel de logro de algún aprendizaje es un proceso en el que la evaluación y el propio aprendizaje están simbióticamente unidos, de forma que ambos discurren paralelamente retroalimentándose continua y respectivamente.

Desde esta visión integrada y continua no siempre es necesario (ni posible) que la comprobación del nivel de logro se realice mediante la evaluación de ejecuciones, sino que en determinados momentos la evaluación de algún tipo de competencia inicial o intermedia, puede realizarse mediante alguna de las formas tradicionales. Con ello estarnos incentivando el uso complementario y combinado de los distintos procedimientos que poseemos, Atendiendo también a criterios de relación coste-beneficio.

Entendemos que desde las nuevas tendencias, la planificación evaluativa debe estar prevista como parte integrante de los propios procesos de enseñanza-aprendizaje en el marco del diseño curricular. Y es desde esta planificación que se incorporan, según sea la naturaleza, el momento y la oportunidad el tipo de testigo que se introducirá para recoger la información evaluativa deseada.

Bajo esta concepción, tienen cabida todos los procedimientos y se apuesta por una pluralidad de uso de los mismos. Creemos y hacemos nuestra consideración desde una perspectiva práctica, que la instrumentación clásica seguirá siendo la más usada; pero, sin embargo, el peso que tendrá en la toma de decisiones será mucho menor que la basada en ejecuciones.

Todo ello nos mueve a presentar en nuestro libro una breve presentación de los instrumentos más utilizados desde las dos perspectivas. Ambas series no pretenden ser en absoluto exhaustivas ni en el número ni en su contenido, especialmente en el caso de la instrumentación alternativa donde la experiencia existente es menor. Hay en el mercado muchos manuales dedicados exclusivamente a ellas con los que sería, en el espacio de este capítulo, absurdo competir.

5.1. Las pruebas objetivas

Las pruebas objetivas son un conjunto vertebrado de preguntas claras y precisas que demandan del alumno una respuesta igualmente breve, generalmente limitada a una elección entre una serie de alternativas que se le proporcionan.

La denominación de objetivas hace referencia a las condiciones de aplicación de la prueba y al tratamiento e interpretación de los datos, pero no indica que sean más objetivas desde el punto de vista de una mejor valoración del rendimiento del alumno.

En el proceso de construcción de una prueba objetiva hay que tener en cuenta los siguientes aspectos fundamentales: la selección de los contenidos, la redacción de las preguntas o ítems (que supone tener en

cuenta la tipología existente), la corrección, puntuación y presentación de la prueba.

En la selección de los contenidos, hay que asegurarse de que todos los que se consideren relevantes respecto al dominio que se va a medir estén especificados de forma operativa y referenciados frente a los objetivos o capacidades a desarrollar en los alumnos.

Se debe determinar el nivel en que se medirán, y establecer el peso que cada contenido ha de tener en el conjunto de la prueba. La concreción de este apartado se realiza mediante lo que se conoce como la tabla de especificaciones, que no es más que una tabla de doble entrada (contenidos-objetivos), en la que se especifica el número de preguntas en las intersecciones. Las normas más habituales que se deben tener en cuenta en la selección de los contenidos pueden sintetizarse en las siguientes:

- La prueba debe considerar los contenidos que se quieren medir y que son explícitos y conocidos por los alumnos. No se deben incluir contenidos que no estén programados.
- Las preguntas deben abarcar todos los puntos importantes del dominio.
- El número de preguntas o ítems para cada aspecto del dominio deberá ser proporcional a la importancia de los contenidos que se van a evaluar.
- Deben estar fijados previamente los objetivos alcanzados y las capacidades desarrolladas cuyo logro se desea comprobar.
- Las preguntas de la prueba han de estar colocadas de manera que encierren una dificultad progresiva. Suelen colocarse primero las más sencillas, e ir incrementando la dificultad hasta llegar al máximo, para finalizar con unas pocas cuestiones de menor dificultad.
- La prueba puede someterse a la crítica de otros especialistas en la materia.
- La mayor parte de las preguntas se derivan de los objetivos específicos propuestos en la programación.
- La elaboración de las preguntas debe poner de manifiesto las habilidades conseguidas por el alumno. De ahí que sea importante una buena definición de los objetivos operativos en términos de conducta.
- Las preguntas deben ser representativas de los contenidos y deben medir los objetivos propuestos, de modo que la prueba en su conjunto tenga una muestra de preguntas representativa de los objetivos y comportamientos que se pretende evaluar.

Características y tipología de los ítems

Antes de entrar en consideraciones respecto a cómo llevar a cabo la redacción de las preguntas o ítems se debe analizar previamente su estructura y características básicas y, después, enunciar los tipos de preguntas más usuales en el ámbito educativo.

Suele entenderse por pregunta una declaración, asunto o tema sobre el que se va a trabajar, en este caso, sería el tema sobre el que va a trabajar el alumno. Algunos autores prefieren utilizar el término ítem en lugar de utilizar términos más propios a nuestra cultura, como serían los de “pregunta” o “cuestión”, porque el ítem no se limita a enunciarse únicamente en forma interrogativa, sino que puede adoptar la forma de una afirmación, una instrucción que se da al alumno, una orden para realizar una tarea, etcétera. En nuestra exposición, los utilizaremos indistintamente.

