

Algunas técnicas de evaluación

Lorenzo García Aretio
Titular de la CUED
Editor del BENED

La evaluación del rendimiento académico de los estudiantes se lleva a cabo a través de las diversas pruebas y formas de evaluar ya comentadas en editorial anterior. Como ya señalamos allí, una prueba de evaluación es un reactivo que propone el profesor a los estudiantes con el fin de que éstos pongan de manifiesto la posesión de determinadas competencias, léanse conocimientos, habilidades, destrezas, etc. Y para medir esto pueden utilizarse distintos instrumentos. Trataremos de enumerar los tipos y características de algunos de los que más pueden interesar en la enseñanza a distancia.

Deseamos insistir en que los estudiantes pueden, deben, aprender del propio planteamiento de la evaluación. La evaluación integrada plenamente en el proceso de aprendizaje, se estructura como una variable sustancial que debe propiciar y mejorar este proceso, siempre que esté bien diseñada, bien aplicada y bien comunicada, a efectos de retroalimentación eficaz.

Vamos a limitar nuestra exposición a los que entendemos son instrumentos más usuales para la evaluación de logros de objetivos y competencias en la enseñanza a distancia. Por ello, por ejemplo, prescindiremos de la prueba oral, raramente utilizada en esta modalidad de enseñanza, salvo en algunas pruebas de idiomas que, en todo caso, pueden realizarse también utilizando los sistemas digitales y de telecomunicaciones.

Las pruebas abiertas o de respuesta libre

Las pruebas abiertas pueden recoger las denominadas de ensayo, las preguntas de respuesta libre, las composiciones, los trabajos, etc. En las pruebas de ensayo concretamente, se suele requerir a los estudiantes que muestren determinados productos de su aprendizaje, más complejos que la simple respuesta de un dato. En las composiciones escritas, pruebas de ensayo o de respuesta libre, pueden solicitarse al estudiante dos tipos de contestación:

- una respuesta a cada *item* o cuestión sin limitación de espacio, preguntas de *respuesta extensa*, o
- una respuesta limitada en la que se marca el espacio (media cuartilla, ocho líneas, una caja no extensible en el formulario digital en la Web, etc.), que serían cuestiones de *respuesta restringida* o breve.

Igualmente se pueden plantear cuestiones de estos tipos:

- Obligatorias o *impuestas* o, por el contrario dar un grado de *opcionalidad* al alumno de elegir "X" preguntas de entre las varias que se proponen.
- Pruebas en las que al estudiante se le va a dar o no posibilidad de *utilizar material* de apoyo para responder. Téngase en cuenta esta particularidad, dado que en las pruebas de evaluación a distancia será difícil conseguir que el alumno se sustraiga a consultar su material instructivo. Otra cuestión son las pruebas presenciales o, incluso, las realizadas telemáticamente con tiempo marcado para cada respuesta y para toda la prueba. No resulta en estos casos fácil consultar material, sencillamente por la falta de tiempo.

De igual modo debe considerarse el grado de repentización en la respuesta que se exige al estudiante. Así la respuesta puede ser:

- *Inmediata*, que se daría más frecuentemente en los ejercicios de autocomprobación y en las pruebas presenciales, y
- *Diferida*, más propia de las pruebas de evaluación a distancia o de los trabajos prácticos, de investigación o de amplia elaboración.

Tipos de preguntas en pruebas de respuesta libre

Podríamos enumerar una serie de posibles preguntas que ponen en juego diferentes procesos mentales que deberían estar relacionados con los logros previstos a nivel de competencias y objetivos. Veamos:

- Definir conceptos, ideas, principios, objetos...; justificarlos, proponer otros...
- Responder a quién, cuándo, cuánto, cuál, dónde, cómo, por qué...?
- Enumerar y/o describir hechos, objetos, fenómenos, personas...
- Contrastar semejanzas y diferencias, ventajas e inconvenientes, fortalezas y debilidades...
- Explicar relación causa-efecto, origen-consecuencia, inferencias o generalizaciones desde datos previo o por aplicación de normas o reglas.
- Argumentar a favor o en contra de determinados planteamientos.
- Analizar o sintetizar algún acontecimiento, procedimiento o argumento.
- Interpretar textos (comentario de textos), gráficos, tablas, audios, vídeos...
- Ordenar en cuanto a tiempo, espacio, valor...
- Evaluar propuestas, argumentos, procedimientos, obras...
- Etc.

