



UNIVERSIDAD AUTÓNOMA
DE CHILE

 UNIDAD DE EDUCACIÓN
EN CIENCIAS DE LA SALUD
Universidad Autónoma de Chile



Guía de Apoyo Docente

EVALUACIONES ESCRITAS EN CIENCIAS DE LA SALUD

Construcción y revisión de instrumentos de evaluación
escrita

TABLA DE ESPECIFICACIONES

SELECCIÓN MÚLTIPLE

DESARROLLO BREVE

DESARROLLO EXTENSO

Unidad de Educación en Ciencias de la Salud
Facultad de Ciencias de la Salud
Universidad Autónoma de Chile - 2026

LA EVALUACIÓN ESCRITA IMPORTA, Y MUCHO

Una prueba es válida cuando mide lo que se propone medir, confiable cuando produce resultados consistentes, y alineada cuando responde al perfil de egreso y al programa de asignatura.

La evaluación escrita es uno de los instrumentos más utilizados en la formación universitaria en ciencias de la salud.

Su valor no radica únicamente en asignar calificaciones: **una evaluación bien construida orienta el aprendizaje, evidencia el dominio de competencias y retroalimenta el proceso docente.**

VALIDEZ

Mide lo que se propone medir. Alineación con RA y CE del programa.

CONFIABILIDAD

Resultados consistentes entre distintos evaluadores y ocasiones.

ALINEACIÓN

Coherencia entre el perfil de egreso, la enseñanza y la evaluación.

EVALUAR BIEN TAMBIÉN SE APRENDE

Si al intentar aplicar estos criterios por primera vez te sientes abrumado/a por la cantidad de elementos a considerar, ¡tranquilo/a, es completamente normal!

Construir una buena evaluación escrita es una habilidad pedagógica en sí misma y, como toda habilidad, se desarrolla con práctica, reflexión y retroalimentación.

No buscamos perfección desde el primer intento, buscamos progreso sostenido.

Elige un paso a la vez: esta semana, prueba usar una tabla de especificaciones antes de redactar tu próxima prueba. Observa si cambia la distribución de tus preguntas. Ajusta. La siguiente vez, revisa tus distractores con el checklist.

¡Tú también estás aprendiendo a evaluar mejor!



TABLA DE ESPECIFICACIONES

La tabla de especificaciones es el mapa estratégico de toda evaluación escrita. Permite distribuir proporcionalmente los ítems entre las unidades del programa, asegurar variedad de niveles taxonómicos y garantizar la validez de contenido.

No es un documento opcional: es el instrumento que garantiza que tu prueba mide coherentemente lo que el programa declara que los estudiantes deben aprender. **Asegura alineamiento constructivo.**

Antes de redactar cualquier ítem, construya su tabla de especificaciones en base a su programa de asignatura.

Le recomendamos completar la tabla de especificaciones antes de iniciar cada unidad, esto le dará un panorama de cómo distribuir los contenidos en el aula o talleres para el logro del RA respectivo.

¿Qué evaluar?					¿Cómo Evaluar?									
Resultado de Aprendizaje (RA)	Ponderación planificada	Criterios de Evaluación (CE)	Sub ponderación según puntaje asignado	Contenidos relevantes	Nivel Cognitivo				Cantidad de ítemes	Tipo de ítem	Ubicación del ítem en el instrumento	Puntaje del ítem	Ponderación según puntaje asignado	
					Análisis	Aplicación	Meta - cognición	Autoregulación						
RA 1		CE 1	19%							SM		2	4%	
										SM		2	4%	
										DB		6	12%	
		CE 2	42%							DB		6	12%	
										DB		6	12%	
										DE		10	19%	
		CE 3	38%							SM		2	4%	
										SM		2	4%	
										DE		16	31%	
RA 2		CE 1	0%										0%	
														0%
														0%
		CE 2	0%										0%	
														0%
														0%
		CE 3	0%										0%	
														0%
														0%
Total	0%		100%					0			52	100%		

Consideraciones:

- Escriba el tipo de ítem en la columna correspondiente con abreviaturas. Ejemplo: "SM" (selección única) "DE" (desarrollo extenso) "DB" (desarrollo breve).
- El tipo de pregunta debe tener relación con el nivel taxonómico del criterio de evaluación y por lo tanto debe tener coherencia con las columnas de "nivel cognitivo".
- La distribución de % debe estar en relación con la importancia del contenido/unidad.



