

Redescubrir el Concepto de Evaluación

Grupo Cerpa

Resumimos aquí algunas de las consideraciones que, sobre la problemática de la evaluación, se recogen en R. Neira, T. (Coord.) (2001). *La Evaluación en el Aula*, 2ª ed. Oviedo: Ed. Nobel, y que no suelen dejar indiferentes a los lectores.

Empezamos transcribiendo opiniones que suelen oírse en los centros de enseñanza referentes a aspectos varios relacionados con la evaluación, antes de proponer una definición o, al menos, una descripción de la evaluación, indispensable para abordar su estudio.

En cuanto a las opiniones, destacan en primer lugar los juicios negativos sobre los exámenes, en los que, para algunos, se concentran todos los "diablos" de la educación. Así, se dice que los exámenes producen más efectos negativos que positivos, que son represivos, que reflejan un estilo de enseñanza conservador y autoritario, que producen ansiedad y provocan secuelas muy negativas en los alumnos.

Por el contrario, otros acuden a su experiencia y al sentido común para demostrar que la mayoría de los alumnos no sufre estos daños, ya que, normalmente, las malas calificaciones se asignan a los alumnos que no han estudiado y sus problemas no se van a solucionar suprimiendo los exámenes. En este sentido, casi nadie dice que en la enseñanza reglada puede haber instrucción sin evaluación; es decir, no puede haber enseñanza y aprendizaje si no hay unas metas que alcanzar, y si éstas no se valoran objetivamente.

Desde otro punto de vista, se critica la tendencia a emitir juicios de valor sobre los alumnos a partir de una información muy elemental, sin datos contrastados, a pesar de las importantes consecuencias académicas, personales y sociales que suelen conllevar. También, se identifica con demasiada frecuencia evaluación y calificación. Con el agravante de que a veces la evaluación se usa con fines distintos para los que fue diseñada. Así por ejemplo, surge en los profesores la tentación de preparar a los alumnos para lo que se les va a exigir en las pruebas, o de enseñar las res-

puestas de lo que se va a preguntar o hasta subir la nota. En otros casos, se conceden notas altas como premios y bajas como castigos, convirtiéndolas así en instrumentos de control en clase. Últimamente, sin embargo, sobre todo en la educación obligatoria, se ha desvinculado el éxito del esfuerzo, disfrutando todos los alumnos de los mismos resultados (todos pasan curso) independientemente de su trabajo, por lo que los exámenes están dejando de tener las connotaciones apuntadas.

Se impone, por tanto, someter éstas y otras opiniones a un análisis riguroso para devolver a la evaluación el papel que le corresponde dentro del proceso de enseñanza-aprendizaje. Un método sería tener respuestas claras a una serie de cuestiones; a saber: tendríamos que preguntarnos si es posible evaluar, y en caso de que nuestra respuesta fuese, como es presumible, afirmativa, seguir preguntándonos por el Qué, Quién, Cómo, Cuándo y, sobre todo, Para qué evaluar.

Para contestar a la primera pregunta, es decir, si es posible, conveniente, necesario y útil evaluar, es imprescindible llegar antes a una definición o al menos a una descripción de evaluación.

Para redescubrir el concepto de evaluación y sus funciones auténticamente educativas, es útil rastrear el significado que se le ha ido atribuyendo a lo largo de la historia de la educación. Los autores de R. Neira, T. (Coord.) (2001). *La Evaluación en el Aula*, 2ª ed. Oviedo: Ed. Nobel, ponen de manifiesto en el capítulo 2, que el desarrollo y la constitución de las prácticas evaluativas a lo largo de la historia de la educación han supuesto la pervivencia selectiva de una serie de enfoques y modelos, que se han ido sucediendo, reiterando y superponiendo hasta llegar a la realidad actual.

Por razones de brevedad, y aun corriendo el riesgo de simplificar excesivamente la realidad y de dibujar una caricatura, a veces más expresiva que los detalles de una fotografía, ensayamos aquí una clasificación de los enfoques y modelos de evaluación estudiados en el citado capítulo 2, en dos grandes tendencias;

bien entendido que los rasgos que se atribuyen particularmente a ambas no significan en la realidad exclusión en la otra tendencia, sino énfasis propio.

En un extremo, podemos hablar de una **tendencia cuantitativa**, en la que se encuentran autores tales como Tyler, Thorndike, Skinner, Gagné, etc. Su enfoque educativo en general, y respecto a la evaluación en particular, se enmarca dentro de los siguientes rasgos: El fin de la educación es alcanzar un repertorio de objetivos; la organización escolar está orientada hacia una mayor eficacia, con objeto de lograr mejoras progresivas en las metas y métodos; el currículum está centrado en la estructura de las disciplinas; el profesor es un instructor especialista de su materia; la función principal de la evaluación es medir, jerarquizar y seleccionar, dando prioridad al rigor sobre la riqueza evaluativa; cada etapa educativa se concibe como una preparación para la siguiente; por lo que, a medida que se va ascendiendo en el sistema educativo, la evaluación va eliminando alumnos hasta quedar pocos en el vértice, en una concepción piramidal de la educación.

En concreto, para estos autores, la evaluación es esencial para la educación, porque, en principio, afirman que al menos los productos importantes de la educación, explicitados en los rasgos esenciales de los objetivos de aprendizaje, son medibles. Lo explican de la siguiente manera: si son productos importantes, deben producir un cambio de comportamiento observable; es decir, una persona que ha cambiado considerablemente, debe comportarse de una manera claramente diferente; si es observable, es medible, es decir, existe una base para medirlo; porque lo fundamental de toda medida es una observación verificable en relación con un referente.

A las objeciones y ataques que se suelen formular contra la evaluación cuantitativa, responden con diversos argumentos. Por ejemplo: decir que lo que un "alumno sabe y puede hacer es más importante que una calificación" implica la afirmación falsa de que la calificación no tiene relación con lo que sabe y puede hacer el alumno; y decir que "los exámenes no tienen valor educativo" implica que las calificaciones derivadas de los exámenes no tienen relación con las metas y objetivos propuestos como síntesis de la educación deseada. La realidad, concluyen, es que los exámenes, y en general los instrumentos de la evaluación, ayudan a calificar, y la calificación pretende ser un resumen de los resultados de un proceso de aprendizaje; es decir, la valoración del estado interno del alumno, la "medida" de lo que es, sabe y puede hacer, indispensable para una toma de decisiones adecuada.