Recomendaciones para su evaluación

Al elaborar este tipo de pruebas deberían considerarse, al menos, estas cuestiones:

- Los **objetivos** pretendidos y el tipo de proceso mental que se procura realice el estudiante que va a ser evaluado.
- La **extensión de la materia** que se pretende examinar, porque se hace difícil comprobar mediante este sistema el aprendizaje de excesiva cantidad de materia, dado el problema que supone el que la prueba sea **representativa** de todo lo que se quiere evaluar.
- El **tiempo** que se concede al estudiante para responder. Como al fin y al cabo en este tipo de pruebas son pocas las preguntas que se pueden hacer a los estudiantes, convendrá extremar su preparación con el fin de que midan el objetivo o competencia pretendidos, sean representativas, tengan una formulación sencilla y unívoca (todos han de entender lo mismo) y puedan ser previsiblemente contestadas por los concretos destinatarios de ellas que hayan dedicado un suficiente tiempo al estudio.

Si las preguntas que se hacen requieren una respuesta muy extensa (2 o 3 folios) podrán realizarse muy pocas en función del tiempo concedido. Pudiera parecer conveniente elaborar una prueba con más cuestiones de respuesta breve con el fin de que el muestreo de contenidos sea más amplio y su evaluación menos compleja. En todo caso, dependerá de la materia a evaluar y de las competencias (habilidades, destrezas, etc.) de las que se desea comprobar su logro.

Criterios que el evaluador puede considerar para otorgar la correspondiente calificación al estudiante, pueden ser entre otros posibles y según las características de la prueba, la apreciación de los siguientes:

- Respuesta ajustada a lo que realmente se ha preguntado y no a otras cuestiones cercanas o laterales al tema en cuestión.
- Dominio de la terminología y precisión conceptual.
- Justificación o argumentación sobre lo que se afirma o se niega.
- Capacidad para interpretar los datos facilitados y para organizar la información de forma sistemática.
- Riqueza y pertinencia de las ideas y profundidad en su tratamiento.
- Elaboración personal, crítica y fundamentada de lo aprendido.
- Originalidad, creatividad, imaginación...
- Calidad de la relación entre los conceptos: diferencias, semejanzas, causas, efectos, etc.
- Nivel de las conclusiones, pertinencia, consistencia y valoraciones críticas, si procede.
- Proyección de la problemática acometida a las situaciones reales.
- Estructura y argumentación.
- Citas y referencias bibliográficas.
- Corrección en la expresión gramatical y sintáctica.
- ...

El estudiante debería conocer la importancia relativa que el evaluador concederá a los diferentes criterios de valoración hechos públicos.

En el caso de tener que corregir, por ejemplo, un examen presencial de numerosos estudiantes, basado en este tipo de pruebas, sugerimos uno de estos métodos:

a) Método analítico.

- El evaluador elabora las respuestas de un examen modelo con las ideas o pasos fundamentales que considera debe responder el estudiante.
- Se asignan coeficientes de importancia a cada idea o paso.
- En función de ellos se van adjudicando puntuaciones a cada respuesta emitida por los alumnos.
- Se sugiere corregir la misma pregunta de todos, después otra y así sucesivamente.

Igualmente aconsejable es calificar, sin pausa, en una sola ocasión, todas las respuestas dadas a una determinada pregunta.

b) Método global. Según se leen las pruebas se agrupan en un número determinado de categorías (generalmente 5 ó 6). Por ejemplo, ejercicios de:

- Calidad superior
- Calidad por encima del promedio.
- Calidad media.
- Calidad por debajo de la media.
- Calidad inferior.

Quizás si son muchos trabajos pueda ser aconsejable, al principio, agruparlos en sólo tres categorías, buenos, normales y malos, que posteriormente podrían subdividirse. Puede resultar de interés adjudicar porcentajes aproximados a cada uno de los "montones", para así acercarnos a una distribución normal, por ejemplo, porcentualmente: 10-20-40-20-10, siendo el 40% el correspondiente al grupo de pruebas de calidad media. Aunque éste sería un procedimiento más ajustado a una evaluación normativa más que criterial. Nosotros apostamos más por esta última (ya se explicitó hace un par de editoriales en este mismo lugar).

