

PRUEBAS O EXÁMENES ESCRITOS

El examen escrito es un instrumento clásico de evaluación de los aprendizajes. No obstante, existen numerosas variantes por lo que respecta a la forma del mismo, los tipos de respuestas exigidas y las modalidades de corrección.

El examen escrito constituye la forma de examen más conocida y más extendida. En esta prueba cronometrada, efectuada bajo control, el profesor pide a los alumnos que respondan por escrito y con sus propias palabras a una o varias cuestiones relativas al programa de la asignatura, independientemente de si se les concede o no el derecho a poder consultar distinta documentación o material de apoyo.

Antes del examen final los profesores proponen, a menudo, uno o dos exámenes a lo largo del curso, lo que puede permitir a los estudiantes evaluar mejor su progresión en el aprendizaje. La duración de estos controles es, en general, inferior a la del examen final. Además, estos mini exámenes no abarcan más que una parte de los contenidos cuando el examen final impone generalmente a los estudiantes una revisión global de todos los contenidos.

En un examen escrito, el profesor puede recurrir a dos categorías de cuestiones: cuestiones para ser corregidas subjetivamente y cuestiones de respuesta y corrección objetiva.

Examen para corrección subjetiva. En este tipo de examen, el profesor debe, para corregir las respuestas de los alumnos, juzgar estas respuestas para determinar si satisfacen los criterios preestablecidos. El profesor puede poner cuestiones cuyas respuestas sean largos desarrollos –numerosas líneas e incluso numerosas páginas-, pero puede, igualmente, exigir respuestas breves –cinco líneas o menos-

Tipos de preguntas y de exámenes abiertos

Es útil tener a la vista los diversos tipos posibles de preguntas o exámenes abiertos. El término pruebas abiertas es muy genérico y tanto las ventajas como las limitaciones de estas pruebas dependen de; tipo de prueba o examen abierto. No conviene olvidar que los diversos tipos de exámenes no son mutuamente excluyentes.

Los tipos presentados aquí (sugeridos por Beard y otros, 1974, que recogen prácticas habituales en algunas universidades) son indicativos; se puede pensar otras modalidades.

1. Preguntas que requieren una respuesta larga. Las ventajas (el alumno se expresa y aprende a expresarse, a organizar ideas, etc.) y las limitaciones (sobre todo relacionadas con la corrección) que suelen aducirse sobre estas pruebas se refieren normalmente a estas respuestas largas.

2. Preguntas de respuesta corta. Comparadas con las preguntas de respuesta larga, estas preguntas participan menos de sus ventajas (apenas hay oportunidad de organizar nada) y de sus limitaciones (se corrigen con menor dificultad). Si las respuestas son muy breves, participan más de las características de las preguntas objetivas.

Una modalidad de estas preguntas consiste en hacer una serie de preguntas de respuesta breve en torno a una información o problema común.

En este tipo de exámenes (preguntas abiertas de respuesta corta) la fiabilidad es mayor (entendida como correlación entre evaluadores), y pueden tener una mayor fiabilidad que las preguntas objetivas que cubren el mismo contenido (estudios citados por Heywood, 1977).

3. Exámenes largos de libro abierto. Exámenes largos (tres horas) con posibilidad de utilizar libros, textos, diccionarios; una crítica que se puede hacer a este tipo de exámenes es que comprueban la habilidad de trabajar contra reloj (Beard y otros, 1974).

4. Exámenes hechos en casa. Trabajo largo, de una longitud predeterminada, preparado en casa (por ejemplo durante dos semanas). El alumno tiene que localizar información, organizarla, etc. Tiene además la oportunidad de descubrir cosas para él desconocidas, se ve sometido a las mismas presiones que un autor que debe publicar algo, etc.

5. Exámenes con tema anunciado. los alumnos llegan preparados. En un departamento de bioquímica uno de los tipos de exámenes en uso consiste en anunciar a los alumnos de qué se van a examinar seis semanas antes.

Lo que se pretende en estos casos es que los alumnos aprendan a buscar información por sí mismos, que consulten con compañeros, etc. En cuanto a los resultados de este tipo de exámenes, algunas investigaciones muestran que los alumnos quedan mucho mejor, pero el orden resultante es el mismo que con otros tipos de exámenes (los mejores son los mismos) (Beard y otros, 1974).