Desde un punto de vista opuesto, hablaríamos de la **tendencia cualitativa**, en la que se encuentran autores tales como Dewey, Parlett, Stake, Stenhouse, etc. Su enfoque educativo en general, y respecto a la evaluación en particular, se enmarca dentro de los siguientes

rasgos: El fin de la educación es el desarrollo personal y social; la organización escolar debe estar orientada a favorecer la flexibilidad en programas, objetivos, métodos, etc.; el Currículum está centrado en el cliente, el alumno, en sus necesidades individuales, sentimientos, intereses, estilos de aprendizaje, etc.; el profesor es el motivador y facilitador de todo el proceso de aprendizaje; la función principal de la evaluación es la de diagnosticar y servir de "feedback" para mejorar el rendimiento, dando prioridad a la riqueza sobre el rigor evaluativo; se concibe la educación como permanente en un sistema cerrado del que nadie sale; en vez de seleccionar, la evaluación busca las condiciones que capaciten a cada individuo a alcanzar su techo de desarrollo. Estas afirmaciones recogen algunas de las ideas que ya en abril de 1952 intentó exponer Rogers en la Universidad de Harvard; la conferencia, que se hizo famosa aunque duró apenas 10 minutos, tuvo la virtud de exasperar a los presentes y a muchos de los que luego tuvieron la oportunidad de enterarse de su contenido, y por ello se conoce como la "bomba pedagógica" de Rogers.

Para los autores con tendencia cualitativa, la evaluación educativa clásica ejerce una influencia distractora y hasta nociva sobre alumnos, profesores y el proceso de aprendizaje, porque, en todo sistema físico (partículas) o social (personas), cuando se introduce un elemento extraño para medirlo, se perturba, y puede llegar a destruir el sistema. Ejemplos son el termómetro que modifica la temperatura de un vaso de agua, el encuestador que distorsiona con su mera presencia los comportamientos de los encuestados, o el examen para el alumno.

Además, afirman que los comportamientos mejores no se pueden describir porque apenas si son observables; sólo se miden los resultados menos nobles de la educación porque suelen ser más fácilmente observables.

Sostienen, también, que el individuo es capaz de dirigirse a sí mismo y de encontrar en su propia naturaleza su equilibrio (homeostasis) y sus valores, sin que nadie decida por otros. La "alienación" fundamental del hombre consiste en no ser fiel a sí mismo. Para agradar, para conservar la atención de los demás, el hombre llega a falsificar su experiencia; queda alienado por intentar adaptarse a valores que no son los suyos. El alumno actúa en los exámenes para "agradar" a otros, sus padres y profesores.

En resumen, concluyen que para educar no se deben emitir juicios de valor: no a los exámenes y a las calificaciones. La evaluación debe reducirse a la información y diagnóstico de la realidad escolar.

Parece difícil formular una síntesis de ambas tendencias. Sin embargo, en un intento de buscar las claves que permitan explicar las diferencias y entender-

las mejor, podemos hacer una reflexión sobre los aspectos más subrayados en cada una de ellas.

A veces se pueden confundir medida y predicción (toda decisión lleva implícita una predicción). No es lo mismo medir la creatividad, la motivación, el interés, y por supuesto unos resultados en ámbitos más familiares como el cognitivo y psicomotriz, que predecir, a partir de esta medida, los efectos que se producirán en el futuro. Y la misma predicción tiene diversos alcances: en tiempo, a corto (para la próxima evaluación) y a largo plazo (para el resto de la vida del alumno); y en profundidad, reduciéndose a comportamientos escolares más concretos o llegando a los más humanos, personales y sociales.

La divergencia entre las dos tendencias extremas e incluso las matizaciones intermedias que puedan hacerse pueden atribuirse a que la "cuantitativa" se fija preferentemente en la medida y en una predicción a corto plazo (por ejemplo, para la próxima evaluación, para el próximo parcial, para el próximo curso), que ven como posible: Sí a la evaluación, porque es posible medir aprendizajes; y la "cualitativa" en la predicción a largo plazo (por ejemplo, para su vida profesional), que ven como imposible: No a la evaluación, o más exactamente, no a la evaluación cuantitativa porque no podemos predecir nada y menos imponer nuestras decisiones a otros; hay que "inventar" una nueva evaluación.

Dejando aparte los planteamientos maximalistas de los que eliminan la valoración, los juicios de valor, la toma de decisiones sobre otros y reducen la evaluación al acopio de información neutral, creemos que la evaluación educativa comprende ambos aspectos: evaluar es medir, y evaluar es predecir, aunque más bien a corto y medio plazo. Por lo que optamos por centrar la evaluación no exclusivamente en la medición de los resultados finales, sino en la recogida de información sobre todas las variables que intervienen en el proceso de aprendizaje, como subrayan las tendencias cualitativas, pero cuantificadas en una medida, como reclaman las tendencias cuantitativas. Todo ello encaminado a una toma de decisiones posterior a la medición, con objeto de lograr una "optimización", tanto del proceso como de los resultados finales del aprendizaje. En consecuencia, proponemos y adoptamos como definición de Evaluación Educativa la siguiente:

La evaluación educativa es la medida o comprobación del grado de consecución de objetivos, lo que comporta una recogida de información para emitir un juicio de valor codificado en una calificación, con vistas a una toma de decisiones.

En el documento, ¿Cómo evaluar?, de esta misma página Web, estudiamos estos tres elementos constitutivos de la evaluación educativa: en primer lugar, la Recogida de Información, con un estudio de los distin-

tos tipos de pruebas o instrumentos para recoger dicha información; en segundo lugar, la Calificación y, finalmente, la Toma de Decisiones.

