

Cómo citar la publicación:

García Aretio, L. (15/05/2020). Las pruebas abiertas o de respuesta libre. *Contextos universitarios mediados*. (ISSN: 2340-552X), <https://aretio.hypotheses.org/4283>.

## Las pruebas abiertas o de respuesta libre

Lorenzo García Aretio  
UNED

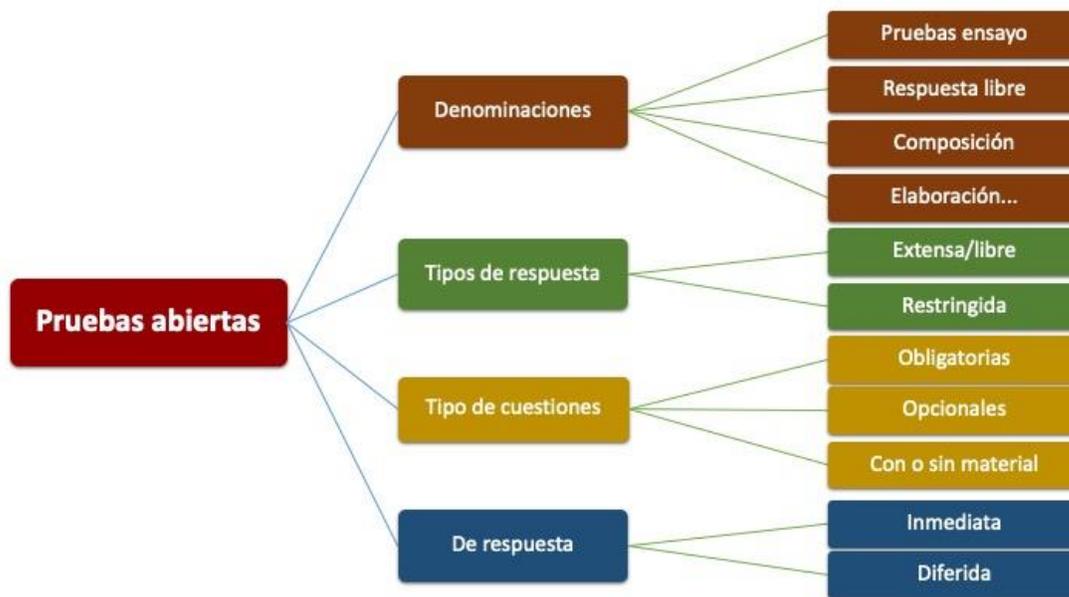
En [entrada anterior](#) ya me referí a los diferentes instrumentos y técnicas de evaluación de los que nos podemos servir a la hora de comprobar, medir, valorar, hasta qué punto, o a qué nivel, se alcanzaron los diferentes objetivos de logro académico. Las competencias y los objetivos pretendidos, estamos obligados a dar fe de que se superaron o no. Definirlos, aunque se haga a la perfección, no supone logro. Nos guste o no, estamos obligados a conocer sus niveles de consecución. En esa entrada anterior aludida, ya me referí a algunos tipos de instrumentos o técnicas de evaluación: *orales, comentarios de texto, observación (escalas, listas de control), portafolios, blogs, wikis, prácticas*. En esta entrada y en otras tres siguientes, trataré de abordar los otros instrumentos.

Ahora, trataré de *actualizar* lo publicado sobre el tema hace más de 25 años ([García Aretio, 1994](#)). En el concepto de **pruebas abiertas** podrían contemplarse las denominadas de **ensayo**, las preguntas de *respuesta libre*, las *composiciones*, los trabajos de *elaboración*, etc. En esta técnica evaluadora se suele requerir a los estudiantes que muestren determinados resultados de su aprendizaje más complejos o elaborados que la simple respuesta de un dato. Pueden solicitarse tres tipos de contestación:

- una respuesta a cada *ítem* o cuestión sin limitación de espacio, preguntas de *respuesta extensa y libre*, o
- una *respuesta acotada* 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 podrían plantear cuestiones de estos tipos:

- *Obligatorias* o *impuestas*. Ha de intentar responderse a todas.
- Dar un grado de *opcionalidad* al estudiante para elegir «X» preguntas de entre las varias propuestas.
- Pruebas en las que al estudiante se le va a dar o no posibilidad de *utilizar material* de apoyo para responder.



En cuanto esta última tipología, téngase en cuenta que en las pruebas de evaluación a distancia será difícil conseguir que el estudiante se sustraiga de consultar sus materiales de estudio. Otra cuestión son las pruebas presenciales o, incluso, las realizadas en línea, 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. Esa variable tiempo ha de cuidarse en extremo en este tipo de pruebas en línea. Resulta una buena forma de dificultar la consulta de materiales, si es que no se permite su uso, así como la intercomunicación con otros sujetos.