6. Trabajos, proyectos, tesinas, papers, etc. para hacer en casa. Estos

trabajos pueden hacerse individualmente o en pequeños equipos; no responden a la situación de examen clásico. Son muy útiles porque con ellos se consiguen objetivos importantes (búsqueda y selección de información, lectura, organización, etc.) Una topología útil de este tipo de trabajos puede encontrarse en Meyers (1987).

Esta tipología tiene más que ver con tipos de exámenes abiertos que con tipos concretos de preguntas. Modelos de preguntas pueden encontrarse en diversas fuentes; aunque se trate de preguntas pensadas para una asignatura determinada, pueden adaptarse a otras materias muy diferentes o al menos pueden sugerir otros modos de formular preguntas que estimulan un tipo de estudio interesante (pueden verse modelos en Madigan y Brosamer, 1990; Berrenberg, 1990; Poe, 1990).

Ventajas e inconvenientes

Ventajas y limitaciones de las preguntas abiertas

VENTAJAS	INCONVENIENTES
<p>1. Es fácil de preparar:</p> <p>a) requiere menos preguntas que las pruebas objetivas</p> <p>b) es fácil redactar buenas preguntas;</p>	<p>1. La corrección es laboriosa; es difícil; corregir con objetividad; la calidad de estas pruebas depende en buena parte de la calidad de la corrección. En la corrección influyen:</p>
<p>2. Es barato; las preguntas se pueden dictar, escribir en la pizarra; si se fotocopian ocupan menos espacio que las pruebas objetivas;</p>	<p>a) el contexto, la evaluación se ve influida por los ejercicios precedentes...</p> <p>b) el conocimiento previo del alumno...</p> <p>c) rasgos particulares de cada ejercicio (letra, presentación, etc.)</p> <p>d) fatiga, tiempo disponible, etc.</p>
<p>3. La formulación de la pregunta puede adaptarse para comprobar casi cualquier tipo de objetivo;</p>	
<p>4. Admiten varias modalidades (una pregunta de respuesta amplia, varias preguntas de respuesta breve en torno a un mismo tema o información, etc., exámenes de libro abierto, preparados de antemano, etc.</p>	<p>2. Tienen menos fiabilidad que las pruebas objetivas (en general):</p>
<p>5. Es el único método válido para comprobar objetivos importantes como capacidad de expresión, organización, creatividad, etc., también son el método válido para comprobar capacidad de aplicación, análisis, evaluación, etc. (al menos en niveles complejos).</p>	<p>a) admiten menos preguntas y se evalúa sobre una base más limitada;</p> <p>b) preguntas distintas hubieran dado resultados distintos;</p> <p>c) correctores distintos o un mismo corrector en momentos distintos pueden dar resultados distintos;</p> <p>d) una pregunta mal entendida afecta de manera desproporcionada a la calificación;</p>
<p>6. No se puede adivinar la respuesta sin conocerla-</p>	<p>e) el alumno puede responder a lo que sabe, no a lo que se le pregunta...</p>
<p>7. Potencian la capacidad de expresión, organización, originalidad, etc. aunque no se trate de objetivos en sentido propio (aunque esto dependerá de las normas del profesor, tipo de corrección y feedback etc.; también se pueden crear malos hábitos).</p>	<p>3. No cabe el adivinar, pero puede influir la suerte, se puede copiar, etc.</p>

VENTAJAS	INCONVENIENTES
<p>8. Pueden condicionar un buen método de estudio (visiones amplias, completas y bien estructuradas de los temas, relaciones interdisciplinarias, etc.) aunque también esto dependerá en parte del profesor (también pueden condicionar hábitos memorísticos de repetición de textos y apuntes).</p> <p>Cautelas:</p> <ol style="list-style-type: none"> 1. Seleccionar muy bien las preguntas, 2. Redactar bien la pregunta (que oriente). 3. Preparar clave de corrección 	<ol style="list-style-type: none"> 4. Frecuentemente se corrigen en función de la cantidad de conocimientos y se tiene menos en cuenta la capacidad de expresar y organizar ideas, la presentación, etc. también pueden condicionar un estudio memorístico, poco profundo; además para evaluar la capacidad de organización harían falta más organizaciones... 5. Las preguntas tienen el peligro de ser muy generales o ambiguas, con preguntas muy concretas y de respuesta muy breve se obvian muchas dificultades, pero se le quita al alumno la posibilidad de expresarse, organizar la respuesta, mostrar visiones de conjunto, etc., 6. El que cree buenos hábitos de estudio no significa que sea el mejor método de evaluación de muchos objetivos, que se pueden comprobar mejor de otras maneras.