El ítem consta de una base que presenta una situación—problema y una serie de alternativas de respuesta. Una o varias de estas alternativas deben ser correctas. Las alternativas que son respuestas plausibles pero incorrectas se denominan "distractores"

La respuesta correcta debe colocarse aleatoriamente entre las alternativas, cuyo número ideal se sitúa entre cuatro o cinco, y nunca deben ser menos de tres. De esta forma, se pueden controlar aceptablemente los efectos del azar. Estas pruebas suelen conocerse adjetivadas por el tipo de ítem que las caracteriza.

Existe una gran variedad de ítems pero los más usados en la práctica docente son: los de selección simple, selección múltiple, selección de la mejor respuesta, selección de la respuesta incorrecta, de base común, de ordenamiento, de identificación de gráficos y localización en mapas, de verdadero o falso y de emparejamiento.

Ítems de selección simple

Se caracterizan porque a la pregunta le siguen un conjunto de respuestas entre las que sólo hay una que sea cierta, y otras (los distractores) que son falsas. El alumno deberá contestar evidentemente la correcta.

Ítems de selección múltiple

En los de selección múltiple el sujeto debe elegir las diversas respuestas que sean verdaderas entre distintas alternativas.

Ítems de selección de la mejor respuesta

En el modelo de selección de la mejor respuesta, todas son verdaderas y se debe determinar cuál es la mejor.

En este caso, los distractores son todos verdaderos.

Ítems de selección de la respuesta incorrecta

En este caso se le pide al alumno que marque la respuesta que no es correcta. Los distractores se corresponden justamente con respuestas correctas. En las instrucciones debe estar muy claro que se tiene que elegir la incorrecta, ya que, por lo general, existe una tendencia natural a elegir la respuesta correcta.

Ítems de base común

En los ítems de base común se aprovecha una misma información para generar, alrededor de ella, conjuntos diversos de preguntas.

Ítems de ordenamiento

En relación con los ítems de ordenamiento, el alumno se enfrenta a una serie de hechos o conceptos que aparecen desordenados, que debe ordenar con arreglo a un criterio que previamente ha sido establecido. Este criterio debe estar muy claro, lo mismo que las instrucciones para realizar la tarea. Los criterios más usuales son los siguientes:

Histórico-temporal. En este ítem de ordenamiento se presentan al alumno una serie de acontecimientos históricos o eventos que deberán ser ordenados con criterio cronológico.

—Espacial. Demandan situar los elementos del problema de forma ordenada, en el marco de un espacio determinado. Se utilizan en su construcción términos como: mayor-menor, este-oeste, cercano-lejano, superior-inferior, etcétera.

—Causal. Se solicita del alumno que ordene de forma concatenada los elementos de un conjunto, unidos por la relación causa-efecto.

— Funcional. Exige establecer la secuencia funcional de una serie de hechos fenómenos o actividades.

— Lógico. En este caso, el criterio de dependencia que permite la ordenación se fundamenta en una base racional.

Ítems de identificación de gráficos y localización en mapas

En estos casos se presenta inicialmente un gráfico o un mapa, y se pide al alumno que realice algún ejercicio de identificación o de localización.

Ítems de verdadero o falso

En el caso de los ítems de verdadero o falso, la respuesta de los alumnos exige decidir por separado, respecto a la corrección o incorrección de un grupo de afirmaciones.

Ítems de emparejamiento

Las preguntas de emparejamiento o pruebas por pares se componen de dos listas, una de premisas y otra de respuestas, e instrucciones claras para emparejar ambas listas. Se puede usar una gran variedad de combinaciones premisas-respuestas: fechas y sucesos, conceptos y definiciones, escritores y obras, magnitudes y unidades, cantidades y fórmulas, etcétera.

Ítems de problemas científico-matemáticos

Este tipo de ítems constituye un formato frecuente en materias en las que se estudian relaciones cuantitativas. Es una prueba peculiar que está a medio camino entre las pruebas de formato objetivo y las de formato libre.

Suelen presentar un problema que hay que procurar que represente sólo lo que se quiere comprobar, y en el que no se introduzcan simplificaciones, pero tampoco complejidades no intencionadas. Las respuestas que se proponen deben tratar de comprobar la calidad del proceso seguido y el resultado obtenido.

Estas pruebas constituyen una buena forma de comprobar la comprensión y la aplicación, en contraste con la mera memorización de los hechos.

Corrección y puntuación de la prueba

Posteriormente, hay que proceder a corregir la prueba y a puntuarla. Se debe asignar un valor a cada ítem de acuerdo con la aportación que se considera que hace a la puntuación global; en el supuesto de máxima homogeneidad entre ellos, se otorga un punto por ítem. La suma de las puntuaciones parciales ofrece la

denominada puntuación directa.

Sin embargo, hay que tener en cuenta que en los ítems de elección múltiple el alumno tiene la probabilidad de acertar por razones puramente aleatorias, probabilidad que estará en función directa con el número de alternativas. Esta desviación puede compensarse, al menos desde un punto de vista puramente matemático, aplicando a la puntuación directa (A) un factor corrector según la siguiente fórmula:

Puntuación corregida = $A - E / (n-1)$ donde:

A = número de respuestas acertadas (puntuación directa).