De igual modo debe considerarse el grado de repentinización en la respuesta que se exige al estudiante. Así la contestación 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.

En todo caso, deberá cuidarse la *redacción* de las preguntas, de manera que sean meridianamente claras, coherentes con los objetivos y competencias, de diferente grado de dificultad, etc.

### ***Tipos de preguntas en pruebas de respuesta libre***

Aún, en esta época de evaluación de aprendizajes por competencias, se observan exámenes o pruebas, por ejemplo, compuestos por cuatro preguntas extraídas, casi literalmente, de los epígrafes o apartados de las unidades didácticas o de estudio recomendadas. En este caso, además de facilitar mucho las posibilidades de plagio o copia, estaríamos evaluando exclusivamente la capacidad de *recuerdo* o la *memoria fotográfica* del estudiante, ¿dónde dejamos todos los demás objetivos y competencias?

No es eso. A título de ejemplo, me gustaría sugerir 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 estos ejemplos:

- *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...
- *Comparar* y *Contrastar* semejanzas y diferencias, ventajas e inconvenientes, fortalezas y debilidades...
- *Explicar* relación causa-efecto, origen-consecuencia, inferencias o generalizaciones desde datos previos.
- *Aplicar* normas, principios o reglas generales.
- *Determinar finalidad* o *propósito* de un autor, personaje, técnico u operario al tomar decisiones u organizar una actividad.
- *Argumentar* a favor o en contra de determinados planteamientos.
- *Discutir simbólicamente*. Argumentos en pro o en contra, relaciones de esos argumentos con otras ideas o conclusiones.
- *Analizar* o *sintetizar* algún acontecimiento, procedimiento o argumento.
- *Clasificar*, *categorizar*. A partir de los conceptos e ideas sugeridos, buscar categorías, clasificar los hallazgos...
- *Interpretar textos* (comentario de textos), gráficos, tablas, audios, vídeos...
- *Interpretar tablas, gráficos, imágenes*.
- *Ilustrar* y *ejemplificar*. Desde ideas o conceptos previos planteados, indicar ejemplos originales.
- *Ordenar* en cuanto a tiempo, tamaño, espacio, valor...
- *Criticar* y *Evaluar* propuestas, argumentos, procedimientos, obras. Adecuación, corrección, o importancia de una situación, con razones para mejorar o argumentos para abandonar.
- *Elaborar* respuestas imaginativas y creativas ante un problema o situación...

### **Recomendaciones para su evaluación**

Una vez seleccionados esos procesos cognitivos que deseamos evaluar, no vendría mal distribuir el número de preguntas entre ellos y entre los bloques o unidades motivo de evaluación, como les sugerí con la tabla de especificación. Así, al elaborar este tipo de pruebas deberían considerarse, al menos, estas cuestiones:

- Los **objetivos** y competencias pretendidos y el tipo de proceso mental que se procura realice el estudiante que va a ser evaluado;
- la extensión de la materia, los **contenidos**, 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. En este tipo de pruebas son pocas las preguntas que se pueden proponer a los estudiantes. En el caso de pruebas en línea, la variable tiempo resulta *esencial* con el fin de dificultar la consulta de materiales, si es que no se permiten, así como la intercomunicación.
- convendrá extremar su preparación con el fin de que 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.

- Definir con claridad los **criterios** de evaluación que siempre deben ser conocidos por los estudiantes con suficiente antelación.

Si las preguntas que se hacen requieren una respuesta muy extensa (2 o 3 folios, por ejemplo) podrán proponerse muy pocas en función del tiempo concedido. Pudiera parecer conveniente elaborar una prueba con más cuestiones de respuesta más breve con el fin de que el muestreo de objetivos y contenidos sea más amplio y su evaluación menos compleja. En todo caso, dependerá de la materia a evaluar, de las competencias (habilidades, destrezas, etc.) de las que se desea comprobar su logro y de la redacción de la prueba. De ahí que cabría, que una sola cuestión bien planteada pudiera requerir poner en juego numerosos objetivos y procesos mentales y, a la vez, ser integradora de varias unidades de estudio

Si se hubiese elaborado una rúbrica o matriz de valoración, se habrían podido seleccionar criterios, que el evaluador puede considerar para otorgar la correspondiente calificación al estudiante. Podrían apreciarse algunos de entre éstos (*en los primeros propongo ejemplo de escala de valoración*):

- *Respuesta ajustada a lo que realmente se ha preguntado y no a otras cuestiones cercanas o laterales al tema en cuestión (mínimamente ajustada – poco ajustada – ajustada – bastante ajustada – plenamente ajustada).*
- *Dominio de la terminología y precisión conceptual (nulo --- --- máximo).*
- *Justificación o argumentación sobre lo que se afirma o se niega (1 – 2 – 3 – 4 – 5).*
- *Capacidad para interpretar los datos facilitados y para organizar la información de forma sistemática (mínima – poca – aceptable – bastante – máxima).*
- *Riqueza y profundidad en su tratamiento (etc.).*
- *Pertinencia de las ideas expuestas.*
- *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.*
- ...