2

Orientaciones para la redacción de preguntas abiertas

- 1) Reservar las preguntas abiertas para comprobar resultados que no se puedan comprobar mejor mediante pruebas objetivas.
- 2) Formular con mucho cuidado estas preguntas para que centren la atención del alumno en la tarea deseada.

Utilizar uno o varios verbos: describe y justifica; explica por qué; qué sucedería si..... y por qué; como resolverías esta situación.....según.....; haz una síntesis de.....; explica las ventajas y limitaciones de.....; según estos criterios evalúa si.....; de estas dos proposiciones, con cuál estaría de acuerdo..... y por qué.

Se debe mencionar siempre la necesidad de una buena organización de las ideas y de una presentación clara. La tarea que se le pide al alumno debe estar bien definida.

Las preguntas muy genéricas o ambiguas pueden despistar a los alumnos que saben, pero no adivinan el enfoque que quiere el profesor.

- 3) Una buena práctica consiste en introducir en las preguntas, al menos en algunas, información nueva (texto, citas, datos); estas preguntas requieren algo más que reproducir el material del texto o apuntes y comprueban con más facilidad algo más que la pura memoria.
- 4) Puede ser preferible, como criterio general, preguntas específicas de respuesta relativamente breve, así se cubre un espectro más amplio de conocimientos y habilidades.

- 5) Es preferible evitar el uso de preguntas opcionales; los alumnos lo prefieren, pero no se evalúa en todo con la misma base, dejan algunos temas sin estudiar; puede favorecer hábitos no deseables de estudio.
- 6) Antes de corregir, se debe pensar en el método de corrección-calificación (global, analítico)
- 7) Hay que decidir de antemano que se va a hacer con deficiencias y errores aunque sean irrelevantes en relación a lo que se está comprobando (saber o no saber) , como son las faltas de ortografía, la presentación....
- 8) Una práctica recomendable es corregir sin saber el autor y pregunta por pregunta.
- 9) El profesor debe ser consciente y prevenir que un alumno no responda inteligentemente, utilizando estrategias como escribir por lo menos algo, aunque sólo sea reformular la pregunta y decir que es muy importante; o bien, insistir y extenderse en la importancia del tema; o bien mostrar acuerdo con los puntos de vista del profesor siempre que parezca apropiado, etc. Con preguntas bien redactadas se evitan mejor este tipo de respuestas que perjudican al tímido, honrado, poco imaginativo...
- 10) Al menos en los exámenes importantes es preferible disponer de más de un corrector.

Guía para evaluar preguntas

1. ¿Se utilizan estas preguntas exclusivamente para comprobar objetivos de cierta complejidad? (capacidad de organización, de pensar por cuenta propia, etc.) (la suposición implícita de los autores es que para comprobar otros tipos de objetivos son preferibles las pruebas objetivas).
2. ¿Están las preguntas redactadas de manera que comprueben los objetivos pretendidos? (validez).
3. ¿Presenta cada pregunta una tarea clara al alumno? ¿Comienzo las preguntas con expresiones como compara, contrasta, da la razón de por qué..., aporta ejemplos originales de..., explica cómo..., predice qué pasaría si... ? (La pregunta debe orientar sobre el tipo de respuesta requerido, sin limitarse a enunciar un tema). ¿Conocen los alumnos qué puntos deben desarrollar en su respuesta? (el alumno tiene obligación de responder, no de adivinar la verdadera pregunta).
4. ¿Hay tiempo suficiente para responder a todas las preguntas? ¿Son los alumnos conscientes de las limitaciones de tiempo?
5. ¿He utilizado material nuevo e interesante en este tipo de preguntas (citas, casos, ejemplos, etc.)?
6. ¿Tengo escritas una serie de orientaciones para los alumnos sobre cómo responder a este tipo de preguntas?
7. Si son varias las preguntas ¿difieren en dificultad y complejidad?
8. Antes de empezar a corregir, ¿tengo claro el esquema de la re-

- spuesta correcta?
9. Si los alumnos pueden escoger entre varias preguntas, ¿tengo claro que no las voy a corregir y evaluar comparando a unos alumnos con otros?

Pedro Morales (1995). Tipos de pruebas- los exámenes orales y las preguntas de respuesta abierta. Bilbao: ICE de la Universidad de Deusto.