¿CÓMO EVALUAR?

Para evaluar, hay que llevar a la práctica cada uno de los tres elementos de la evaluación; es decir, en primer lugar, recoger datos fiables del grado en que los alumnos alcanzaron unos objetivos; en segundo lugar, codificar en una calificación la valoración de dichos datos y, por último, tomar las decisiones pertinentes, todo lo cual pasamos a estudiar resumiendo los capítulos 4 a 16 de R. Neira (Coord.) (2001). La Evaluación en el Aula, 2ª ed. Oviedo: Ed. Nobel.

I. RECOGIDA DE INFORMACIÓN. INSTRUMENTOS DE EVALUACIÓN

La recogida de información es condición indispensable para evaluar; sin ella, la evaluación quedaría en mera "opinión" del profesor. Los instrumentos para recoger dicha información pueden clasificarse con diversos criterios; así, según la forma de las preguntas, se distinguen tres clases de Instrumentos de Evaluación: Las Pruebas Objetivas, las Pruebas Libres y las Pruebas Diagnóstico.

1. Pruebas Objetivas

Las Pruebas Objetivas (PO) tienen como denominador común que el alumno elige la respuesta entre una serie de alternativas que se le proporcionan. Algunas de sus variedades son las siguientes:

1.1. Verdadero/Falso (V/F):

El alumno califica la formulación de cada proposición como verdadera o falsa.

- A. La sal disuelta en agua se puede recuperar evaporando el disolvente [V]
- B. Los versos de arte menor son los usados por poetas "menores" o de segundo orden [F]

1.2. Múltiple Elección de Respuesta (MER):

El alumno debe seleccionar una de entre un conjunto de posibilidades propuestas. Por su importancia, ampliamos su análisis más adelante.

1.3. V/F Múltiple:

En estos ítems, los alumnos responden identificando la corrección o incorrección de cada afirmación por separado. Estas pruebas se pueden considerar como varios ítems V/F con una base común, o también como un ítem MER con varias alternativas correctas.

La "Monadología" refleja una concepción de la sustancia que, según Leibnitz, podría ser calificada como:

- a) Una teoría mecanicista [F]
- b) Una teoría panteísta [F]
- c) Una teoría dinamicista [V]
- d) Una teoría pluralista [V]

La "Monadología" refleja una concepción de la sustancia que, según Leibnitz, podría ser calificada como una teoría:

- a) Mecanicista
- b) Panteista
- *c) Dinamicista
- *d) Pluralista

1.4. Respuestas Combinadas:

Son una variante que sintetiza los formato V/F y MER. Se proponen varias alternativas como en los ítems V/F múltiples, pero luego se agrupan para pedir, en formato MER, la combinación correcta.

Entre los hechos de intervención de Felipe II en la política interior de Francia, se pueden citar:

- A- La candidatura de Eugenia de Montijo al trono francés.
- B- El tratado de Vervins.
- C- El tratado de la Isla de los Faisanes, en Bayona.
- D- El tratado de Chateau-Cambrésis.

Elegir la combinación correcta:

- a) A y B
- b) B y C
- *c) B y D
- d) A y D

1.5. Emparejamiento (EM):

Estas preguntas, conocidas también como pruebas por pares, se componen de dos listas, una de premisas y otra de respuestas, con instrucciones para emparejar ambas listas.

Escribir la unidad al lado de la magnitud que corresponda:

- 1. Resistencia eléctrica _____
- 2. Intensidad de corriente _____
- 3. Carga eléctrica _____
- 4. Diferencia de potencial _____
- 5. Potencial _____
- 6. Potencia eléctrica _____
- 7. Trabajo eléctrico _____

- (A) Amperio
- (C) Culombio
- (V) Voltio
- (W) Watio
- () Ohmio

1.6. Clasificación (CL):

El formato de Clasificación consiste, de manera semejante a los ítems EM, en una lista de respuestas que hay que clasificar entre las premisas que se proporcionan; pero difieren en que, mientras en los típicos ítems EM cada respuesta se empareja con una

sola premisa, en los ítems CL la misma premisa se puede emparejar con más de una respuesta. Instrucciones: Localizar cada suceso (respuestas ordenadas con números) con el período en que sucedieron (premisas ordenadas con letras):

- A- Antes de Cristo (a.C.).
- B- Después de Cristo pero antes de la invasión islámica (711 d.C.).
- C- Después de la invasión islámica, pero antes de que Colón llegase a América (1492).
- D- Después de que Colón llegase a América, pero antes de la Revolución francesa (1789).
- E- Después de la declaración de la Revolución francesa (1789).

- 1. La batalla de Las Navas de Tolosa C
- 2. El asentamiento agrícola en el Levante español A
- 3. La Escuela de Traductores de Toledo C
- 4. La expulsión de los moriscos D
- 5. El nacimiento de Francisco de Quevedo D
- 6. La instauración de la I República española E

1.7. Respuesta Doble:

Las Preguntas de Doble Respuesta son una combinación de V/F, donde se comprueba la memorización, y un posterior Emparejamiento (EM) para comprobar la comprensión.

En las afirmaciones siguientes, señalar si son verdaderas [V] o falsas [F]:

- El origen de una era glacial debe buscarse en:
 - A- La glaciartosis producida por los polos [F]
 - B- Las variaciones de la órbita terrestre [V]
 - C- Las variaciones de inclinación del eje terrestre [V]
 - D- La fusión de los rayos ultravioleta y el CO2 [F]
 - E- El efecto albedo [V]
 - F- El efecto invernadero [F]

Ahora escoger, de entre las siguientes, las causas que expliquen cada una de las afirmaciones señaladas como verdaderas.

- 1. La eclíptica terrestre.
- 2. El movimiento de campos magnéticos polares, que suponen diversas radiaciones incidentes en la meteorología fría.
- 3. La diferencia de potencial calorífico de polos y ecuador.
- 4. La reflexión de los rayos solares durante el verano por la nieve acumulada durante el invierno.
- 5. La filtración espacial de variantes térmicas estelares.
- 6. La desviación del movimiento rotacional, que supone la incidencia solar en zonas más o menos amplias de los polos.