Sea el que fuere el método utilizado, durante la corrección debe aprovecharse para realizar cuantos **comentarios** se estimen oportunos, con el fin de rectificar errores o fijar más los aprendizajes. Así siempre el estudiante conocerá por qué se le asignó una determinada calificación. Esta comunicación, por otra parte, se convierte en una insustituible fuente de motivación.

Normas a los estudiantes

Debe aprovecharse, bien sea la guía del curso o carrera, la guía didáctica de la materia o asignatura, el cuaderno de evaluación a distancia, las orientaciones en la plataforma digital o cualquier otro medio de comunicación de los utilizados en la enseñanza a distancia, para ofrecer a los estudiantes unas directrices que les orienten en la elaboración de sus respuestas a la prueba o pruebas. Podríamos señalar algunas:

- Leer rápidamente todas las preguntas y fijarse en la importancia y dificultad de cada una.
- Si se trata de pruebas con limitación de tiempo para contestar, distribuir éste de acuerdo con la prevista dificultad de las preguntas.
- Recapacitar sobre lo que realmente se pide en la pregunta con el fin de contestar estrictamente a aquello que se propone (comparación, análisis, síntesis, descripción, etc.).
- Un esquema de los puntos fundamentales de la respuesta en cuestión ayudará a organizar los contenidos y evitar una contestación desordenada o meramente memorística. Según la extensión requerida para responder, puede ser conveniente incluir este esquema en la cumplimentación de la prueba.
- Cuidar la utilización correcta y clara de la terminología propia.
- Evitar divagaciones y vaguedades que no responden a lo preguntado.
- Ser conscientes, en fin, de los criterios de evaluación que van a emplear los evaluadores y que deberían siempre ser conocidos con antelación por parte de los estudiantes.

Ventajas e inconvenientes de las pruebas de respuesta libre

Apuntamos una serie de ventajas e inconvenientes que pueden esgrimirse en pro o en contra de esta forma de evaluar. Hacen referencia bien a la redacción de las preguntas, bien a la respuesta por parte de los estudiantes o a la posterior calificación que se asigna.

VENTAJAS

- Son fáciles de preparar porque requieren menos preguntas que las pruebas objetivas.
- Soportan bajos costos. Incluso, en pruebas presenciales, se podrían dictar.
- Se puede medir cualquier tipo de objetivos del ámbito cognoscitivo y recoger aspectos fundamentales del estilo de aprendizaje del estudiante.
- Son insustituibles para comprobar cómo organiza el estudiante los conocimientos y cómo los expresa y dispone, dado que se le concede libertad para que exponga sus ideas con absoluta espontaneidad.

- Permiten apreciar la originalidad, creatividad e imaginación.
- Se puede apreciar la capacidad de los sujetos evaluados para emitir juicios y medir indirectamente actitudes, valores y opiniones.
- Potencian las capacidades de expresión y organización al requerir que el estudiante se exprese en forma gramaticalmente correcta, lógica y coherentemente.
- Pueden fomentar un estilo de estudio más profundo que requiera comparar, analizar, sintetizar, ordenar, etc.
- Es prácticamente imposible acertar por azar

INCONVENIENTES

- Soportan la dificultad de representar a la totalidad de los objetivos al tratarse necesariamente de una muestra corta de cuestiones. Se evalúa sobre una base muy limitada.
- Puede responderse a lo que se sabe más que a lo que se pregunta.
- Es cierto que no se puede adivinar la respuesta por azar, que sí puede suceder en las pruebas objetivas, pero puede ser fácil copiar.
- Existe desventaja para el alumno que tiene mala letra o es lento escribiendo o tecleando.
- Son difíciles de evaluar dado que frecuentemente sus preguntas admiten diversidad de respuestas e interpretaciones al estar formuladas, en algunos casos, de forma general y ambigua.
- Se suele calificar más la cantidad que la capacidad de organizar ideas.
- Detentan escasa validez y fiabilidad porque preguntas distintas es habitual que ofrezcan resultados distintos; correctores distintos suelen calificar de forma diferente; y una cuestión mal entendida por los alumnos, afecta mucho a la puntuación final.
- Con suma frecuencia es inevitable apreciar una fuerte influencia de la subjetividad, rigor, fatiga, prisas y estado de ánimo del evaluador.
- No lo debe corregir otro que no sea el profesor responsable del aprendizaje de los estudiantes.