E = número de respuestas erróneas.

n = número de alternativas de respuesta en cada ítem

Los detractores de este sistema de corrección niegan que sea adecuado utilizar la puntuación corregida en lugar de la directa. Alude al factor inhibitor que tiene para el alumno la sensación de penalización si se equivoca.

Sin embargo, es razonable pensar que era el alumno que conoce bien la materia, el efecto real (aunque no el psicológico) de la penalización es muy débil, dado que el alumno no se enfrenta a un ítem con una probabilidad de acierto al azar de $1/n$, sino con una muy superior, dado que hay alternativas que pueden desestimarse de entrada.

Por el contrario, sí no se introduce el factor corrector, el "plus" que reciben en aciertos los alumnos peor preparados es muy elevado como para ignorarlo.

En cualquier caso, antes de aplicar una prueba objetiva se debe indicar al alumno si se utilizará o no la corrección al azar, ya que esta información le es necesaria para decidir qué opciones tomará en la resolución de la prueba.

Presentación de la prueba

En la presentación de las pruebas deben resolverse tres aspectos, que aunque secundarios, han de ser tratados con atención: las instrucciones de realización de la misma, el formato de las preguntas y el de las respuestas.

Las instrucciones hacen referencia a la presentación de las preguntas y de las respuestas, a la duración de la aplicación de la misma, al rigor en las expresiones y, en definitiva, a todos los aspectos que el alumno deberá tener en cuenta para su ejecución.

Las preguntas y las respuestas pueden presentarse integradas en un mismo formato o por separado. Es evidente que en el primer caso, al presentar unidos ambos aspectos, se favorece la realización. Pero si se pretende utilizar la misma prueba en más de una ocasión, será necesario reproducirla en su totalidad, por cuya razón, en algunos casos, se prefiere separar las hojas de respuesta del cuaderno de preguntas. Es evidente que esta cuestión no representa ningún problema cuando el soporte está computarizado.

Crítica a las pruebas objetivas.

Desde su introducción en la práctica educativa, se han sucedido numerosas críticas a su uso, de entre las cuales destacan las siguientes:

—Artificialidad. Se especula que este tipo de prueba favorece fundamentalmente la memorización y que incluso la enfatiza con detalles triviales, por lo que su acción se reduce a la medición de conocimientos artificiales.

—Selección versus construcción. La elección de la mejor alternativa como actividad inherente a este tipo de prueba refuerza más el pensamiento selectivo que los procesos mentales dirigidos a la construcción del conocimiento.

Adivinación y copia. El azar puede constituirse en elemento distorsionador de la medición de los conocimientos, ya que la elección puede estar más dirigida por la capitalización del azar que por la búsqueda de la respuesta correcta. La brevedad y la estereotipificación de las respuestas puede facilitar la copia entre alumnos.

—Coste. Su preparación y diseño es costoso, pues precisa una elaboración compleja, que exige una gran dedicación en tiempo y en recursos humanos e incluso materiales, ya que su presentación requiere cierta formalización, que la hace evidentemente más cara.

La mayoría de estas críticas son discutibles; en cualquier caso, gran parte de ellas pueden ser atribuibles a la incapacidad del constructor de la prueba para servirse convenientemente de sus innumerables posibilidades. Por otra parte de entre sus ventajas más reconocidas pueden destacarse las siguientes:

—Es de fácil aplicación y corrección.

—Tiene la posibilidad de abarcar amplios dominios de aprendizaje.

- La puntuación de la prueba se ve menos alterada por factores extraños al proceso de corrección.
- Admite la posibilidad de múltiples análisis estadísticos de los resultados.
- Facilita el análisis de la información sobre alumnos concretos, grupos de alumnos, logros de objetivos por los mismos, etcétera.
- Posibilita la identificación y cuantificación del nivel de consecución de los objetivos propuestos a los alumnos y la determinación de los objetivos que alumnos concretos o grupos de alumnos deberán trabajar de cara a su refuerzo, afianzamiento o ampliación.

Por la sencillez de su aplicación y corrección, facilita que se pase a gran cantidad de alumnos, y hacerlo en tantas ocasiones como sea preciso, con lo que se facilita todo tipo de estudios y valoraciones en términos absolutos y comparativos.

6. PROCEDIMIENTOS MIXTOS DE EVALUACIÓN DE LOS APRENDIZAJES

Es difícil establecer el límite entre los procedimientos tradicionales y los alternativos. En ocasiones su ubicación no depende tanto de su formato como del uso estratégico que se haga en el contexto concreto de evaluación. Los procedimientos que muy brevemente presentamos a continuación, aunque históricamente nacieron enclavados en la aproximación tradicional, por su estructura interna consideramos que se podrían adaptar también, modificando algunos principios estratégicos, a la evaluación de ejecuciones de los aprendizajes.

6.1 Pruebas libres escritas

Las pruebas de respuesta abierta constituyen una modalidad muy popular y utilizada; entre sus ventajas está el hecho de que se adaptan a cualquier objetivo de aprendizaje. Sus normas más básicas de construcción son las siguientes:

Especificación y referenciación de los contenidos

Es necesario analizar el dominio de enseñanza-aprendizaje con el fin de establecer con precisión los contenidos y las competencias que se pretende evaluar y su relación con los objetivos curriculares.