7. El enfriamiento progresivo producido por la capa de CO₂, que impide la normal refrigeración ambiental del planeta.

Solución: B1, C6, E4.

1.8. Respuesta Limitada (RL):

En este tipo de preguntas, se pide una respuesta corta; por lo que, aunque son preguntas abiertas, suelen considerarse como PO.

¿Cuáles son los morfemas del verbo? R: Tiempo, modo y aspecto.

Si se trazan dos rectas perpendiculares a la misma recta, las dos rectas son, R: Paralelas.

1.9. Problemas Científico-matemáticos:

Los Problemas Científico-Matemáticos, a veces mal llamados Problemas Numéricos, son un formato de prueba frecuente en materias en las que se estudian relaciones cuantitativas. No se suelen clasificar entre las PO, ya que, aunque excepcionalmente se puede presentar en formato MER para que el alumno elija entre varias alternativas, por lo común se pide detallar tanto el proceso para solucionarlo como el resultado. No obstante, es frecuente considerarla como variante del formato de RL, ya que la respuesta que el alumno debe construir suele estar muy acotada.

¿Cuál es el área del rectángulo cuya base es 3 cm y su altura 4 cm?

2. Pruebas Libres

Lo específico de las Pruebas Libres (PL) es la libertad de expresión de los alumnos, que han de organizar sus propios conocimientos, seleccionar lo más importante, integrando en un todo armónico los elementos de sus respuestas. No se reducen a elegir una respuesta, como en las PO. Se pueden identificar las siguientes variedades:

2.1. Respuesta Abierta (RA):

Las preguntas de Respuesta Abierta o Examen Clásico, son el formato de PL más popular, sin que esto quiera decir que estas pruebas constituyan siempre el mejor instrumento de evaluación. Por su frecuente utilización, ampliaremos más adelante algunos de sus aspectos.

2.2. Pruebas Mixtas:

Las Pruebas Mixtas combinan las ventajas del examen clásico con las de las PO. No son dos exámenes sin relacionar, sino uno solo presentado mediante dos formas secuenciadas y complementarias; es decir, primero el alumno ha de responder a un examen tradicional y, luego, ha de resumir y transcribir en PO lo que ya escribió en su RA, con lo que se consigue tener constancia escrita de su razonamiento.

2.3. Examen Oral:

Los exámenes orales han sido los primeros instrumentos de evaluación utilizados. Consisten en un diá-

logo entre un examinador, que plantea preguntas generalmente no relacionadas entre sí, y un alumno, que debe repensar sus respuestas.

2.4. Prueba Práctica Real:

En este tipo de prueba, el estudiante tiene que realizar tareas que requieren la aplicación de destrezas en circunstancias semejantes a las de la vida profesional, para que el profesor pueda observarlas directamente.

2.5. Simulación:

Su fin es medir, por simulación tan fiel como sea posible, la solución de una situación-problema que evoluciona a lo largo del tiempo, y que representa vicariamente un caso real. Sustituye a la observación directa del comportamiento del alumno ante una situación real, especialmente en aquellos casos en que su duración es larga, o encierra peligros para el alumno (es el caso de prácticas de laboratorio) o para la situación simulada (como en algunos casos clínicos).

3. Pruebas Diagnóstico:

La aplicación de estos instrumentos de evaluación supone el dominio de una serie de técnicas particularmente útiles en aquellas áreas donde no es fácil hacer valoraciones por métodos convencionales. Recogemos aquí de manera resumida los aspectos de aplicación más frecuente en contextos académicos de dos de estas técnicas: la Observación y la Encuesta.

1. Observación:

La observación, que puede ser espontánea o estructurada, se muestra como uno de los métodos más eficaces de recogida de información sobre la conducta.

La mayoría de las observaciones que hacen los profesores en el aula son espontáneas. Tales observaciones son muy útiles para una evaluación formativa y suelen constituir el primer paso para, más tarde, hacer una observación planificada. Sin embargo, uno de sus principales defectos es el proporcionar una "visión túnel": vemos a los alumnos que llaman más nuestra atención, y pasan desapercibidas actuaciones menos llamativas, que quizás sean importantes, de otros.

Las observaciones estructuradas o planificadas pretenden obtener informaciones más precisas, por lo que dan resultados más objetivos y son de mayor valor educativo que los obtenidos en una observación espontánea. Para ello, es preciso seguir unas pautas y utilizar unos instrumentos apropiados, tales como Plantillas de observación y Escalas.

2. Encuesta:

La Encuesta pretende obtener datos a partir de un interrogatorio. Se suelen clasificar en dos grupos: Entrevista y Cuestionario.

La *Entrevista* es una comunicación entre dos o más personas, que se lleva a cabo presencialmente y con

unos objetivos previamente establecidos. Por medio de la Entrevista pueden evaluarse aptitudes y actitudes, e incluso objetivos de aprendizaje, en cuyo caso se convierte en un Examen Oral.

El Cuestionario es un tipo especial de Entrevista. Puede utilizarse en una relación presencial, en cuyo caso nos encontraríamos ante una Entrevista estructurada; o puede ser contestado sin intervención directa del encuestador, lo que se podría considerar como una Entrevista a gran escala.

En R. Neira, T. (Coord.) (2001). La evaluación en el aula, 2ª ed. Oviedo: Ed. Nobel, se estudian pormenorizadamente cada una de estas clases de pruebas; aquí recogemos sólo algunos aspectos destacables de dos de las más utilizadas; a saber, las Pruebas de Múltiple Elección de Respuestas (MER) y las Pruebas de Respuesta Abierta (RA) o Exámenes clásicos.

1.2. Pruebas de Múltiple Elección de Respuestas

Los ítems de Múltiple Elección de Respuestas (MER) son PO en las que el alumno debe sólo seleccionar una de entre un conjunto de posibilidades. Constan de una Base y de alternativas (a veces también se les llama distractores); por ejemplo:

BASE: De las actividades que realiza el profesor, ¿cuál pertenece a la evaluación?