Las pruebas objetivas y sus tipos

Las pruebas objetivas son un tipo de reactivo que propone el profesor a los alumnos con el fin de que éstos pongan de manifiesto o demuestren la posesión de determinadas competencias, con la particularidad de que la puntuación nunca depende de la opinión o juicio del profesor ya que, incluso, pueden ser corregidas mecánicamente

Algunas características generales de las pruebas objetivas son éstas:

- En ellas se utilizan gran número de *items*, preguntas o cuestiones.
- Las respuestas correctas no deben admitir discusión. Sólo una es la verdadera.
- La formulación de los *items* es breve.
- En lugar del recuerdo de la respuesta correcta, exigen del sujeto reconocerla.
- Generalmente pueden responderse rodeando por medio de un círculo el número o la letra de la respuesta correcta, subrayándola, marcándola en el mismo cuestionario donde se encuentran las preguntas o en una hoja de respuestas preparada al efecto. En las plataformas o entornos virtuales resulta fácil marcar la respuesta correcta.

Aunque **clasificaciones de pruebas objetivas** se han hecho numerosas y minuciosas muchas de ellas, nosotros vamos a caracterizar sólo aquellas que consideramos son de uso más común y factibles de utilizar en la enseñanza a distancia y lo haremos sólo de forma muy sintética:

- *Pruebas de respuesta libre y limitada.* Requieren una respuesta muy breve y exacta que en este caso ha de crearla totalmente el sujeto sin que existan sugerencias en la pregunta, contestando mediante una palabra, número o indicación. Ejemplo de ello son los textos mutilados.
- *Pruebas de selección de respuesta de doble alternativa.* Son una de las modalidades de las de selección de respuesta en las que la propia respuesta forma parte del *item*. El alumno debe identificar la opción correcta. Pruebas tipo verdadero-falso, correcto-incorrecto, sí-no...
- *Pruebas de selección de respuesta de alternativa múltiple.* Es el tipo de pruebas objetivas más utilizado y recomendable. La pregunta ofrece generalmente de tres a cinco posibilidades de respuesta. El alumno deberá encontrar entre ellas la respuesta correcta a la cuestión planteada.
- *Pruebas de asociación o respuesta por pares.* Se trata de establecer una determinada relación lógica por grupos de elementos establecidos en dos o más columnas de palabras, números, símbolos, frases, etc. En realidad son como una variante de *items* de múltiple elección de respuesta ya que de las varias preguntas que se presentan hay que elegir la respuesta correcta que también se ofrece en otra columna.
- *Pruebas de ordenamiento.* Precisan que el alumno coloque en un determinado orden una serie de elementos presentados sin orden. Resulta muy válida para evaluar objetivos que impliquen comprender la secuencia de una serie cronológica o lógica.

Dentro de las pruebas objetivas podemos considerar las denominadas *adaptativas y autoadaptadas*. Naturalmente este tipo de pruebas están pensadas más bien para ser utilizadas con ayuda del ordenador. En nuestro caso, a través de las plataformas virtuales. Las preguntas pueden presentarse al estudiante adaptadas a su nivel de conocimientos y habilidad. En las autoadaptadas el estudiante elige el nivel de dificultad de la prueba e, incluso, de cada pregunta.

No es nuestro propósito entrar a detallar los pormenores, características, ventajas e inconvenientes de cada una de estas pruebas objetivas.

Pruebas prácticas o de ejecución

En este tipo de pruebas se exige al estudiante que realice una determinada actividad con el fin de comprobar el nivel de logro de la competencia en cuestión. Este tipo de pruebas, incluso en la enseñanza a distancia, suelen requerir la presencia del estudiante en el laboratorio, taller o aula. No se descartan, sin embargo, las posibilidades de las tecnologías, por ejemplo a través del software de simulación. Tipos distintos de actividades que pueden ser motivo de pruebas prácticas podrían ser:

- Trabajos de laboratorio, utilización de aparatos, herramientas, instrumentos y materiales...
- Montaje y desmontaje de objetos, aparatos e instrumentos de carpintería, fontanería, electricidad, electrónica, mecánica, relojería...
- Elaboración de trabajos de pintura, dibujo, delineación, modelado, informática...
- Muestra de prácticas de cocina, peluquería, belleza y estética, corte y confección, mecanografía, programas informáticos, aeromodelismos, decoración y diseño, encuadernación, fotografía...