En consecuencia, el estudiante debería conocer la importancia relativa que el evaluador concederá a los diferentes criterios de valoración hechos públicos. Y por ello resulta bien apropiado el uso de rúbricas.

### **Otras formas de valoración de la prueba.**

Vayamos al caso de tener que corregir, por ejemplo, un examen presencial de **excesivo número** de estudiantes, basado en este tipo de pruebas. Las rúbricas pueden resultar harto complejas para aplicarlas a tantos estudiantes. Sugerimos uno de estos métodos:

#### **a) Método analítico.**

- El evaluador elabora las *respuestas tipo* 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, según relevancia (previamente esta ponderación debería ser conocida por los estudiantes hasta donde sea posible y recomendable).
- En función de ellos se van adjudicando *puntuaciones* a cada respuesta emitida por los estudiantes.
- Para ser más objetivo, 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. Previsiblemente cambiará menos el estado de ánimo del evaluador, salvo que sean demasiados estudiantes, y a los últimos en ser evaluados..., ¿cómo estará el ánimo?

En el caso de asignaturas o cursos impartidos por equipos docentes, podría arbitrarse el que cada miembro del equipo corrigiese una misma pregunta de todos los estudiantes. Así los criterios y valoraciones podrían fluctuar menos.

#### **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 estudiantes pueda ser aconsejable, al principio, agrupar los trabajos en sólo tres categorías, *buenos, normales y flojos*, 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, en el caso de cinco categorías, porcentualmente, la típica campana: 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 *criteria*. Ya aposté claramente por esta última (explicitado en [entrada anterior](#)).

### **c) Método ponderado**

En esa apuesta por lo  *criterial* por encima de lo  *normativo*, en algún tipo de exámenes, convendría señalar las ponderaciones y valoraciones que se asignan en la prueba respecto a las diferentes preguntas e, incluso, criterios.

- ¿Qué valoración tendrá en la prueba una pregunta meramente teórica?,
- ¿Será su valor en el examen idéntico al de una pregunta de aplicación, o práctica?

En consecuencia, se deberán establecer públicamente esas valoraciones:

- ¿todas las cuestiones tienen el mismo peso en la calificación, o diferente?,
- ¿se puede superar un examen al obtener 5 puntos (sobre 10) con sólo las respuestas a preguntas teóricas?, ¿y al revés?

Podrían establecerse puntuaciones mínimas o de corte con respecto a cada parte del examen, teórico o práctico, o respecto a alguna pregunta esencial en la que no se admitirá, por ejemplo, obtener menos de 5 puntos, aunque en las otras cuestiones se obtuviese una alta puntuación.

Sea el que fuere el método utilizado, durante la corrección debe aprovecharse para realizar cuantos **comentarios** se estimen oportunos, con el fin de *realimentar*, 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 y autorregulación. Ya me referiré a este importante tema de la realimentación o *feedback*.

### **Normas a los estudiantes**

Debe aprovecharse, bien sea la guía del curso o carrera, la guía didáctica de la materia o asignatura, la prueba 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 destacar 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 previsible 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 del área de saber.
- 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.

- Resulta esencial considerar el tiempo total concedido para la resolución de la prueba que, en todo caso, debe ser lo más ajustado posible, con el fin de dificultar al máximo acciones fraudulentas.

### ***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 diversos tipos de objetivos y habilidades de orden superior 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.
- 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 y contenidos, al tratarse necesariamente de una muestra corta de cuestiones. Se evalúa sobre una base muy limitada.
- Exigen mucho tiempo para la corrección.
- 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, sobre todo si la prueba exige mera memorización.
- 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. De ahí la necesidad de criterios claros, *crúbricas*?
- Se puede tender a calificar más la cantidad que la capacidad de organizar ideas.

- Detentan escasa validez y fiabilidad porque:
  - preguntas idénticas es habitual que puedan ofrecer respuestas distintas pero aceptables;
  - correctores distintos suelen calificar de forma diferente;
  - una cuestión mal entendida por los alumnos (quizás por redacción deficiente), puede afectar en 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 las debería corregir otra persona que no sea el profesor o equipo responsable del aprendizaje de los estudiantes.
- ...

**Referencia:** García Aretio, L. (1994). *Educación a distancia hoy*. UNED.