Elaboración de las instrucciones

Se debe tener en cuenta que las instrucciones han de ser siempre muy claras, y tanto más cuanto menor sea el nivel del alumno. Las instrucciones se han de referir a la presentación de las preguntas y a la orientación del tipo de respuesta que se espera, al rigor en el uso de las expresiones científicas, a la duración del examen y a todos aquellos aspectos que faciliten la estandarización en el pase de la prueba.

Dificultades atribuibles

Conviene reflexionar previamente sobre las limitaciones más importantes de este tipo de prueba, con el fin de proponer recomendaciones que traten de evitarlas o superarlas. Una de las mayores dificultades que presentan estas pruebas es que cubren dominios muy restringidos del conocimiento. Por ello, se recomienda que se utilicen de forma complementaria a otros tipos de ejecuciones, y se tenga presente que, si bien no cubren grandes áreas, pueden hacerlo con gran profundidad, poniendo en juego un gran número de capacidades y actitudes de los alumnos, lo que permite analizar con relativa facilidad ámbitos que con otras pruebas resultaría mucho más complejo.

Otra dificultad atribuible es la indefinición de las tareas propuestas en las preguntas. Es fácil tropezar con pruebas de este tipo en las que no está especificada la tarea indicadora del logro ni la base para juzgarlo. Hay que señalar que, cuanto más detalladas y explícitas sean las preguntas, más objetivas y fiables podrán ser las respuestas y su corrección. También se atribuye a este tipo de prueba la subjetividad en la corrección, aunque si se toman las debidas precauciones, puede suavizarse en gran medida este efecto.

Respuesta pilotada

Con ello se quiere indicar la necesidad de que, anticipadamente a la realización de la prueba, se diseñe el itinerario principal y las posibles ramificaciones por donde, aproximadamente, deberían discurrir las respuestas correctas.

El hecho de pilotar la respuesta ofrece ventajas múltiples. En primer lugar, es una ocasión para que el docente reflexione sobre la calidad de lo que pregunta y la corrección y claridad en la forma de preguntarlo, lo cual puede sugerir matices adicionales en la redacción de la pregunta, haciéndola más o menos fácil y más discriminante.

Otra ventaja contrastada consiste en que el itinerario marcado, estructurado alrededor de los puntos críticos, ofrece una referencia básica para establecer el sistema de corrección y, en el caso de la existencia de varios examinadores o incluso para uno solo, facilita la uniformidad en la aplicación de los criterios de calificación.

Respuestas breves

También es recomendable que la prueba permita un sistema de respuestas breves. Por lo general, se considera que los logros de aprendizaje pueden apreciarse tanto mejor cuanto mayor sea el número de preguntas o tareas propuestas en un examen; sin embargo, también es cierto que tener que enfrentarse a una pregunta compleja y amplia entraña la integración coherente de elementos variados de aprendizaje.

Surge entonces el dilema: ¿Pocas preguntas de amplio recorrido, o muchas y más delimitadas? Resulta complejo ofrecer una receta única al dilema; no obstante, se recomienda dar preferencia a la elaboración de pruebas con más preguntas y con un alto grado de especificación. Las opciones de amplio recorrido deberían ser cuidadosamente sopesadas por el docente o profesor tratando de justificar su necesidad en función de las características específicas del tipo de aprendizaje que se pretende medir.

Tareas definidas

Algunos docentes, de forma ocasional, proponen preguntas de contestación indeterminada, argumentando que lo más importante no son las conclusiones que saquen los alumnos sino la evidencia en que las basan y la lógica de sus argumentos. Sin embargo, no es aconsejable hacer preguntas en cuya respuesta los expertos puedan no estar de acuerdo.

Las preguntas de contestación indeterminada y controvertida son más apropiadas de actividades simultáneas a la instrucción. Pero en las pruebas debe definirse lo que se le pide al alumno tan completa y específicamente como sea posible, para ayudarlo a organizar sus conocimientos y a aplicar sus capacidades. Para ello es conveniente expresar las preguntas sin ambigüedades de manera que los alumnos comprendan totalmente lo que se espera que hagan.

Preguntas opcionales

Por lo general, no es recomendable formular preguntas opcionales a los alumnos en el marco general de la prueba, con el fin de evitar la dispersión, lo cual debilitaría la posibilidad de comparar entre ellos y objetivar los criterios de corrección y calificación. Sin embargo, las preguntas opcionales pueden justificarse cuando los alumnos han tenido diversas oportunidades de instrucción o itinerarios formativos diferenciados en distintos temas y siempre son, en todo caso, una oportunidad de mostrar la riqueza de sus aprendizajes.

Criterios de corrección.

Una vez realizada la prueba debe procederse a la corrección de la misma, y los criterios que se adopten para tal fin están íntimamente ligados a la primera norma que se ha expuesto: la de pilotar la respuesta estableciendo cuál debe ser el recorrido clave que se ha de seguir, pues ello facilitará la asignación de los pesos sobre el conjunto de la puntuación de cada uno de los aspectos que se espera contesten adecuadamente, lo cual permitirá avanzar en la composición de la puntuación global de la prueba. Esta actividad es fundamental si realmente se quiere estandarizar o tipificar en mayor o menor medida este tipo de pruebas.

Una vez elaborada la prueba, se eliminan los puntos en los que se detectan deficiencias o combinaciones de deficiencias del siguiente estilo: omisión de elementos importantes; afirmaciones incorrectas; inclusión de afirmaciones correctas, pero con poca o ninguna relación con la pregunta; inducción a conclusiones no válidas, ya sea por errores en el razonamiento o por una errónea aplicación de principios, o una mala expresión escrita que oscurece el desarrollo y la exposición de las ideas.