DISTRACTORES:

- Seleccionar contenidos.
- Formular objetivos.
- Elegir metodología.
- Recoger datos del proceso educativo

RESPUESTA:

Existen variedades como las siguientes: Escoger la única respuesta correcta (la variedad más aconsejable), Escoger la mejor respuesta, Eliminar la opción incorrecta y las de Base Común (se aprovecha una misma información para hacer varias preguntas sobre ella).

A continuación, aportamos algunas *normas para elaborar la Base del ítem* (posteriormente, trataremos de las Respuestas), cuya función es familiarizar al alumno con el problema que se le propone:

- Es deseable evitar el uso de las mismas preguntas y problemas que se usaron durante la explicación; más bien, conviene introducir situaciones problemáticas nuevas.
- Evitar negaciones y, si se usan, destacarlas.
- Evitar distractores irrelevantes.
- Es preferible redactar el ítem, no como frase incompleta, sino en forma interrogativa; esto ayuda al alumno a entenderlo mejor.
- La respuesta a un ítem MER no debe admitir diferencias de opinión. Sin embargo, se pueden hacer

preguntas que no tengan respuesta completamente correcta para que el alumno escoja la mejor.

- El enunciado de la Base debe incluir una idea completa; es decir, si se eliminan los distractores, el ítem MER debe transformarse automáticamente en una prueba de Respuesta Abierta. Ésta es la característica más determinante de los buenos ítems MER.

Una vez redactada la Base del ítem, hay que escribir la respuesta correcta, así como un conjunto de distractores o respuestas incorrectas; es decir, alternativas lógicas y plausibles. He aquí algunas *sugerencias para generar buenos distractores en ítems MER*:

- Hay que vigilar que los distractores y la respuesta correcta sean homogéneos internamente, es decir, deben referirse a contenidos semejantes.
- Su redacción debe ser breve y concisa, ya que la brevedad simplifica la tarea del alumno centrando su atención en las diferencias esenciales entre las alternativas.
- No dar pistas en unos ítems para otros.
- Los distractores deben ser también homogéneos externamente; es decir, es necesario cuidar la corrección gramatical (género y número). La forma interrogativa en la Base ayuda a no introducir incoherencias gramaticales.
- Hay que evitar reiteraciones de palabras innecesarias e introducir los elementos comunes en la Base. Esto coincide con la norma de formular una idea completa en la Base.
- Aunque de ordinario las alternativas deben presentarse al azar, es conveniente disponer las cifras (números, fechas, etc.) en orden de mayor a menor o viceversa.
- Cuanto más distractores, más discriminará el ítem; pero los buenos distractores no son fáciles de formular; por lo que se recomienda proponer más ítems con menos distractores, que menos con más distractores. Aunque no hay un número mágico de alternativas, la práctica común aconseja escribir 3 o 4 opciones o, lo que es lo mismo, una respuesta y 2 o 3 distractores para cada ítem.

2.1. Pruebas de Respuesta Abierta

Las preguntas de Respuesta Abierta (RA) se conocen también con los nombres de Ensayo, Tradicionales o Examen Clásico.

Hay muchas razones que justifican su popularidad; entre otras, porque se adaptan a todo objetivo de aprendizaje, imposibilitan responder por azar y es más difícil "copiar". En todo caso, los buenos exámenes clásicos suponen la aplicación de una técnica muy minuciosa en su elaboración y, sobre todo, en su corrección.

Así, a la hora de construir preguntas de RA, es necesario tener a la vista las variables que se quieren me-

dir. Subrayamos aquí esta condición, ya que sin objetivos explícitos es casi imposible construir PO; pero se puede creer, erróneamente, que se pueden formular preguntas de RA sin hacer una detenida y previa consideración de lo que se pretende evaluar.

Sin embargo, es la corrección, y su posible subjetividad, una de las causas de la falta de fiabilidad y, por tanto, una de las críticas más frecuentes a las RA. Son numerosos los factores que influyen en que los resultados sean menos fiables; entre otros, fatiga, prisa, humor, primeras impresiones, "efecto halo", desgaste de la sensibilidad (un error juzgado como grande en el primer examen puede parecer insignificante después de haberlo leído muchas veces en los exámenes siguientes), el mucho tiempo que hay que invertir y el grado de atención que requiere del corrector.

Para paliar estas deficiencias, se proponen una serie de normas y métodos de corrección tales como el método holístico, el método analítico, etc.

a) **En el método holístico**, se asigna una puntuación teniendo en cuenta la calidad global de la respuesta, lo que se puede perfeccionar haciendo una corrección por grupos; a saber, el profesor, con una primera lectura superficial de cada respuesta, hace una clasificación provisional y coloca los exámenes, por ejemplo, en 3 grupos, [Alto-Medio-Bajo]; después de esta primera lectura rápida, se vuelven a leer más detenidamente los exámenes de cada montón para colocarlos en grupos subdivididos como [Muy buenos-Buenos-Satisfactorios-Malos-Muy malos] para asignarles las notas definitivas.

b) **En el método analítico**, se descompone cada pregunta en partes, encerrando en cada parte una serie de categorías, que deben ser identificadas y calificadas separadamente. La presencia o ausencia de estos elementos en las respuestas dadas por el alumno determina la puntuación; es decir, el profesor, al corregir, añade o suma puntos por cada elemento que aparece, o resta puntos por cada elemento inexacto, irrelevante o que falta.

A primera vista, el método analítico parece mejor; sin embargo, no siempre es así, ya que en algunos casos es tan laborioso y las ventajas son tan poco significativas que no merece la pena ponerlo en práctica, pues se obtienen los mismos resultados y las calificaciones son tan fiables con el método holístico, más simple y más rápido.

c) **La corrección horizontal supone corregir por preguntas**, en vez de por alumnos. Esto significa que el profesor corrige primero las respuestas de todos los alumnos a la pregunta 1, antes de pasar a la pregunta siguiente. Al concentrarse en una cuestión cada vez, ayuda a no tener que hacer transposiciones mentales al ir cambiando de pregunta.

d) **Si es posible, hay que ocultar a quien corrige**

la identidad del alumno. Esto puede reducir la posibilidad de que prejuicios, el nombre, la fama o el "efecto halo", asociado con las calificaciones del mismo alumno en otras preguntas del examen, influyan en la asignación de calificaciones.