Corrección de pruebas abiertas

Factores que influyen frecuentemente en la corrección e exámenes abiertos

- 1º *La letra, presentación, ortografía...*
- 2º *El contexto,- lo que vamos viendo nos influye en lo que corregimos después; no es lo mismo corregir un examen normal después de varios exámenes muy buenos o muy malos...*
- 3º *Lo que ya sabemos del alumno, nuestras expectativas. cómo va en general ... tendemos a ver lo que esperamos ver...*
- 4º *Cansancio, tiempo disponible...*
- 5º *El propio estilo de corregir.- tanto por lo que respecta al método (una visión de conjunto para hacernos una idea global, o un enfoque más analítico, distinguiendo diversos aspectos ...) como por lo que respecta al talante personal (tendencia a la benignidad o a cierta dureza...*

3

Enfoques para corregir pruebas abiertas

ENFOQUE GLOBAL (holístico)

Lo característico es seguir la siguiente norma:

- a) Se clasifica según una apreciación de conjunto más o menos global.
- b) Se califica o se valora.

Se pueden distinguir tres tipos o modalidades:

- a) Criterios referidos al grupo:
 - Clasificación en tres grupos: medio, por encima de la media y por debajo de la media.
 - Los dos montones extremos se pueden ir subdividiendo (los que están muy bien, los que están muy mal) quedando 5 subcategorías.
 - Los casos dudosos se dejan aparte hasta el final.
 - Se dejan aparte los que por alguna razón se desea examinar con más cuidado (suspensos, dudosos, etc.)

Este enfoque es posiblemente más adecuado para corregir trabajos de cierta amplitud, cuando hay poco tiempo o es difícil pensar en otra alternativa, se presta menos para corregir *exámenes convencionales*, con varias preguntas relativamente breves. Estos sistemas de corrección-clasificación basados en una clasificación previa son más útiles:

- a) Cuando se trata de corregir redacciones o trabajos en las que el alumno tiene una mayor libertad para escoger su propio enfoque.
- b) Cuando hay que corregir un gran número de trabajos. De hecho estos métodos son muy utilizados cuando el número de alumnos es muy grande (como sucede con frecuencia en trabajos de redacción hechos como parte de examen de ingreso en una universidad).

b) Criterios según juicios predeterminados

En vez de los tantos por ciento anteriores, se pueden especificar previamente juicios de valoración como criterios de inclusión en un grupo u otro, como *excelente, bastante bien, suficiente, deficiente y muy deficiente*. El proceso es más claro si hay algunos *criterios definidos* para la inclusión de los ejercicios en un grupo u otro. Esta modalidad se presta más que a la anterior para corregir preguntas abiertas en exámenes convencionales.

Se pueden también especificar una serie de juicios pensados de antemano y aplicables a las probables respuestas de los alumnos, como las que figuran en el cuadro X (*adaptado* de Heywood, 1989).

Los juicios de cuadro X y su *gradación* de mejor a peor pueden ser discutibles, después de todo responden a una situación específica, pero dan una idea del procedimiento. Posiblemente es más útil disponer de menos juicios previos.

Estos esquemas pueden admitir múltiples formas, unas más sencillas y otras más complicadas. Posiblemente pueden bastar cuatro o cinco descripciones de posibles respuestas. Taylor (1994) presenta un ejemplo similar para corregir en este caso una redacción o ensayo hecho a partir de la lectura de una obra literaria. En la clave se especifican en primer lugar los *criterios* que se van a tener en cuenta (se identifica el carácter del protagonista, se citan pasajes que justifican el asignarle ese carácter o personalidad, se describe el influjo del carácter del protagonista en el desarrollo de la historia, etc.). Una vez identificados los criterios se describen esquemáticamente cinco posibles modelos de respuesta o de ensayo.

Cuadro X

Juicios predeterminados para evaluar respuestas abiertas

10. Respuesta extraordinaria, hay originalidad, muestra pensamiento independiente, etc.	6. Muestra comprensión pero poca información.
9. Óptima organización, aporta en la respuesta toda la información que se puede esperar.	5. Hay cosas relevantes, pero no muestra una comprensión adecuada.
8. Buena organización de lo más	4. No responde realmente a la pregunta, muestra cierta com-

importante.	presión general de la materia.
7. Buen uso de lo más importante.	Confuso, mal organizado, no hay nada relevante.

Estas descripciones de los trabajos (o respuestas abiertas a un examen) de los alumnos se pueden hacer de muchas maneras y se pueden adaptar a temas específicos. Si se van a describir tres respuestas (o modelos) tipo, el mejor podría tener las características descritas en el cuadro Xi (adaptado de un ejemplo mencionado por Heywood, 1989, y pensado para corregir y evaluar 4500 ejercicios).