6.2. Pruebas libres orales

Las pruebas orales fueron los primeros procedimientos de evaluación de los aprendizajes utilizados. Sin embargo, por razones de economía de tiempo, han caído en desuso y en la actualidad su utilización es muy limitada.

Estructuralmente, son una prueba libre consistente en un diálogo entre uno o varios examinadores y un examinado, con objeto de comprobar en el alumno el dominio de unos conocimientos y competencias determinadas. La opinión que merecen es muy semejante a la de la prueba escrita, y son válidos la mayoría de los principios anteriormente expuestos.

7. PROCEDIMIENTOS ALTERNATIVOS DE EVALUACIÓN DE LOS APRENDIZAJES

Presentamos a continuación propuestas o procedimientos evaluativos que se ajustan más a la evaluación divergente, no predeterminada, a que hacía relación Torrance y Pryor (1995), en la que el instrumento utilizado o la estrategia aplicada, facilita que el énfasis de la evaluación se centre fundamentalmente en lo que aprende el alumno y en cómo lo hace y donde los alumnos deben aceptar parte de la responsabilidad de su propio proceso evaluador.

7.1. El portafolios

El portafolios es posiblemente el procedimiento más popular dentro de la evaluación alternativa. Consiste (Martín, 1997) en una compilación de trabajos del alumno, recogidos a lo largo del tiempo, que aportan evidencias respecto de sus conocimientos, habilidades e incluso de su disposición para actuar de determinadas maneras.

El portafolio contiene una muestra del trabajo del alumno e incluso de evidencias escritas respecto del pensamiento y de los sentimientos del alumno que son suficientemente comprensivos como para generar una representación de cómo lo está haciendo (Chapin y Mesick, 1996).

La muestra de materiales contenidos en el portafolio debe cubrir un extenso período de tiempo, usualmente el curso académico. De esta manera, los alumnos, profesores y padres disponen de suficientes evidencias concretas como para poder analizar y determinar el grado de desarrollo y crecimiento del alumno durante un período específico.

A modo de ejemplo, los alumnos podrían introducir en su portafolios una composición escrita realizada al principio del curso y otra al finalizarlo. Alumnos y profesores podría analizar y ejercer una crítica constructiva respecto del cambio existente entre ambos y determinar si ha habido mejora. Se puede observar también si ha habido evolución en el grado de sofisticación de sus escritos; todo ello proporcionará evidencias de cómo se están produciendo los procesos de aprendizaje y el nivel de logro adquirido en los mismos.

Un beneficio clave del portafolio es que permite al estudiante presentar sus creaciones al profesor, para que pueda juzgar su trabajo de una manera global e integral y no fragmentada, y desconectada de otros aspectos de su propia personalidad. Otra de las grandes ventajas reside en que al tener que escoger los materiales que debe incluir en el portafolio le obliga a tomar conciencia de cuáles son los que mejor reflejan la evolución de los aprendizajes más significativos, lo que evidentemente supone profundizar en su grado de conocimiento y comprensión sobre la intencionalidad de la intervención educativa por un lado y de sus propios niveles de logro por otro.

El alumno con esta actividad obtiene algo más que unas puntuaciones, obtiene evidencias respecto de cómo lo he hecho en realidad a lo largo del curso escolar.

Provee también a los alumnos, profesores y padres de elementos sobre los que conversar de forma concreta e inteligente. Al debatir los contenidos del portafolio, todas las partes toman conciencia clara del grado de desarrollo y progreso del alumno, como persona y como estudiante (Martín, 1997).

Lo materiales que debe contener el portafolio es un aspecto a negociar entre el profesor y el alumno. No existe ninguna lista prefijada. Los alumnos pueden incluir materiales tales como: resúmenes de sus actividades e informes de investigación. Pueden servirse de fotografías o vídeos realizados por ellos, trabajos por ordenador, dibujos o manualidades. Pueden incluir incluso resultados de pruebas objetivas u otros tipos de tests si los han realizado, deberes o tareas escolares realizadas en casa, etc.; el punto central sobre el que pivota la decisión sobre la recolección de materiales, es que la información incluida en el portafolios debe propiciar la generación de una película evaluativa completa de las habilidades, competencias, conocimientos, disposición para actuar y disposición afectiva del alumno (Martín, 1997).

7.2. Otros tipos alternativos de evaluación de ejecuciones

Es del todo imposible listar todas las posibles ejecuciones que podrían ser elementos susceptibles de constituir procedimientos evaluativos alternativos. Es posible generar tantos como creatividad tengan los evaluadores. Así, pueden utilizarse prácticas reales, en las que los alumnos tuvieran que realizar tareas que requieren la aplicación de destrezas en circunstancias semejantes e incluso idénticas a las requeridas en la vida profesional.

Otro tipo podría ser el desarrollo de proyectos, donde durante un período de tiempo el alumno debería realizar una actividad compleja, con un objetivo determinado, que puede incluso tener carácter profesional. Con este tipo de actividad se pueden evaluar habilidades de orden superior que impliquen ejercer responsabilidades, adquirir compromisos personales, poner en práctica hábitos de trabajo individual o de grupo, dominio interdisciplinar de las materias e incluso demostrar capacidades comunicativas. Dentro de este tipo de actividad

se podrían desarrollar investigaciones de campo, experimentos, diseños de organización, etc.