II. LA CALIFICACIÓN

Después de haber estudiado el primer elemento de la evaluación, es decir, la recogida de datos con un estudio de los distintos instrumentos para recoger dichos datos, abordamos ahora el segundo, la valoración, codificada generalmente en una calificación. Si bien en la práctica docente se valoran los datos conforme se van recogiendo durante la corrección de pruebas -se califica al corregir-, conviene tener en cuenta que el mero resultado de dicha información, aunque sea numérico, no es aún una calificación.

La calificación es un juicio de valor en el que se comparan los datos obtenidos en la recogida de información, que reflejan el "cómo es" de lo que se va a evaluar, con unos determinados parámetros de referencia, que indican el "cómo debería ser"; es decir, los criterios con que se van a comparar los datos obtenidos.

Elegir esta base de referencia es una decisión fundamental en la tarea de todo profesor y conlleva consecuencias significativas, por lo que es importante entender las diferencias de los procedimientos para corregir pruebas que se presten a un tipo de valoración más que a otro. Aunque son varias las formas para llegar a la valoración de datos, se suelen citar dos: valoración "normativa" y "criterial".

1. **Valoración "Normativa"**: La valoración "normativa" supone comparar los resultados de una persona con su grupo de referencia.

Se reparten los alumnos en categorías y, a cada categoría, se le asigna una calificación. Mediante una valoración "normativa", los alumnos quedarán siempre clasificados en dos grupos: los admitidos y los excluidos; sin la seguridad de que los admitidos sean, además, competentes.

2. **Valoración "Criterial"**: Se basa en unos criterios absolutos que permitan valorar la competencia de una persona y describir la posición del aprendiz respecto a los objetivos de un programa y no de un grupo. Se califica según el porcentaje de objetivos alcanzados.

Este tipo de valoración permite evaluar la competencia de una persona en una tarea dada, sin tener que compararla con otras. Es posible, por tanto, declarar aptos a todos los alumnos; y también, teóricamente, suspender a todos.

En general los profesores, por razones teóricas, prefieren estas calificaciones absolutas a las relativas; sin embargo, encuentran difícil, y en algunos casos arbitraria, la tarea de establecerlas. Los métodos para

hacer valoraciones absolutas las estudian en el capítulo 14 los autores de R. Neira (2001).

III. TOMA DE DECISIONES

Nos centramos ahora en la toma de decisiones como tercer elemento esencial de toda evaluación. Una decisión se define como una elección entre distintas acciones posibles. Una decisión, también las educativas, la puede tomar una persona, como en estos momentos es el caso del profesor único en la Educación Primaria y la Enseñanza Universitaria, o un grupo, como es el caso de las Juntas de Evaluación exigidas por el ordenamiento legal para la Educación Secundaria.

En los párrafos siguientes, resumimos la problemática de la Toma de Decisiones, a través de aspectos tales como: Clases de decisiones, Proceso racional de toma de decisiones y Herramientas de formulación y selección de alternativas, para terminar con algunas sugerencias sobre un proceso práctico aplicable a las Juntas de Evaluación.

1. Clases de decisiones: Simplificando el análisis, se podría afirmar que existen dos clases de decisiones: las selectivas y las modificativas.

a) Las decisiones selectivas están relacionadas con las funciones atribuidas a una Evaluación Sumativa. Se suelen considerar como decisiones selectivas, entre otras, las siguientes: La promoción o repetición de un curso o asignatura, la concesión de premios, etc.

b) **Las decisiones modificativas están relacionadas con la Evaluación Formativa**, ya que tienen en cuenta el rendimiento del alumno para determinar actividades de recuperación, modificaciones diversificadoras en la programación o en la metodología, etc. Estas decisiones pueden ser ejecutivas, es decir, aquellas en las que se asume la responsabilidad de llevarlas a la práctica, y consultivas, cuando las soluciones que pueden subsanar los malos resultados rebasan la responsabilidad del que las toma, por lo que la decisión se limita a la elaboración de un informe dirigido al órgano competente.

2. El enfoque racional de toma de decisiones: Las decisiones suelen ser más eficaces y la probabilidad de acierto mayor, sobre todo en el caso de la toma de decisiones grupal, cuando se tienen en cuenta unas reglas. Estas reglas se materializan en el enfoque de toma de decisiones racional a lo largo de las siguientes tres fases:

1ª Fase. Problema: Una toma de decisiones es una consecuencia de la aparición de un problema que hay que resolver. Para llegar a una buena decisión es preciso, por tanto, detectar el problema, analizarlo para recopilar datos que ayuden a definir la situación problemática y traducirlo a objetivo en términos de resultado deseado.

2ª Fase. Soluciones: Esta segunda fase empieza estableciendo una relación entre el problema suscita-

do y sus posibles causas. A continuación, hay que identificar soluciones, sugeridas precisamente por las causas del problema, que sirvan para alcanzar el objetivo propuesto. Una vez exploradas las soluciones opcionales, se trata, por fin, de elegir una alternativa que ayude mejor que cualquier otra a superar las posibles deficiencias.

3ª Fase. Plan de Acción: Una cosa es tomar una decisión y otra llevarla a la práctica. Por lo tanto, una vez seleccionada la mejor alternativa, es necesario elaborar un plan de acción en el que queden organizadas unas tareas y un sistema de seguimiento.