Cuadro XI

Modelo de la mejor respuesta (pregunta abierta o ejercicio, Heywood, 1989)

1)	El desarrollo se centra en el tema propuesto,
2)	Tiene una organización clara,
3)	Los puntos principales se desarrollan de manera lógica y se apoyan en datos o ejemplos,
4)	Selecciona bien palabras y frases, usa correctamente las expresiones transicionales y mantiene un tono o estilo consistente

Es normal que estos trabajos (por ejemplo redacciones en exámenes de admisión) los evalúen varios profesores, y en estos casos es importante que las instrucciones sean muy claras. Las *instrucciones a los correctores* que acompañan a la respuesta-tipo anterior (y otras dos descripciones, la de una respuesta normal y suficiente y una tercera respuesta, redacción o ejercicios deficiente, son estas,.

Lea la redacción rápidamente para tener una impresión global de su calidad en relación con los tres modelos propuestos y asignar una puntuación según su parecido a uno de estos modelos. No se requiere analizar la redacción, sino compararla con los modelos tipo.

Madigan y Brosamer (1991) presentan otros ejemplos de juicios pensados previamente para evaluar preguntas abiertas en exámenes-, es útil reproducir, un poco libremente, uno de los esquemas de corrección-calificación que proponen (y que adaptamos con cierta libertad). Estos criterios de calificación están puestos en el cuadro Xii. Se supone que las respuestas van a caer en una de estas cinco clasificaciones,- la puntuaciones posibles van en este caso de 0 a 20, con lo que a cada descripción o categoría de clasificación le puede corresponder un número distinto de puntos, a juicio de profesor corrector, que puede matizar más su juicio. Los 20 puntos se podrían haber repartido por igual entre las cuatro categorías, pero el profesor en este caso ha decidido hacer otro reparto.

Cuadro XII
 Criterios para evaluar **preguntas** abiertas (Madigan y Brosamer, 1991)

Puntos	Criterios
0-5	La respuesta es básicamente irrelevante para lo que se pregunta, o muestra una importante confusión conceptual
6-13	La respuesta deja fuera aspectos importantes de la pregunta, o contiene errores fácticos importantes
14-19	Responde de manera apropiada a todas las partes de la pregunta, pero no con suficiente profundidad, o incluye material irrelevante o errores menores.
20	La respuesta es completa, concisa, y contiene ejemplos u observaciones personales. No incluye material irrelevante.

Es útil caer en la cuenta en algunos de los aspectos que figuran en estos criterios: Incluir o no todo lo importante, incluir o no material irrelevante, aportar o no ejemplos u observaciones personales, etc. Los Autores incluyen también un esquema semejante para evaluar el estilo de cada pregunta (*writing skill*), que tiene un valor menor en la calificación o puntuación (de 0 a 5 puntos). Los puntos en que se basan los criterios para asignar de 0 a 5 puntos son: *legibilidad*, coherencia y unidad de cada párrafo, construcción apropiada de párrafo, etc.

b) Criterios según modelos determinados.

No se trata de buscar respuestas tipo, sino de escoger buscándolas intencionadamente, tres ejercicios de otros tantos alumnos que sirvan de referencia (muy bueno, suficiente, deficiente) y los demás se van clasificando según e parezcan más o menos a los modelos de referencia o dejándolos en los espacios intermedios. El seleccionar modelos entre varios correctores (por consenso) es un procedimiento que se desarrolló a principios del siglo XX, sobre todo para corregir redacciones (Palmer, 1.961).

El que estos métodos sean adecuados dependerá de la situación (número de ejercicios, tiempo disponible, importancia o peso de estas correcciones en la calificación final, etc.) y tipo de ejercicios (que pueden prestarse más o menos a esta clasificación previa). Siempre será verdad que este tipo de corrección-calificación, que puede parecer extremadamente simple y rápida es preferible a cualquier alternativa claramente peor (como no corregir, hacerlo de cualquier manera, etc.)

ENFOQUE ANALITICO

Este enfoque es más adecuado para exámenes convencionales con varias preguntas. Se trata de preparar una clase de corrección para cada pregunta y atenerse a ella en la corrección de cada pregunta. En cada pregunta se diferencian aspectos, partes que en principio se pueden corregir por separado y que en conjunto abarcan todos los aspectos de una respuesta correcta..