Se pueden desarrollar también ejecuciones en contextos de simulación, bien sea por ordenador o con medios más convencionales. Una simulación supone el planteamiento de un problema cambiante que representa vicariamente un caso real y que debe ser resuelto por el alumno o equipo de alumnos. Cada toma de decisiones respecto del problema retroalimenta el mismo y ofrece los cambios producidos por las soluciones aportadas, con lo que permite analizar la bondad de las propuestas y ver las consecuencias originadas por nuestra acción.

Podríamos añadir a la lista un conjunto innumerable de actividades como puedan ser la resolución de problemas, los estudio de casos, la generación de debates, la confección de un periódico escolar, la creación de una emisora de radio escolar, etc., etc., tal como señalábamos al principio tantas como seamos capaces de imaginar, ya buen seguro que no alcanzaríamos a agotar las propuestas posibles.

En otro orden de cosas y para finalizar este apartado, simplemente señalar que en el marco de la evaluación alternativa, no tan sólo se considera la posibilidad de desmarcarse del marco tradicional, mediante el uso de reactivos que provoquen situaciones de ejecución, de naturaleza sustancialmente distinta a las de las pruebas objetivas, también se modifica las situaciones de prueba atendiendo a la introducción de otros agentes en la función de evaluador, hacemos referencia a la autoevaluación y a la evaluación entre iguales.

Algunos autores sugieren la importancia de que el alumno, especialmente los pertenecientes a grados superiores cobren mayor protagonismo en sus propios procesos de evaluación llegando a responsabilizar casi exclusivamente de la misma. Consideran que es la única vía posible para alcanzar una verdadera reflexión crítica y un aprendizaje en profundidad (Boyd y Cowan, 1985).

Por evaluación entre iguales se entiende aquella disposición evaluativa en la que los sujetos consideran la cantidad, el nivel, el valor y la calidad de los productos surgidos de situaciones de aprendizaje y correspondientes a compañeros de estatus similar (Topping, 1998).

No queríamos obviar esta última consideración referente a la participación de otros agentes evaluativos distintos del colectivo de profesores, pero tampoco pretendemos desarrollarla. Simplemente señalar que creemos que en la medida que se van alcanzando niveles superiores de enseñanza, se hace indispensable la participación de los alumnos, con algún grado de intensidad, en la evaluación de sus propios aprendizajes o de los compañeros. También consideramos que la cultura evaluativa es una cuestión de grado y que conviene iniciarla desde edades tempranas. Todo dentro de un marco de evidente prudencia.

8. EVALUACIÓN DE REFERENCIA NORMATIVA Y EVALUACIÓN DE REFERENCIA CRITERIAL

Los juicios evaluativo se basan necesariamente en algún tipo de comparación y para ejercerla precisan del establecimiento previo de referentes. La evaluación moderna se sirve habitualmente de dos clases: el normativo y el criterial. Tipologías que no son ajenas a los enfoques tradicionales y alternativos que venimos debatiendo desde el principio de este capítulo.

La evaluación de referencia normativa es la más común. Bajo esta aproximación, la ejecución de un alumno en una determinada prueba es comparada con la de otros estudiantes que también pasaron la misma prueba. Los estudiantes como grupo establecen lo que denominamos la norma. Los grupos normativos pueden ser establecidos a partir de características tales como la edad, el grado, el nivel, el género, la localización geográfica, clase social, etnia, o cualquier otro factor de categorización.

Las pruebas estandarizadas de rendimiento constituyen probablemente el grupo de pruebas de referencia normativa mejor conocido. Ellos sirven perfectamente para identificar las personas de acuerdo a su variabilidad en alguna habilidad. Sin embargo, no aportan demasiada información respecto de la situación del sujeto en relación con los niveles de logro en algún contenido o capacidad del currículum.

Las pruebas diseñadas bajo los principios de la referencia normativa (como son las estandarizadas), están pensadas para poder evaluar lo que los estudiantes conocen de una materia particular en relación con otros estudiantes en un momento determinado. Sin embargo, no miden efectivamente lo que ellos saben en términos absolutos. Pero, a pesar de estas limitaciones, es bastante habitual comprobar que se utilizan muy a menudo para evaluar los currícula o los procesos de enseñanza-aprendizaje (Findley y Estabrook, 1991).

En términos generales, las consecuencias más importantes que se derivan de aplicar pruebas basadas en estrategias normativas son las siguientes:

—La puntuación individual se interpreta en función de los rendimientos del grupo con el que mantiene una relación de pertenencia.

—La puntuación permite la comparación entre los distintos individuos

— La puntuación ofrece una idea global de la realización del sujeto, pero no permite establecer acciones de mejora o de corrección respecto al propio individuo, al currículum o al proceso de enseñanza-aprendizaje.

La alternativa a la evaluación normativa es la criterial. Este enfoque está diseñado para informar respecto de la calidad de una ejecución del estudiante o demostración de una habilidad o del grado de comprensión de un concepto por relación con algún tipo de criterio previamente establecido.