3. Herramientas de toma de decisiones: Se han elaborado diferentes técnicas para cada una de las fases de la toma de decisiones; a saber, de Análisis, de Selección y de Plan de acción.

a) **Técnicas de análisis:** Son herramientas de exploración cuyo fin es precisamente detectar posibles alternativas de problemas, causas, soluciones, etc.; por ejemplo, el "Brainstorming" para grupos pequeños y el "Phillips 66" para grupos más numerosos.

b) **Técnicas de selección:** Son herramientas que ayudan a abandonar alternativas de problemas, causas, soluciones, etc., para concentrarse en unas pocas; por ejemplo, el Voto ponderado.

c) **Técnicas para diseñar un plan de acción:** A este fin se puede emplear una versión simplificada del PERT ("Program Evaluation and Review Technique), y que se refiere al camino crítico de un calendario de un plan de acción que especifica las fechas, personal implicado, orden de las operaciones, etc., para conseguir una meta determinada.

Es conveniente también tener en cuenta la ley de Murphy: "Todo aquello que tenga posibilidades de fracasar, fracasará". Por eso, el plan se puede completar anticipando los posibles fallos detallando lo que puede fallar, cómo se puede evitar que ocurra y cómo se rectificará el fallo si ocurre.

4. Juntas de evaluación: Todo grupo debe adoptar un proceso racional (inicialmente debe ser sencillo) de toma de decisiones y obligarse disciplinadamente a seguirlo sin improvisar excesivas modificaciones sobre la marcha.

La estructura de cada fase del proceso debe utilizar, en primer lugar, una técnica de análisis que permita formular todos los posibles problemas, causas, soluciones, etc.; para reducirlos posteriormente a las alternativas más viables mediante una técnica de selección. Es decir, un proceso en "zigzag" que sucesivamente se abra de modo creativo a todas las alternativas para luego sacrificar la mayoría de ellas con objeto de concentrarse en el problema más importante y en la solución más viable. A manera de ejemplo, se sugiere el siguiente para un escenario restringido al ámbito de los aprendizajes, que es el más frecuente en las

Juntas de Evaluación.

1. *Decisiones Previas:* Para agilizar el proceso formal de toma de decisiones en la Junta de Evaluación, se deben tomar en el foro adecuado una serie de decisiones previas. Son fundamentalmente tres:

a) Además de clarificar los objetivos, metodología, etc., hay que decidir si se ofrece a los alumnos la opción de elegir entre asignaturas optativas; y, luego, dentro de cada asignatura, si se les ofrece un único programa a todos, o también programaciones adaptadas o diversificadas, en las que los objetivos se calificarían con una gama completa de notas (de 0 a 10) o, sencillamente, con un apto/no apto.

b) Hay que atribuir una función a cada miembro del grupo y diseñar la dinámica de las reuniones.

c) Decidir qué datos se van a llevar a la Junta, es decir, las calificaciones numéricas o literales, simples o matizadas.

2. *Proceso con los pasos siguientes:*

a) Datos sobre alumnos con calificaciones insatisfactorias.

b) Enumerar los problemas planteados por los datos aportados, filtrando sólo los susceptibles de decisiones ejecutivas; es decir, los problemas relacionados con el aprendizaje. El resto de problemas, enviarlos al órgano competente (Jefatura de Estudios, Seminario, Gabinete de Orientación, etc.). Seleccionar, por fin, un problema de aprendizaje.

c) Formular todas las posibles causas del problema, así como todas las soluciones sugeridas por las causas detectadas, para terminar seleccionando la mejor alternativa.

3. *Plan de acción:* Concretar las acciones, personas responsables de cada acción, sistema de seguimiento, recursos y tiempos.

IV. CLASES DE EVALUACIÓN

Son varios los criterios que se pueden adoptar para determinar las distintas clases de evaluación. Entre otros, conviene citar el sistema que se evalúa (dando lugar a la evaluación interna y externa), el momento de aplicación (evaluación inicial, intermedia y final) y, sobre todo, los fines de la evaluación, que dan lugar a la evaluación formativa y a la evaluación sumativa.

1. *Evaluación Formativa:* Si la evaluación se utiliza para mejorar el proceso de instrucción, la evaluación tiene una función formativa, con las características siguientes:

a) Esta evaluación detecta deficiencias en el aprendizaje sin limitarse a jugar un papel sancionador y convertirse en una serie de obstáculos que los estudiantes tienen que vencer.

b) Esta evaluación implica no sólo un análisis de los aspectos parciales no asimilados de los objetivos; sino que se preocupa de las causas de los fallos, tales como aptitudes y capacidades básicas, habilidades aplica-

das, estilo y estrategias de aprendizaje no adecuadas, así como la posibilidad de entorpecer otros factores contextuales, tales como aspectos médicos, psicológicos, familiares y ambientales que pueden influir en el aprendizaje.

c) Lo más importante es su característica de no ser un veredicto final, sino más bien un diagnóstico de la situación del alumno durante su proceso de aprendizaje respecto a cada objetivo por separado para ayudarle a alcanzarlos; llegando, más allá de los objetivos, a analizar las necesidades y valorar hasta qué punto se han satisfecho. Es, por un lado, un "feedback" para los alumnos; pero también es un complemento natural de una buena enseñanza, ya que es un "feedback" para los profesores.

d) Como consiste en evaluar tanto las estrategias de aprendizaje como el rendimiento conseguido por el alumno durante la instrucción, desde el momento en que aborda el aprendizaje hasta que termina, la evaluación debe ser permanente.

En resumen, la evaluación Formativa es una evaluación permanente para determinar el grado de adquisición de cada objetivo, y detectar tanto los aspectos NO asimilados, como las causas para reajustar y "optimizar" los logros y el proceso de enseñanza-aprendizaje.

2. *Evaluación Sumativa:* Si la evaluación se utiliza para tomar decisiones ejecutivas finales sobre los resultados del aprendizaje, la evaluación tiene una función sumativa, con las características siguientes:

a) Las decisiones asociadas a la evaluación Sumativa exigen que el elemento medir quede cuantificado en una calificación criterial final, mientras que el elemento toma de decisiones suele quedar reducido a la decisión selectiva de poder pasar de tema, de curso, etc.

b) La evaluación Sumativa es una acción "después de" que se preocupa de dar un juicio del trabajo del aprendiz en relación con algo fijo, mediante la comparación de los objetivos y los resultados. Pero dicho juicio no da pistas e información suficientes sobre las necesidades de ajuste y cambio para que el alumno llegue a las metas propuestas no alcanzadas; aparte de que no tiene ya la oportunidad de recuperar las eventuales deficiencias de aprendizaje, como sucede en la evaluación Formativa.

c) Los controles son más espaciados que en la evaluación Formativa. Se suelen realizar al final de una Unidad Didáctica, a final de trimestre, de curso, etc.