Cuadro Xi
Ejemplo de claves de corrección

	mal	bien	mal	suficiente	bien
1. Información básica:	0	1	0	1	2
2. Incluye información adicional,	0	1	0	1	2
3. Muestra comprensión, entiende lo que dice:	0	1	0	1	2
4. Organiza bien la respuesta, hay claridad-	0	1	0	1	2
5. Ortografía y limpieza, presentación-	0	1	0	1	2

La clave puede ser tan simple o tan compleja como se desee (si es muy compleja, deja de ser útil,- utilizar *unos y ceros* es lo más oráctico, sobre todo si la clave es muy específica); puede haber una clave distinta para cada pregunta o una común para todas, etc.

Si la respuesta requiere solamente aportar información (nombres, fechas, características, causas, etc.), ésta puede pormenorizarse en la clave. Aunque algunos aspectos se pueden *ponderar* (darles más valor para que *pesen* más), siempre será preferible *tener más información de lo más importante que dar más importancia a una poca información*. (Lo muy importante puede aparecer en la clave de todas las preguntas, o puede haber más preguntas sobre lo más importante).

La puntuación de cada pregunta, y de todo el ejercicio, dependerá de la clave utilizada,- cuatro preguntas, con cinco aspectos diferenciados cada una, dará un máximo de 20 puntos. Esta puntuación se convierte después en una calificación convencional (de 0 a 10). No hay que utilizar necesariamente la *clave de calificación como clave de corrección*, como sería el caso de cinco preguntas calificada cada una de 0 a 2 para que el total dé 10 como máximo y coincida con la *nota*. Se trata de que la *calificación* (o juicio de valor) no *contamine* la corrección.

A veces los criterios se desprenden de la propia definición de los aprendizajes y de la correcta realización de las tareas implicadas, su patrón de respuestas correctas es claro. Por ejemplo, en materias como matemáticas o estadística.

Otras veces es necesario que los criterios sean más amplios y ofrezcan un marco más flexible que facilite su articulación, para poder establecer diferentes niveles de realización.

Ejemplo de corrección de trabajo escrito

Criterios	Peso	Niveles de realización			Puntuación ponderada
Organización y estructura	3			X	6
Actualidad y relevancia	2		X		4
Aportaciones personales	2		X		4
Originalidad	1	X			0
Aspectos formales	1		X		1
Fuentes de información	1		X		1
0 = ejecución deficiente; 1 = Ejecución mediana; 2 = Ejecución superior					

Este sistema (u otros análogos, con clave de corrección previamente establecida) suele tener estas ventajas:

1. Centra la atención del corrector y le facilita el trabajo; *no se deja engañar* cuando el alumno cambia de tema;
2. Minimiza los riesgos de subjetivismo, de influjo de la idea previa que se tiene de alumno, etc.
3. Facilita la calificación, hay *más matiz* en la puntuación total de cada alumno; permite además comparar con más objetividad unas calificaciones con otras;
4. Facilita la comunicación de resultados de una manera más provechosa (porque es más específica) para el aprendizaje de alumno;
5. Facilita la evaluación por objetivos (o por aspectos) de todo el grupo (bien la información, mal la organización de la pregunta en general, etc.) y de metodologías, programas, etc.

Quando se sigue un sistema analítico de corrección, suele recomendarse no corregir *alumno por alumno*, sino *pregunta por pregunta*: la primera pregunta a todos, después la segunda a todos, etc. Las ventajas que suelen aducirse sobre esta recomendación son de este orden:

1. Es más descansado para el profesor, que se centra en una única pregunta, sin tener que saltar mentalmente de una a otra,
2. Se minimiza el influjo de las primeras impresiones, mala letra, etc.
3. A través de una pregunta se va viendo el nivel de la clase,
4. Si se interrumpe la corrección y se continúa en un tiempo distinto, cualquier cambio en el profesor (humor, cansancio,...) afecta a todos los alumnos por igual (a partir de la misma pregunta).

Estas recomendaciones no tienen más valor que el de sugerencias que admiten múltiples adaptaciones a asignaturas y personas distintas.

EJEMPLO DE CORRECCION ANALITICA

1) Se establecen los criterios generales de valoración de la prueba. En general, se le asigna un peso o importancia a cada uno de estos.

2) A cada pregunta se le asigna una cantidad de la puntuación total de la prueba. Así, por ejemplo, si la prueba se califica en base 10, se le puede asignar a una pregunta 2 puntos, a otra 3, etc. En la asignación de esta puntuación pueden intervenir factores como, por ejemplo: importancia de; contenido pedido, transferibilidad de otros contenidos, énfasis puesto en la fase de explicación, complejidad de la pregunta, tiempo necesario para poder contestarla, etc.