8 PASOS PARA CONSTRUIR UNA TABLA DE ESPECIFICACIONES

- 1 Extrae los resultados de aprendizaje (RA) del programa**
 Lee completamente tu programa de asignatura. Identifica todos los RA que se declaran. Estos son tu punto de partida obligatorio.
Pregunta clave: ¿Cuántos RA tiene mi asignatura? (generalmente 1-3 para una asignatura)
- 2 Para cada RA, identifica sus criterios de evaluación (CE)**
 Cada RA tendrá uno o varios CE asociados. El CE es más específico y evaluable que el RA.
 Los CE cumplen dos funciones en tu tabla:
 1. Definen el verbo taxonómico máximo que puedes usar (no pueden superarse)
 2. Especifican qué aspecto concreto del RA será evaluado en la prueba escrita**Pregunta clave: ¿Cuál es el verbo taxonómico del CE? (Ej: "Analiza", "Aplica", "Valora")**
- 3 Asigna una ponderación planificada a cada CE**
 Idealmente este análisis debe ser durante la construcción del Syllabus, previos a iniciar la unidad.
 Esto responde a: ¿Cuánto tiempo de clase/taller dediqué a este CE?
 La ponderación planificada refleja la distribución real de la asignatura:
 - Si dediqué 20 horas a CE1 de un total de 100 horas → 20%
 - Si dediqué 50 horas a CE2 → 50%
 - Si dediqué 30 horas a CE3 → 30%
 Esta ponderación debe ser proporcional y sumar 100%.
Pregunta clave: ¿Cuántas horas (aproximadamente) dediqué en clase/taller a cada CE?
- 4 Identifica los contenidos relevantes asociados a cada CE**
 Cada CE se fundamenta en contenidos específicos del programa de asignatura.
 Pueden ser:
 - Conceptos (Ej: mecanismos fisiológicos)
 - Procedimientos (Ej: técnicas de dispensación)
 - Normativa (Ej: regulaciones vigentes)
 - Habilidades (Ej: comunicación con pacientes)**Pregunta clave: ¿Qué contenidos mínimos del programa debo evaluar en este CE?**
- 5 Define el nivel taxonómico máximo por CE**
 Aquí es donde el CE establece el techo cognitivo de tu evaluación. Recuerda los 6 niveles de Marzano.
 Regla clave: El CE define dónde debe llegar como mínimo tu evaluación. No puedes quedarte solo en niveles inferiores.
 Ejemplo lógico:
 - Si el CE dice "Analiza", no puedes medir solo "Recuperación" y "Comprensión"
 - Si puedes mezclar: algunos ítems en Análisis (lo obligatorio) + algunos en niveles inferiores si necesitas verificar prerrequisitos**Pregunta clave: ¿Cuál es el verbo del CE y qué nivel de Marzano le corresponde?**
- 6 Decide cuántos ítems necesitas y de qué tipos**
 A) Número de ítems por CE que depende de tu ponderación planificada
 - Un CE con 40% de tiempo → más ítems
 - Un CE con 10% de tiempo → menos ítems
 B) Tipo de ítems según el nivel taxonómico: Coherencia. No puedes medir "Análisis" solo con selección múltiple simple.
Preguntas clave: ¿Cuántos minutos totales tengo para la prueba? ¿Cuántos ítems me permiten cubrir todos los CE? ¿Qué tipo de ítem es más apropiado para cada nivel taxonómico?
- 7 Asigna puntaje a cada ítem y calcula la ponderación según puntaje**
 La ponderación según puntaje es el reflejo numérico de:
 - Si CE1 tiene 3 ítems de 5 puntos c/u → 15 puntos → 15% de la prueba total
 - Si CE2 tiene 8 ítems de 5 puntos c/u → 40 puntos → 40% de la prueba total
 Esta ponderación según puntaje debe coincidir (aproximadamente) con tu ponderación planificada. Si planeaste 40% para CE2, debe tener ~40% del puntaje total.
Pregunta clave: ¿El puntaje asignado refleja fielmente cuánto tiempo dediqué a cada contenido?
- 8 Verifica la coherencia final**
 ✓ ¿Todos los RA están representados?
 ✓ ¿Cada CE tiene al menos un ítem que mida su verbo taxonómico?
 ✓ ¿Los ítems de mayor complejidad corresponden a los CE con mayor ponderación?
 ✓ ¿El tiempo total de la prueba es realista (cada ítem × tiempo estimado)?
 ✓ ¿La cobertura de contenidos es equilibrada?
Si algo no cuadra, vuelve a los pasos 4-6 y reajusta.