En resumen, la evaluación Sumativa es la evaluación final que emite un juicio sobre el alumno en relación con el resultado del aprendizaje, reflejado en una calificación globalizada de un repertorio de objetivos, y de la que se sigue una toma de decisiones ejecutiva.

En realidad, los profesores practicamos de modo espontáneo la evaluación Sumativa, y aprovechamos la información en ella obtenida para llevar a cabo ocasio-

nalmente el "feedback" de la evaluación Formativa. Sería conveniente incrementar de forma progresiva esta última hasta aproximarse a la utopía que defiende la permanencia en el tiempo de la evaluación Formativa y la reserva de momentos para la evaluación Sumativa.

V. EL PROCESO DE LA EVALUACIÓN

La evaluación debe realizarse según un proceso que sirva de guía de trabajo para construir una prueba, aplicarla, valorarla, tomar las decisiones oportunas e informar de ellas a los alumnos, lo que puede materializarse mediante los pasos siguientes:

1. Planificación: Para planificar un examen, deben seleccionarse los objetivos importantes de la programación, lo que condiciona el número de preguntas.

2. Formato de las preguntas: Además de determinar el número de preguntas de un examen, hay que decidir la forma de cada una de ellas. En la obra citada, se ofrecen sugerencias para seleccionar el formato de las preguntas según la capacidad que encierra el objetivo que se va a evaluar.

3. Construcción de las pruebas: Cada formato supone unas normas de construcción.

a) Dificultad: En general, tanto los ítems demasiado fáciles como los demasiado difíciles no contribuyen especialmente a discriminar los logros de los alumnos.

b) Orden de las preguntas: Como el tiempo de examen suele ser restringido, es preferible presentar las preguntas en orden de dificultad creciente, pues es lógico suponer que las preguntas más fáciles del principio reducirán la ansiedad de los alumnos ante el examen. Parece también lógico agrupar las preguntas referidas a un mismo conjunto de contenidos. **4. Ejecución:** La celebración de los exámenes plantea pocos problemas; no obstante, tienen tan importantes consecuencias que no está de más que profesores y alumnos sepan cómo pueden realizarlos mejor. A continuación transcribimos algunas consideraciones que pueden ser útiles.

a) Instrucciones: Dado que el tiempo de examen resulta siempre escaso para la mayor parte de los alumnos, es necesario aprovecharlo al máximo. Por ello conviene que las instrucciones, tales como tiempo disponible, cuántas y qué tipo de preguntas va a tener, cómo se va a calificar, cómo usar la hoja de respuestas, instrucciones específicas para responder alguna pregunta, material que se permite utilizar (diccionario, calculadora, etc.), estén dadas de antemano y mejor por escrito, que el material que se haya de entregar a los alumnos esté bien organizado con objeto de repartirlo rápidamente, y que se eviten en lo posible los avisos de última hora.

b) Exámenes anunciados o por sorpresa: Es preferible, para que los estudiantes saquen de su estudio el máximo provecho, no sólo preparar concienzudamente el examen para los alumnos, sino también preparar

a los alumnos para el examen. Por ello, conviene que éstos sepan siempre con tiempo suficiente que van a tener un examen y en qué va a consistir.

c) Habilidades para hacer exámenes: No es infrecuente que los alumnos que no "saben" hacer exámenes rindan por debajo de lo esperado. Por ello, el profesor ha de procurar entrenarles y ayudarles.

d) Exámenes con o sin libros: Las pruebas en las que se pueden utilizar libros y apuntes se ven a veces como un modo de motivar al alumno para que no estudie superficialmente de memoria, sino para que asimile e integre los conocimientos; asimismo, obligan al profesor a construir preguntas que sirvan por lo menos para evaluar comprensión y aplicación. Por todo ello, este tipo de pruebas tiene su razón de ser. Además, los alumnos se dan pronto cuenta de que, si no van al examen bien preparados, los libros les sirven de poco: el buscar hechos y fórmulas puede distraerles y quitarles un tiempo necesario para responder adecuadamente las preguntas.

e) Copiar: El profesor tiene la obligación de evitar que los alumnos "copien". Este problema, que a veces sugiere versiones jocosas, tiene indudables componentes éticos y, por ello, debe considerarse con seriedad.

5. Corrección y Calificación: Una vez finalizado el examen, se procede a su corrección y calificación, según el formato de las pruebas y criterio de valoración.

6. Toma de Decisiones: Una vez obtenidos los resultados y valorados con una calificación, procede tomar las decisiones oportunas, individual o grupalmente, siguiendo las observaciones previamente comentadas.

7. Revisión: Es práctica habitual el ofrecer una última oportunidad al alumno para que revise las valoraciones y decisiones que el profesor hizo de sus ejercicios y, así, pueda aportar nuevos datos para alcanzar una mejor evaluación.

Para ello, es imprescindible planificar estas sesiones de revisión previamente con una información detallada sobre los errores detectados en las respuestas y sus posibles explicaciones; de lo contrario, se corre el peligro de desperdiciarlas con largas y estériles discusiones sobre los merecimientos de un punto arriba o abajo.

Hasta aquí el largo camino que va desde planificar el examen hasta que se informa al alumno de sus logros y de las decisiones tomadas. Aunque a primera vista puede parecer muy complicado, la verdad es que la mayoría de los profesores llegan fácilmente a automatizarlo. El peligro es más bien que el proceso de la evaluación se convierta en una rutina que no tenga en cuenta las precisiones que intentamos recoger en esta breve exposición. ©

Tomado de: <http://www.grupocerpa.com/gcpages/gcquevaluar.htm>