Para comprobar la calidad de los aprendizajes, de la destreza y habilidad adquiridas, al presentar un determinado trabajo para la evaluación, habrán de considerarse tanto los procesos y secuencia de movimientos necesarios para ejecutarlo como el producto final.

Para llegar a evaluar determinados aprendizajes de carácter práctico, además de que realicen un determinado trabajo completo, se puede pedir a los estudiantes que identifiquen o reconozcan la adecuación o no de un determinado procedimiento o utilización de herramienta o materiales, o bien que desarrollen una determinada actividad mediante simulación o con instrumentos y materiales a escala distinta de la real.

Otros tipos de instrumentos para la evaluación

Señalaremos otros instrumentos o tipos de pruebas que pueden utilizarse para medir el grado de logro de determinadas competencias, objetivos, contenidos, etc. Unos más y otros menos propicios para ser utilizados en la enseñanza a distancia.

Podremos considerar la **observación directa y sistemática** del comportamiento del estudiante a lo largo del periodo que se pretende evaluar. Para ello se pueden utilizar escalas, listas de control, etc. Las **listas**, por ejemplo, nos permiten registrar si un evento sucedió, sucede o que una determinada cualidad o característica del mismo se encuentra presente. Se usa para analizar las veces que sucede algo, si se recoge o no una determinada cualidad, etc. El estudiante, mediante este instrumento, puede también evaluar sus propias actividades en función de si se encuentran o no los criterios señalados en la lista para el desempeño de un determinado trabajo

Más posibilidades en los entornos virtuales

Otra forma bastante adecuada en los entornos virtuales de aprendizaje es el **análisis de la cantidad y calidad de la producción y participación** de los estudiantes en los diferentes ámbitos de interacción que se posibilitan. Si esta recogida de información y análisis se realizan de forma sistemática, se está propiciando bien un elemento adecuado para la evaluación continua.

Estas técnicas de observación y análisis se pueden y deben realizar a nivel individual pero también considerando los previsibles grupos de trabajo, equipos, trabajos colaborativos, etc. Veamos.

Un **portafolio** es un registro del aprendizaje, mediante la recopilación sistemática y organizada de evidencias con el fin de supervisar la evolución del conocimiento, las habilidades y las actitudes de los estudiantes. Se trataría de algo así como el antiguo cuaderno de clase. En el portafolio (*portfolio*) se recogen productos del estudiante de muy diferente índole. Aquí pueden incluirse lo que otros denominan *memorias de reflexión de los estudiantes* sobre su propio proceso de aprendizaje. Los portafolios digitales abren inmensas posibilidades para la enseñanza y para la evaluación.

Aunque son de pretensiones diferentes, algunos han sustituido los portafolios por los **blogs educativos o edublogs**. Estos blogs o bitácoras, como ya escribimos en otro editorial del BENED (septiembre de 2005), suponen o pueden suponer otro elemento de gran relieve, tanto como herramienta educativa que como instrumento para la evaluación, que ahora nos ocupa. Con respecto a la evaluación, señalábamos en el citado editorial que mediante el uso de bitácoras, se puede, por ejemplo, comprobar:

- El nivel de análisis y de crítica en la defensa de las propias posiciones y la consistencia de cuanto se afirma. Es evidente que el desarrollo del pensamiento crítico es uno de los grandes valores de esta herramienta.
- De qué manera cada alumno organiza las ideas que expone.
- La calidad de los vínculos o enlaces reseñados, ofrecidos como fuentes de primera mano y no reinterpretada.
- La calidad de los comentarios o críticas realizados a los trabajos propuestos por el profesor o por otros compañeros.
- La frecuencia, oportunidad y calidad de las diferentes intervenciones.
- La disposición mostrada por cada estudiante para integrarse en grupos y realizar trabajos colaborativos.
- La propia valoración de los trabajos colaborativos realizados y expuestos por los diferentes grupos de estudiantes.