3) Valoración de las respuestas dadas a cada pregunta. Cada respuesta se compara con la contestación ideal de la clave para calificar. Hay dos maneras de puntuar:

A) Se puntúa cada respuesta dándole una puntuación para cada criterio. La puntuación de la pregunta se obtiene sumando los productos de las puntuaciones en base 10, otorgadas por el profesor para cada criterio, por sus respectivos pesos y multiplicando este total por el valor relativo de la pregunta en la prueba.

A continuación se recoge un ejemplo de calificación de un sujeto hipotético en una prueba de ensayo formada para cuatro preguntas, con el valor de cada una en base 10.

EJEMPLO DE CALIFICACION DE UNA PRUEBA DE PREGUNTAS ABIERTAS					
Preguntas	Valor Pregunta/total	Contenido (0,5)	Organización (0,3)	Claridad expr. (0,2)	Puntuación Pregunta En base 10
1	3p	8x0,5	6x0,3	7x0,2	2,16
2	3p	5x0,5	6x0,3	5x0,2	1,59
3	2p	3x0,5	4x0,3	9x0,2	0,9
4	2p	8x0,5	7x0,3	8x0,2	1,54
Total prueba = 10					Puntuación alumno = 6,19

Así, por ejemplo, la puntuación en base 10 asignada a la primera pregunta (2,16) es iguala:

$$\frac{[(8 \times 0,5) + (6 \times 0,3) + (7 \times 0,2)]}{3} = 2,16$$

B) Otra manera de calificar cada pregunta es asignándole una puntuación global, teniendo en cuenta el peso de los criterios, pero sin matizarla puntuación otorgada a cada uno de estos. Se puede asignar una puntuación numérica o indicar un nivel global de realización. Si se establecen cinco niveles, estos podrían ser: muy deficiente, deficiente, suficiente, bien y muy bien.

4) La puntuación total de; alumno en la prueba viene dada por la suma de los puntos obtenidos en las preguntas.

El seguimiento de los pasos anteriores permite:

- Prevenir o reducir al mínimo la influencia de factores extraños y se proporcionan puntuaciones más fiables.

- Favorecer la explicación al alumno de las razones de su calificación, ya que hace explícitos sus aciertos o errores.

4

Sugerencias a seguir durante el proceso de calificación

Al iniciar el proceso de calificación propiamente dicho, conviene seguir una serie de sugerencias encaminadas a conseguir una mayor confiabilidad, como por ejemplo:

- Mezclar las pruebas antes de iniciar la calificación.
- Evaluar pregunta por pregunta. Calificar una pregunta al mismo tiempo en todas las pruebas o sujetos con el objeto de disminuir el efecto de halo y que el corrector se familiarice con una sola clase de criterios de calificación.
- Procurar evaluar en una única sesión (o en el menor número de sesiones posible) todas las respuestas dadas a una determinada pregunta.
- Evaluar de forma anónima.

Otras sugerencias a tener en cuenta siempre que sea posible:

- Si las decisiones son a tomar son importantes, se habría de procurar tener dos evaluaciones independientes de la prueba, ya sea mediante dos calificaciones o la misma persona en dos ocasiones diferentes. Eso es necesario para comprobar la fiabilidad de la calificación (inter o intra calificador)
- Emitir comentarios y corregir errores. Eso permite analizar los errores comunes, proporcionar una realimentación efectiva del alumno y justificar calificaciones individuales.

Valoración comparativa de las pruebas objetivas (*tipo test*) y de las preguntas de respuesta abierta (*tipo tradicional*) en relación a su validez para comprobar determinados objetivos (siguiendo con cierta amplitud la *taxonomía de objetivos* de Bloom, Hastings y Madaus, 1973)

dos objetivos (siguiendo con cierta amplitud la *taxonomía de objetivos* de Bloom, Hastings y Madaus, 1973)