Las pruebas construidas mediante este tipo de enfoque nos permiten establecer el nivel de logro del alumno por referencia con un objetivo educativo específico. Nos indican lo que los alumnos pueden o no pueden hacer respecto de un contenido específico, habilidad o actitud. Señalan cambios en los niveles de logro de los aprendizajes a lo largo del tiempo en contraste con las pruebas normativas que sólo capturan información evaluativa sobre los aprendizajes en momentos específicos.

Las pruebas de referencia criterial focalizan sus reactivos hacia tareas o competencias específicas que pueden ser identificadas perfectamente en los apartados de los currícula. También proporcionan información respecto del grado de dominio del alumnado en relación con materiales concretos. Todo ello facilita que se puedan establecer diagnósticos personalizados y eficaces, perfectamente identificables en términos de contenidos curriculares concretos, lo que facilita la tarea de conocer a fondo la situación de cada alumno y la de orientarlos en relación con sus niveles de ejecución.

Podríamos resumir indicando las características más importantes de la evaluación de referencia criterial:

—La interpretación de la ejecución individual se obtiene fundamentalmente por contraste respecto a algún criterio fijado previamente y aceptado como valioso.

—De la ejecución individual nos interesa la composición analítica (los componentes) del contenido objeto de evaluación y la especificación de los procesos implicados en la ejecución.

—Los resultados así obtenidos permitirán posteriormente la puesta en marcha de estrategias de mejora a nivel individual, curricular o institucional.

—La selección de los elementos que componen las pruebas construidas de acuerdo al enfoque criterial, se realiza sobre planteamientos de validez de contenido y no basada en distribuciones teórico - matemáticas.

Las tecnologías para la construcción de pruebas de referencia normativa o criterial son los suficientemente complejas como para no poder ser abordadas en este trabajo. Sin embargo, existen en el mercado magníficos e innumerables manuales al uso, para entrar con detalle en los procesos de diseño y construcción.

9. A MODO DE CONCLUSIÓN

En las sociedades postindustriales se exige cada vez más que la inserción del individuo se produzca equipado con una educación de calidad. Educación en la que el conocimiento no es el único activo importante, sino que deberá contemplar también la formación en otras capacidades, actitudes y valores en la línea de los expresados por Delors (1996) (saber conocer, hacer, convivir y ser).

Ante estos nuevos retos, la evaluación del alumnado, como práctica más extendida para regular la calidad educativa, no puede permanecer anclada en estrategias o enfoques insuficientes o desfasados.

Es necesario ampliar el dominio o los contenidos objeto de evaluación; a los aprendizajes académicos hay que incluir en la evaluación aquellas habilidades, competencias y actitudes de reconocido valor, especialmente las relacionadas con las denominadas capacidades transversales de los sujetos (capacidades-competencias genéricas transferibles a múltiples situaciones).

El énfasis en la evaluación moderna sobre los procesos de enseñanza-aprendizaje se ha desplazado definitivamente desde la enseñanza hacia los aprendizajes, entendidos en ese sentido amplio a que antes aludíamos. Lo que interesa primordialmente no es lo mucho que el profesor enseña sino lo que el alumno realmente aprende.

La denominada evaluación de logros ("*outcomes evaluation*") se ha impuesto como paradigma, y debe entenderse el logro desde una perspectiva doble: como una ampliación del dominio evaluable y como una enfatización de la importancia del logro como asunción efectiva y satisfactoria de un aprendizaje.

No se puede gestionar eficazmente la evaluación, si se pretende atender a sus nuevas responsabilidades, desde la acción individual del profesor o concibiendo las asignaturas como entidades aisladas. La gestión de la evaluación moderna exige el trabajo mancomunado de los profesores y sitúa el currículum como referente básico del diseño educativo y evaluativo.

En consecuencia, la nueva gestión no puede fundamentarse en la recogida de información centrada fundamentalmente en la instrumentación clásica; las pruebas de ejecución basadas en los planteamientos alternativos de la evaluación, deberán complementar o sustituir los modelos más tradicionales, si se quieren procedimentalizar estrategias que capturen información realmente sustantiva.

En los procesos de referenciación evaluativa, las pruebas de referencia criterial deben constituirse en la alternativa más común a aplicar en la nueva instrumentación.

Hay que incorporar a otros agentes, como puedan ser los propios alumnos, en las responsabilidades evaluativas. Sólo desde procesos de autoevaluación es posible desarrollar a fondo una verdadera autonomía en los aprendizajes. Corresponsabilizar en alguna medida a los alumnos en los procesos evaluativos constituirá un importante reto en el futuro.

La lógica de la evaluación, por sí sola, no garantiza la asunción de calidad. Es preciso establecer el nexo

racional entre tres grandes procesos:

- El proceso de enseñanza-aprendizaje.
- El proceso evaluativo.
- El proceso de toma de decisiones.

Desgraciadamente acostumbramos a gestionarlos por separado. Sólo cuando el proceso evaluativo alimenta y se alimenta del de enseñanza-aprendizaje y cuando la información evaluativa que se va generando surte los planes estratégicos, considerados como marcos globales para la toma de decisiones de una institución, se garantiza mínimamente la conducción correcta de los procesos de calidad.