También, en otro editorial en el que aludíamos a las **Wiki** (abril, 2006), señalábamos que una *wiki* puede ser buen instrumento para la evaluación de los estudiantes, tanto a nivel individual como de grupo. Un profesor puede valorar los progresos que un estudiante o grupo pueden hacer con respecto a la tarea encomendada que progresivamente se va plasmando en la *wiki*.

En los formatos de enseñanza y aprendizaje digitales, evidentemente, se amplían las posibilidades. Además de todo lo señalado, ya sabemos que en línea resultan más sencillas, por su automatismo en la evaluación, determinadas pruebas de autoevaluación. También se facilita la evaluación por pares, evaluación en equipos y colaborativa, evaluación de debates en línea, simulaciones, solución de problemas, aprendizaje basado en proyectos, etc. En estos últimos casos los estudiantes han de dar respuesta a problemas o proyectos que tienen la particularidad de exigir el dominio de muchas ideas, conceptos, relaciones, etc., probablemente estudiados en unidades de aprendizaje muy diferentes y que han de presentarse de manera coherente, ordenada y precisa.

Las discusiones y debates en línea, más por la calidad de las intervenciones de los estudiantes que por su cantidad, deberían considerarse en el cómputo final de la calificación. Estas intervenciones deberían, o podrían, ser calificadas también por parte de los pares.

Existe software apropiado para facilitar procesos de evaluación mediante sistemas digitales, en línea, que permiten generar diferentes modalidades de pruebas, administrarlas y corregirlas automáticamente. Los bancos de preguntas de diferente índole son un elemento importante para manejar este tipo de tecnologías. Estos modelos tecnológicos facilitan incluso proponer exámenes diferenciados por estudiante (esto tiene sus ventajas pero también sus inconvenientes). Los resultados de la prueba, tanto a nivel individual como de grupo, pueden facilitar información de

gran valor y, además, puede ser inmediata, según los casos, tanto para el docente como para el estudiante.

Preparación de una prueba ordinaria

Si nos referimos a los tipos más convencionales de pruebas, tanto en los ejercicios de autoevaluación de las Unidades Didácticas como -fundamentalmente- en las pruebas de evaluación a distancia y -en su caso- presenciales, deberá cuidarse hasta el extremo la elaboración de la prueba.

Ya contamos con que una prueba ha de tener:

- **Fiabilidad**, es decir, que mida con un grado correcto de precisión.
- **Validez**, es decir, que mida aquello que se pretende medir y no otra cosa.

Pero también es necesario que la prueba sea **representativa** de los objetivos y contenidos considerados en los materiales de estudio. Porque suele ser habitual en la construcción de pruebas el incluir excesivo número de preguntas que exigen sólo evocación de contenidos informativos, e incluso, excederse en determinadas cuestiones que sólo afectan a una determinada área o parte del total de los contenidos que se pretenden evaluar.

Para evitar estos problemas reales que disminuyen notablemente la calidad de la evaluación, resulta aconsejable elaborar previamente una **tabla de especificación** que nos garantice que se van a ponderar todos los objetivos y contenidos de forma adecuada.

En definitiva, deberemos considerar:

- Las competencias y objetivos pretendidos en la parte del curso que es motivo de evaluación y la importancia relativa de los mismos.
- Las áreas de contenidos contempladas y la importancia relativa de los mismos.
- La selección del tipo de prueba que conviene para evaluar esos contenidos y objetivos y el número total de *items* de la misma.
- El número de cuestiones que corresponden a cada objetivo o área, que debe ser representativo.

En fin, creemos haber abordado, aunque no en profundidad, parte de la problemática de la evaluación en los contextos no presenciales. Téngase en cuenta de que si se tienen en consideración buena parte de las ideas y sugerencias vertidas en estos tres últimos editoriales del BENED, podríamos sentirnos satisfechos de la evaluación que realizamos de los estudiantes a nuestro cargo.

© *Lorenzo García Aretio* – Editor del BENED y Titular de la CUED

Otros Editoriales del BENED: <http://www.uned.es/cued/boletin.html>