

Exámenes Escritos de Competencias

Ítems de Selección Múltiple

Lineamientos para su construcción y análisis

Lic. Norma Tauro
Dra. María del Carmen Binda Dávila, Dr. Juan Mazzucco,
Lic. Florencia Moore, Lic. Claudia Lázaro

El lugar que ocupan las Sociedades Científicas de cada Especialidad

- Las Sociedades Científicas cuentan con prestigio académico y social, nuclean a profesionales de las diferentes especialidades
 - Son entidades que desarrollan el conocimiento y asesoran a gobiernos, universidades y diferentes organizaciones
 - Promueven y estimulan la capacitación profesional, haciendo hincapié en la importancia de los valores éticos y morales de la profesión
-

La Educación Continua en las Sociedades Científicas

- En las Sociedades Científicas se estimula el desarrollo del conocimiento para alcanzar el perfeccionamiento de sus miembros, que se refleja en el objetivo final de alcanzar el mayor beneficio de los pacientes
- Para ello se implementan actividades de educación continua tales como: congresos, cursos, jornadas, publicaciones de Revistas y Boletines; como programas de aprendizaje permanente

Cómo acreditar que se “sigue siendo” Especialista

La autorización oficial tendrá una duración de 5 años, podrá ser revalidada cada 5 años mediante acreditación de antecedentes:

- demostrar continuidad en la especialidad durante ese lapso
- una entrevista personal
- examen de competencia

La evaluación en las Sociedades Científicas

- Las Sociedades Científicas han puesto en marcha mecanismos de evaluación de los médicos, que son denominados certificaciones y recertificaciones (o mantenimiento)
 - Estas evaluaciones son voluntarias, periódicas, y son realizadas por los pares
 - Las certificaciones y recertificaciones instalan la necesidad de garantizar la calidad profesional
 - En el área de las Ciencias de la Salud, las evaluaciones sumativas implican una gran responsabilidad ya que su aprobación significa, en ocasiones, reconocer competencias y autorizar determinadas prácticas que se vinculan con la salud de las personas.
 - Los exámenes que se creen y apliquen deben ser objetivos, válidos y fiables.
 - La evaluación no debe ser solo poner un puntaje o aprobar o reprobar., debe ser una instancia de formación profesional.
-

El Examen de Competencia se enmarca en la "Evaluación"

La evaluación es un proceso por el cual se intenta comprobar que se han alcanzado los niveles de calidad esperados

Así, en el marco de la evaluación están incluidos conceptos tales como: medir, comparar, valorar, categorizar, ingresar, admitir, aprobar, promocionar, calificar, puntuar, clasificar, acreditar, certificar, recertificar y otros como, reconocimiento oficial, provisorio o definitivo, estándares, patrones, criterios, objetivos, pruebas, instancias e instrumentos de evaluación, pares evaluadores.....

La Evaluación transita por diferentes Momentos

Evaluar implica: búsqueda de información, establecer un juicio de valor y toma de decisiones.

1º- Toda evaluación incluye la búsqueda de información, con instrumentos válidos y fiables

2º- La información obtenida se compara con un parámetro –competencias y conocimientos que se esperan de un especialista-

3º- Se hace una valoración de los resultados obtenidos

4º- La valoración de esta comparación entre los resultados y los parámetros esperados permite:

5º- La toma de decisiones para elevar la calidad

Sobre los evaluados (otorgar la certificación o no)

Sobre el instrumento de evaluación empleado (optimizarlo, mejorarlo, suprimirlo)

Sobre las instancias de formación continua de los profesionales médicos(Proponer más o distintos tipos de: Cursos, Jornadas, Congresos, Prácticas de simulación, lecturas de publicaciones)

6º- Los resultados son compartidos con los evaluados -feed back-

Existen tres tipos de Exámenes para evaluar el desarrollo profesional continuo

- Exámenes de Desempeño
- Exámenes Orales
- Exámenes Escritos

• Exámenes de Desempeño: permiten evaluar el accionar en el ámbito específico donde se desarrolla la profesión, las destrezas, prácticas y experiencias in situ. Los instrumentos pertinentes, válidos y confiables para evaluar el desempeño son las guías de observación, las listas de cotejo y los registros anecdóticos.