Objetivos	Pruebas objetivas	Pruebas abiertas
Conocimientos sobre todo de MEMORIA: hechos específicos, datos, definiciones. Símbolos, características, clasificaciones, etc.	Las pruebas objetivas son válidas, fáciles de construir para este tipo de objetivos; se puede comprobar mucho en poco espacio y poco tiempo.	El sistema tradicional es válido, pero posiblemente menos cómodo para el profesor, que tiene que ir buscando la información relevante. Sobre todo cuando la pregunta requiere una respuesta amplia. En general las pruebas objetivas, en sus diversas modalidades, son más cómodas para comprobar estos objetivos.
COMPRESION (nivel sencillo) de leyes, conceptos, principios, etc., de relaciones, implicaciones, etc, interpretar, traducir. Identificación de ejemplos y definiciones correctas, etc.	Para estos objetivos de comprensión, interpretación, análisis sencillos, etc., la prueba objetiva puede ser muy válida, pero es más difícil de preparar, hay buenos modelos en la abundante bibliografía sobre evaluación.	El sistema tradicional es válido, el inconveniente suele estar en la corrección, que es muy laboriosa, también puede suceder que se acepte como prueba de comprensión lo que en realidad es una repetición de memoria, de apuntes, etc.; la pregunta debe ser muy clara y orientadora,
Capacidad de ORGANIZAR ideas o temas, de presentar una información de modo coherente, etc., de dar estructura, de exponer de una manera personal, etc.	La prueba objetiva es inútil en este caso (y se trata de objetivos muy importantes).	El sistema de respuesta abierta es en principio el único sistema válido en sus variantes más apropiadas (preguntas de respuesta larga, trabajos hechos en casa, etc.), conviene tener claro qué es una buena organización si es eso lo que se va a evaluar, es frecuente preferir estas preguntas porque permiten al alumno organizar, expresarse, etc., y evaluar luego preferentemente la cantidad de información son especialmente útiles las claves de corrección
ESTILO y presentación.	La prueba objetiva es inútil en este caso (y se trata de objetivos importantes),	El sistema de respuesta abierta es el único sistema válido, pero habrá que tener en cuenta esos objetivos (al corregir, evaluar, e informar) los alumnos pueden acostumbrarse a escribir mal.
Capacidad de ANALISIS, de diferenciar, de descubrir las partes del todo, de identificar relaciones más complejas, etc., de capacidad crítica	Las pruebas objetivas pueden ser válidas, pero difíciles de preparar (habrá que buscar modelos, presentar material nuevo en las preguntas, etc.)	Sistema válido, pero con los inconvenientes dichos para la comprensión, en ambos casos habrá que cuidar la redacción de la pregunta; es útil disponer de claves de corrección que ayuden al profesor a centrarse en el objetivo que desea evaluar.
Capacidad de APLICACIÓN, de resolución de problemas, de aplicar normas y métodos en situaciones nuevas	la prueba objetiva es válida en general 1º para problemas sencillos, 2º para evaluar por separado los pasos y habilidades necesarios en la resolución de problemas complejos; útil con fines diagnósticos, para detectar fallos generalizados ya no es tan válida, o no lo es. En la resolución de problemas complejos	Sistema válido; con frecuencia el único válido; útil disponer de una clave que facilite la evaluación de todo el proceso, por ejemplo, identifica a) el problema, b) el principio o principios, etc. que permitirán resolverlo, etc. (no tiene por qué haber una clave común ni todos los problemas a aplicaciones pertenecen al ámbito de las matemáticas, etc.).
CREATIVIDAD, capacidad de hacer síntesis nuevas, proponer un plan de operaciones, de producir algo original, etc.	la prueba objetiva es inútil en este caso, puede ser útil de manera indirecta, para evaluar capacidades o habilidades que son necesarias para la creatividad.	Único sistema válido; el sistema más apropiado para evaluar creatividad no suele ser lo que se puede preguntar en un examen convencional, sino trabajos hechos en casa, proyectos, etc. es útil tener criterios claros (y una clave) sobre cómo se manifiesta la creatividad en una situación dada, la evaluación no es fácil porque no hay una respuesta correcta.
Capacidad de EVALUACION según un criterio dado: coherencia interna, utilidad, bondad según una norma, etc.	La prueba objetiva puede ser válida y útil, pero difícil de preparar, no siempre habrá con claridad una única respuesta correcta.	Sistema válido, y único cuando el criterio personal es importante, cuando hay que exponer razones, etc.
ACTITUDES, valores, intereses, etc. (dominio afectivo)	Sistema válido en sus diversas modalidades (escalas, etc. hay muchos modelos), facilita la sistematización, interpretación y análisis de las respuestas,	Sistema válido, pero es difícil sistematizar las respuestas buen complemento al menos de los sistemas más objetivos

Pedro Morales (1995). Tipos de pruebas: los exámenes orales y las preguntas de respuesta abierta. Bilbao: ICE de la Universidad de Deusto.