BIBLIOGRAFIA

- BIGUM, G. Y FITZCLARENCE, L. (1994), *Connecting Schools to Global Networks One Way or Another*. Ponencia presentada en la Aplitude 94 Conference, Brisbane.
- BOYD, H. Y COWAN, D. (1985), "Assessment and Evaluation in Higher Education", *Educational Theory*, 10 (3), pp. 42-50.
- Commission on Higher Education (1997), *Framework for Outcomes Assessment*, Middle States Association of Colleges and Schools.
- CHAPIN, J.R. y MESSICK, R.G. (1996), *Elementary Social Studies*, 3^o.ed. New York. Longman.
- CHAPMAN, W. (1991), "The Illinois Experience: State Grants to Improve Schools Through Parent Involvement", *Phi, Delta, Kappa* (enero, 1991), pp. 355-358.
- DELORS, J. et al. (1996), "La Educación Encierra un Tesoro", *Informe a la UNESCO de la Comisión Internacional sobre la Educación para el siglo XXI*. Paris. Ediciones Unesco.
- DOCHY, E. Y MCDOWELL, L. (1997), «Assessment as a tool for learning. *Studies in Educational Evaluation*, 2.3 (4), 279-298.
- FINDLEY, D. y ESTABROOK, R. (1991), "Teacher Evaluation: Curriculum and Instructional Consideration", *Contemporary Education* (verano, 1991), pp. 294-298.
- FRAZEE, B. Y Rudnitsky, R.A. (1995), *Integrated Teaching Methods*. Albany, N.Y. Delmar Publishers.
- FULLAN, M. (1993), *Changes Forces: Probing the Depths of Educational Reform*. London. Falmer Press.
- GIPPS, C. (1994), *Beyond Testing*. Washington, D.C.: Falmer Press.
- HARGREAVES, A. (1989), *Curriculum and assessment reform*. Toronto. OISE Press.
- HARGREAVES, A. (1994), *Changing teachers, changing times: teachers' work and culture in the postmodern age*. Toronto. OISE Press.
- HARGREAVES, A. (1999), Four Perspectives on Classroom Assessment Reform, Barcelona: APAC.
- HARGREAVES, A., Earl, L. Y. RYAN, J. (1996), *Schooling for change*. Philadelphia, N.J. Falmer Press.
- HOUSE, E. (1981), "Three perspective on innovation: Technological, political, and cultural. En Rolf Lehming y Michael Kane (1981), *improving Schools: Using what know*. Beverly Hills. Ca. Sage Publications.
- HUTCHMACHER, W. (1999), "L'avaluació en la transformació de les modalitats de govern dels sistemes educatius. Tendències europees en *Avaluació i Educació*, pp. 15—34. Barcelona. Generalitat de Catalunya. Consell Escolar de Catalunya.
- KHATTRI, N. Y SWEET, D. (1996), 'Assessment Reform: Promises and Challenges, en MB. Kane y R Mitchell (eds), *Implementing Performance Assessment* pp. 1 —21. Mahwah, N.J. Lawrence Erlbaum.
- MARTIN, D.J. (1997). *Elementary Science Methods A Constructivist Approach*. Albany, N.Y.: Delmar Publishers.
- MATEO, J. (1998), "Evaluación e investigación" en J. Mateo (Director), *Enciclopedia Central de la Educación*, pp. 529-648. Barcelona Océano.
- MEHAFFY, G.L. (2000), "The new meaning of teaching in the university of the 21^o. Century", lección magistral presentada en las primeras jornadas sobre Modelos en la Docencia Universitaria. Barcelona Universitat Pompeu Fabra
- MEYER, CA. (1992), "What's the Difference Between Authentic and Performance Assessment" *Educational Leadership*, (mayo 1992), pp. 39-40.
- PAULSON, FL., PAULSON, P.R. y MEYER, CAS. (1991), "What makes a Portfolio a Portfolio". *Educational Leadership* (febrero, 1991), pp. 60-64.
- RODRIGUEZ, S. (2000), *La evaluación del aprendizaje de los estudiantes*. Ponencia presentada al I Congreso Internacional: Docencia Universitaria e Innovación. Barcelona: Institut de Ciències de l'Educació. (U.A.B.,U.B.,U.P.C.)
- STIGGINS, R. (1997), *Student-centered Classroom Assessment*, N.J. Merrill.
- THOMAS, A. y CROSS, P., (1993), *Classroom Assessment Techniques*.

A Handbook for College Teachers. S. Francisco. Ca. Jossey Bass Publishers.

TOPPING, K. (1998), "Peer Assessment Between Students in Colleges and Universities". *Review of Educational Research*, 68 (3), pp. 249-276

TORRANCE, H. (1995), *Evaluating authentic assessment: problems and possibilities in new approaches to assessment*. Philadelphia. NJ. Open University Press.

TORRANCE, H. y Pryor, J. (1995), "Investigating teacher assessment in infant classrooms: Methodological problems and emerging issues. *Assessment in education*, 2 (3), PP. 305-320.

Wolf et al. (1991) "To use their minds well: Investigating new forms of student assessment". En G. Grant (ed.), *Review of Research in education*, 17.

WOLF, D.P. y REARDON, S.F. (1996), "Acces to Excellence through New Forms of Student Assessment" en Baron, J.B. y Wolf D.P. (1996), eds, *Performance Based Student Assessment: Challenges and Possibilities*, Ninety-fifth Yearbook of the National Society of Education Part 1. Chicaho, Ill. University of Chicago Press.