• Exámenes Orales: permiten evaluar el conocimiento y análisis de cierta información, el pensamiento crítico y el razonamiento médico; son útiles para detectar la "estructura cognitiva" del sujeto, las habilidades comunicacionales y la pericia en la exposición oral. Los instrumentos pertinentes, válidos y confiables son las entrevistas y las exposiciones dialogadas.

• Exámenes Escritos: permiten evaluar información, comprensión y aplicación - en términos intelectuales- de datos y casos clínicos; son útiles para detectar la capacidad de exploración, interrelación de conceptos e inferencias.

Los instrumentos pertinentes, válidos y confiables son las pruebas no estructurada, las semi-estructuradas y las estructuradas u objetivas.

Las Pruebas Escritas

Las Pruebas Escritas se clasifican en:

- * NO ESTRUCTURADAS Producciones, posters, publicidades, publicaciones, portafolios
- * SEMI ESTRUCTURADAS: {
 - de respuesta restringida
 - de libro abierto, monografías, tesis
- * ESTRUCTURADAS
U OBJETIVAS {
 - de recordación {
 - recordación simple
 - completamiento
 - de reconocimiento {
 - doble alternativa
 - relación
 - ordenamiento
 - selección múltiple

Metodología para diseñar las Pruebas Escritas: garantizar su calidad

- Respetar los pasos de la Evaluación
- Se debe planificar el examen
- La construcción, selección y aplicación de la prueba debe estar guiada por los Objetivos Educativos: los aprendizajes y competencias que se intentan confirmar
- En todos los casos, antes de aplicar el examen, se debe decidir previamente cuál es la respuesta correcta, clave de respuesta o respuesta esperada. Debe existir consenso entre los expertos sobre cuál es la respuesta correcta
- Las opciones no pueden ser parcialmente correctas/verdaderas o parcialmente incorrectas/falsas
- Los ítems deben estar contruidos con precisión y claridad, sin ambigüedades

Para garantizar su Validez - Objetividad - Confiabilidad

Se debe planificar el examen y garantizar la calidad de las Pruebas Escritas

I Validez

Para garantizar la Validez se debe:

- Decidir la *Tabla de Especificaciones*

Es un gráfico en el que aparecen qué objetivos serán evaluados y en qué nivel: información-comprensión - aplicación

- Decidir previamente la tabla de especificaciones permite consensuar la equidad en cuanto a qué objetivos han de ser evaluados: en ella se vuelca la decisión sobre *qué temas* se evaluarán, *qué cantidad de preguntas* por tema y el *nivel taxonómico* de las mismas.
- Decidir previamente la tabla de especificaciones permite evitar desequilibrios: a veces se hacen muchas preguntas sobre un mismo tema o sobre la información de una unidad temática y pocas sobre su aplicación. Este error se puede prevenir decidiendo con anterioridad una distribución planificada de las competencias y aprendizajes a ser evaluados.

El Mapa de la Especialidad

La Tabla de Especificaciones permite un relevamiento del *Mapa de la Especialidad*

Por medio del Mapa de la Especialidad se realiza una vigilancia epistemológica y un análisis del territorio y las fronteras de las prácticas que se enmarcan en la Especialidad

Tabla de Especificaciones: un ejemplo

Por ejemplo para una prueba de 80 ítems que evalúe un Programa educativo con 4 Unidades Temáticas se puede elaborar la siguiente *Tabla de Especificaciones*

	Información	Comprensión	Aplicación	Total
Tema 1	4 Ítems	6 Ítems	10 Ítems	20
Tema 2	4 Ítems	4 Ítems	12 Ítems	20
Tema 3	2 Ítems	10 Ítems	14 Ítems	26
Tema 4	2 Ítems	6 Ítems	6 Ítems	14
Total	12 Ítems	26 Ítems	42 Ítems	80 Ítems

Niveles taxonómicos que incluye la Tabla de Especificaciones

- *Memoria* recordar un hecho específico, nombrar, identificar, enunciar, enumerar, definir, describir
- *Comprensión* interpretar, ejemplificar, clasificar, comparar, explicar, resumir
- *Aplicación* utilizar el conocimiento en una situación particular, analizar datos y sacar conclusiones, justificar, planificar, tomar decisiones (diagnóstico, conducta a seguir, solicitar estudios, indicar tratamiento)

II Objetividad

Para garantizar la Objetividad

- Se seleccionan o construyen previamente las preguntas/ítems a realizar
- Se decide cuál es la Clave de respuesta o respuestas esperadas antes de aplicar el instrumento
- El consenso entre los expertos sobre cuál es la respuesta correcta, garantiza la objetividad
- Los Ítems deben basarse en bibliografía recomendada y respaldatoria de la clave de respuesta o respuesta esperada

III Confiabilidad

Para garantizar la Confiabilidad

- El equipo evaluador debe:
- seleccionar y/o crear preguntas o casos equivalentes para todos los evaluados, tanto a nivel de abordaje de los contenidos como de los niveles de dificultad
 - garantizar el mismo tiempo de resolución para cada evaluado
 - crear la grilla de respuesta con las claves para ser utilizada por los correctores
-

Pruebas Escritas Estructuradas u objetivas

Qué caracteriza a las Pruebas Escritas Estructuradas u objetivas

- Exigen una respuesta breve y concisa
- Tienen una sola respuesta correcta
- No permiten una interpretación personal del alumno

Ventajas de las Pruebas Escritas Estructuradas u objetivas

- Los puntajes son más objetivos
- Se prestan menos a la ambigüedad
- Son pruebas confiables
- Pueden medir conocimientos con la misma o mayor efectividad que otros tipos de prueba
- Permiten diagnosticar deficiencias en el aprendizaje que necesitan reajustes
- Permiten evaluar a muchas personas en poco tiempo
- Son fáciles de calificar

Desventajas de las Pruebas Escritas Estructuradas u objetivas

- Capacidad limitada para medir dimensiones cognitivas de alto nivel y complejas como la creatividad y la habilidad para resolver problemas
- Imposibilidad de medir destrezas
- Dificultad en su construcción
- Dificultad en la elaboración de distractores adecuados
- Demandan mucho tiempo en su elaboración

Nivel de conocimiento que se evalúa en las Pruebas Escritas Estructuradas u objetivas

1. *Información*: significa la capacidad de recordar hechos, valores, clasificaciones, procedimientos, etc. Implica sólo hacer presente la información apropiada en el momento preciso
2. *Comprensión*: implica el entendimiento de algo y la posibilidad de justificar. Por medio de la comprensión se establecen vínculos entre los conceptos relevantes, las ideas principales, y las situaciones que los ejemplifican
3. *Aplicación*: implica la elaboración de la información o su uso en situaciones particulares y concretas. Para responder bien se debe tener información y comprender el tema. Esto es ir más allá del recuerdo mecánico. Los ítems de aplicación son de gran utilidad para evaluar las capacidades de analizar, realizar juicios críticos y tomar decisiones

Tipos de Ítems que incluyen las Pruebas estructuradas u Objetivas

-Doble alternativa

Verdadero o Falso se basan en el principio lógico de la no contradicción

-Relación

Se intenta indagar el proceso mental por el cual el alumno establece relaciones entre dos categorías de información, se detecta si utiliza criterios de vinculación coherentes y acordes con el contexto, además de verificar si posee la información correcta

-Ordenamiento

Se utilizan para verificar que el alumno conoce la secuencia correcta o las escalas adecuadas

-Selección múltiple

Entre varias posibilidades se debe elegir la opción correcta.

Tiene la utilidad de detectar la capacidad de discriminación en el alumno

Ítems de Selección Múltiple

La Lógica de los *Ítems de selección múltiple* es presentar una situación confusa, desordenada en sí misma -como la realidad que nos circunda- ya que es la mente del hombre la que categoriza, encuadra y enmarca.

La resolución del Ítem permite detectar si se ha alcanzado o no, cierto nivel de *discriminación* entre la opción correcta -clave de respuesta- y los distractores -opciones incorrectas-.

Es decir, ante informaciones con características similares se responde correctamente sólo cuando se es capaz de detectar aquellas características particulares y únicas, que incluso son a veces imperceptibles para quienes no conocen el tema o no han alcanzado los aprendizajes esperados-

La confusión es totalmente válida en este tipo de evaluación, ya que solo podrá responder correctamente quien pudo discriminar, distinguir, separar las respuestas correctas de las que no lo son.

Ítem de selección múltiple

Un ejemplo

De las siguientes seleccione la opción correcta
Una recién nacida de 36 semanas de edad gestacional que nace con un peso de 3000 g, se **clasifica como Recién Nacido**:

- a) A término con peso adecuado para EG
- b) Pre-término con peso adecuado para EG**
- c) Pre-término de bajo peso para EG
- d) Pre-término de alto peso para EG

Las partes que componen un Ítem de selección múltiple

Consigna De las siguientes seleccione la opción correcta

**Base del Ítem
(Situación)**

Una recién nacida de 36 semanas de edad gestacional que nace con un peso de 3000 g, se **clasifica como Recién Nacido**:

**Opciones para
seleccionar
la respuesta**

- | | |
|--|------------|
| a) A término con peso adecuado para EG | distractor |
| b) Pre-término con peso adecuado para EG | CLAVE |
| c) Pre-término de bajo peso para EG | distractor |
| d) Pre-término de alto peso para EG | distractor |



raíz:

deben pertenecer a la misma categoría

*deben ser plausibles

*igual extensión que la clave



En el Ítems de Selección Múltiple se puede solicitar que

- ~ se elija la opción **correcta**
- ~ se elija la opción **incorrecta**
- ~ se exprese la **relación** correcta entre...
- ~ se seleccione la **mejor razón** (la conducta más adecuada)

Cómo organizar la base

La base debe formular claramente el problema y evitar la ambigüedad.

Debe presentar el problema o situación en un contexto que implique aplicar el conocimiento y hacer asociaciones.

Debe comprenderse sin necesidad de leer las opciones.

Debe haber concordancia gramatical entre la base y las opciones

Debe incluir lo estrictamente necesario para comprender el correcto sentido de la respuesta.

Debe evitar hacer preguntas que solo exploren la capacidad de recordar conocimientos abstractos.

Las palabras que eventualmente se repiten en todas las alternativas son parte de la base y no de las opciones.

No debe agregar nada que debilite o confunda la elección de la respuesta correcta.

En la base se evitarán enunciados sugerentes y la utilización de indicios que faciliten la elección de la respuesta correcta.

Se debe cuidar que la base esté libre de "claves" que revelen la respuesta.

Es preferible que los términos o conceptos a definir estén en la base y las descripciones o definiciones en las alternativas.

Se evitará el uso de términos ambiguos: frecuentemente, en alto grado, etc. (todo lo que pueda depender de las subjetividades o varíe según distintos autores).

Solo se preguntará por cantidades o por porcentajes cuando la discriminación cuantitativa sea relevante.

Se evitará utilizar enunciados negativos.

Es más natural preguntar sobre lo que las cosas son y no sobre lo que no son.
Las formas negativas son útiles para reforzar enseñanzas de lo que NO debe hacerse.

En el caso de que se las utilice destacar el NO y de la misma manera cuando se solicita discriminar la EXCEPCIÓN.

Cómo construir las opciones

- Se evitará utilizar ítems que incluyan más de una idea o que evalúen dos aspectos independientes al mismo tiempo.
- Se evitarán ítems en los que la respuesta correcta depende de un detalle insignificante.
- Si bien cada pregunta puede tener entre 3 a 5 opciones, recomendamos que sean 4
- La lista de opciones debe ser homogénea:

- similar extensión
- la misma estructura gramatical
- una terminología semejante
- deben referirse al mismo tipo de cosas

- Tanto la clave de respuesta como los distractores deben ser equivalentes en cuanto a su contenido, deben hablar de la misma categoría de cosas: Todas las opciones deben definir o todas las opciones deben presentar características, intentando que se traten de las mismas clase de cosas.

- Tanto la *clave de respuesta* como los *distractores* deben tener la misma extensión. Por ejemplo, para mejorar la calidad de un Ítems se pueden agregar datos o especificaciones a las distintas opciones, que permitan visualmente observarlas con la misma extensión

- La presentación de las opciones debe guardar un orden lógico, -cifras o cantidades de >< o de <>-.

Los distractores

- deben tener la misma probabilidad de ser elegidos, aun para quien no conozca el tema, es decir no deben autoeliminarse por ser demasiado obvios o absurdos.

- más que inventarse, los distractores, deben recolectarse de los errores comunes que se observan en el proceso de enseñanza.

- Su función no es engañar o poner trampas, es detectar que el evaluado “distingue” la respuesta correcta y precisa entre opciones similares
- Es necesario intentar disminuir la adivinación.
- Todos los distractores deben ser plausibles: posibles de ser elegidos.

- la ubicación de la alternativa correcta debe hacerse al azar.

- se debe evitar utilizar como última opción “Todas son correctas”.

- se debe evitar utilizar como última opción “Ninguna de las anteriores” o “Ninguna es correcta”.

- se debe evitar utilizar como última opción la combinaciones de éstas últimas.

- se debe evitar efectuar una pregunta cuya respuesta correcta no aparezca como alternativa.

Corrección y criterios de aprobación

- El Examen debe ser Correggio con la mayor celeridad posible.
- Se debe decidir previamente cuál será el criterio de aprobación: si será un criterio absoluto o uno relativo
- En el marco de los exámenes de certificación y mantenimiento de la certificación recomendamos que el criterio sea absoluto: el profesional debe responder a niveles de calidad de su desempeño como especialista con pautas externas ya preestablecidas, y que no dependan del desempeño relativo al grupo evaluado.

Dar retroalimentación.

- La retroalimentación se facilita cuando el examen es electrónico e interactivo.
- La devolución es muy importante, tanto si se ha respondido correctamente reforzando y ampliando los aprendizajes alcanzados, como si se lo hizo de manera errónea.
- En el caso de que se haya elegido la respuesta incorrecta se debe argumentar por qué razón no es la respuesta esperada.
- En el caso de que sean preguntas referidas a una publicación científica:
- Tanto para reafirmar lo correcto o para refutar una opción incorrecta, las argumentaciones deben estar explícitas en el texto o inferirse del mismo. Siempre se debe hacer mucha referencia al texto. No hay que olvidar que las preguntas se derivan del mismo y tienen como objetivo asegurar su lectura y comprensión

Análisis de la calidad del Instrumento

Una vez aplicado el Instrumento debe ser revisado, es decir evaluada su calidad.

El análisis de la calidad del Instrumento permite conocer la real efectividad de la educación continua de los especialistas y la calidad de los instrumentos utilizados en la evaluación.

Los exámenes que se creen y apliquen deben ser objetivos, válidos y fiables para que los mismos puedan acreditar que se "sigue siendo" Especialista

En el área de las Ciencias de la Salud, las evaluaciones sumativas implican una gran responsabilidad ya que su aprobación significa en ocasiones reconocer competencias y autorizar determinadas prácticas que se vinculan con la salud de las personas.

El objetivo de evaluar la calidad del Instrumento es garantizar la calidad del proceso de recertificación. Evaluar la calidad del instrumento permite alimentar la *base de datos de ítems* con aquellos considerados "buenos", mejorar los "débiles" y descartar los "malos"

La Calidad del instrumento se observa en:

- *La Tabla de especificaciones*

- la cantidad de temas y de preguntas fue suficiente y equitativo
- fue equilibrado el nivel taxonómico de las preguntas

Memoria
Comprensión
Aplicación

- *Clave de corrección*

- debe existir previo consenso entre los Especialistas

- *Opciones*

- que todas sean plausibles, elegibles
- todas las opciones deben resultar atractivas y homogéneas
- una opción es buena técnicamente si es elegida, por lo menos, por el 5% de los especialistas evaluados
- se dice que un distractor es poco plausible cuando no fue elegido por nadie, es decir no funcionó como distractor
- cuando todas las opciones "funcionan" bien la pregunta es más difícil y discrimina más.

Las preguntas *demasiado fáciles* -respondidas bien por todos- o, por el contrario, *muy difíciles* -respondidas mal por todos- requieren un análisis particular pues discriminan mal entre un mayor y menor nivel de conocimientos.

- Calidad técnica de las preguntas: escala Galofré¹: Anexo

Si una pregunta es respondida correctamente por la mayoría de los evaluados...

Se puede inferir:

- que los especialistas han alcanzado la competencia esperada
- que las publicaciones sobre el tema fueron de calidad, leídas y comprendidas por la mayoría
- que el tratamiento de la temática en los diferentes eventos -Cursos Superiores, Congresos, Jornadas,..- fueron eficaces y aprovechadas por la mayoría
- que la clave de respuesta correcta fue muy obvia y evidente
- se puede considerar la anulación de preguntas por "error técnico"

Si una pregunta es respondida erróneamente por la mayoría de los evaluados...

Se puede inferir:

- que los especialistas no han alcanzado la competencia esperada
- que las publicaciones sobre el tema o no fueron de calidad, o no fueron leídas ni comprendidas por la mayoría
- que el tratamiento de la temática en los diferentes eventos -Cursos Superiores, Congresos, Jornadas,..- no fueron eficaces y aprovechadas por la mayoría
- que la clave de respuesta correcta fue errada y no hay consenso entre los especialistas
- se puede considerar la anulación de preguntas por "error técnico"

¿Cómo hacer mejores exámenes?

Sugerencias generales

Verificar qué tipos de contenidos puede aceptar la estructura de las preguntas de opción múltiple.

Que la cantidad de preguntas esté previamente pautada por la Tabla de especificaciones. La cantidad de preguntas debe depender de la cantidad de temas a evaluar

Que los ítems sean buenas preguntas, con *calidad técnica*

Que las opciones sean entre 3 y 5, plausibles, elegibles -buenas "mentiras"-

Como se está evaluando la formación continua de un Especialista:

las preguntas deben tener una viñeta clínica

las preguntas deben ser sobre contenidos con la suficiente relevancia para justificar su elaboración.

asegurar la objetividad del examen consenso entre los especialistas

analizar los resultados antes de calificar, es decir revisar la calidad técnica del

Instrumento

dar retroalimentación.

¹ Galofré, A. *Instrucciones para calcular un índice de calidad para preguntas de selección múltiple*. Universidad Católica del Norte. Chile. 2006

ANEXO

Calidad técnica de una pregunta de selección múltiple Escala Dr. Alberto Galofré

1. **Presencia de viñeta:** se refiere a la presencia de un caso clínico o procedimiento de laboratorio o problema.
2. **Enunciado completo:** cuando al cubrir las opciones el lector puede responder a la pregunta después de leer el enunciado.
3. **Evitar el uso de negaciones:** como Excepto.
4. **Concordancia gramatical entre enunciado y opciones:** que la pregunta del enunciado y las opciones sigan reglas gramaticales. Por ejemplo si se pregunta por “el” no poner una palabra femenina después en una de las opciones.
5. **Distractores verosímiles:** que cada opción parezca verdadera para el que no sabe mucho.
6. **Extensión similar entre las opciones:** es usual que si existe una opción con excesiva longitud en relación a las otras, tiende a ser la respuesta correcta por tener mayor detalle.
7. **Evitar las opciones ninguna y todas las anteriores:** bastaría saber que dos opciones son correctas para saber que si existe una opción de “todas las anteriores” ésta tiene que ser la respuesta. Ninguna de las anteriores se presta para interpretaciones diferentes del que redacta la pregunta. Por ejemplo, un mejor tratamiento o la mejor conducta a seguir después de dar los antecedentes. En formato de pregunta de “una mejor respuesta” estas opciones complican el proceso de responder.
8. **Opciones ordenadas:** ya sea lógicamente o numéricamente. Si hay números, por ejemplo, 5, 10, 15, 20, 25, ‘estos deben colocarse en ese orden. No es recomendable poner opciones con números sin ordenar. Complica la respuesta innecesariamente.
9. **Opciones homogéneas:** similares en cuanto al contenido.
10. **Nivel taxonómico de Aplicación de conocimiento o superior:** que la pregunta requiera usar un proceso superior a simple recuerdo de un hecho memorizado.