PREGUNTAS DE SELECCIÓN MÚLTIPLE (MCQ)

● **Nivel Taxonómico:** Marzano (Según C.E y R.A.) (Niveles 1 al 3)

● **Tiempo por ítem:** 1-2 min aprox.

1. Anatomía de una pregunta de selección múltiple.

COMPONENTE	DESCRIPCIÓN	EJEMPLO
Enunciado (Stem)	Contexto clínico o situacional que enmarca la pregunta	Paciente de 45 años consulta por disnea de 3 días...
Pregunta (Lead-in)	Pregunta directa y específica que anticipa la respuesta esperada	¿Cuál es el diagnóstico?
Distractores (A, B, C)	Alternativas incorrectas pero plausibles, sin pistas involuntarias	A) Neumonía B) TEP C) Insuficiencia cardíaca
Alternativa correcta (D)	Única respuesta correcta, de extensión similar al resto	D) Crisis asmática

2. Aspectos a considerar en su construcción.

✓ Buenas Prácticas

- Use situaciones clínicas o contextuales auténticas en el enunciado, en coherencia con lo revisado en clases.
- Formule una pregunta directa y específica (lead-in).
- Diseñe 4 alternativas de extensión homogénea.
- Asegure que los distractores sean plausibles y sistemáticamente relacionados.
- El enunciado debe permitir responder sin revisar las alternativas.
- Alinea el nivel taxonómico con los resultados de aprendizaje del programa.

✗ Errores a Evitar

- Alternativas "Todas las anteriores" o "Ninguna de las anteriores".
- Pistas involuntarias: alternativa correcta más extensa.
- Preguntas formuladas en negativo ("¿Cuál NO es...?").
- Patrones de respuesta repetitivos (muchas respuestas en "C").
- Errores de concordancia gramatical entre enunciado y alternativas.
- Ítems formato K ("Solo I y II") tienden a ser confusos y poco válidos.

ÍTEMS DE DESARROLLO BREVE

● **Nivel Taxonómico:** Marzano (Según C.E y R.A.) (Nivel 3+)

● **Extensión esperada:** Frase, dato, cálculo, explicación

● **Instrumento para corrección de la pregunta:** Pauta de cotejo o escala de apreciación.

1. ¿Cuándo utilizarlo?

El ítem de desarrollo breve es adecuado cuando se quiere evaluar un aspecto específico y delimitado: definir un concepto, calcular un valor, identificar un criterio diagnóstico o explicar un mecanismo puntual.

La clave es la delimitación: el estudiante debe saber exactamente qué se le pide y hasta dónde debe llegar en su respuesta.

Debe incluir una pauta de cotejo o escala de apreciación.

2. Recomendaciones de construcción

✓ Buenas Prácticas

- Delimite explícitamente la respuesta esperada: número de elementos, criterios o extensión.
- Use verbos taxonómicos precisos: calcule, identifique, defina, señale.
- Proporcione la información contextual suficiente en el enunciado.
- Elabore pauta de corrección antes de aplicar la evaluación.
- Estratifique el puntaje según el nivel taxonómico exigido.

✗ Errores a Evitar

- Consignas amplias como "explique todo lo que sabe sobre..."
- Preguntas que soliciten opinión personal sin criterio evaluativo claro.
- Preguntas sin argumentación en la literatura.
- Ítems que no especifiquen el número de elementos esperados.
- Enunciados que puedan interpretarse de múltiples formas.

3. Ejemplos comparativos

Inadecuado	<i>"Explique todo lo que sabe sobre la diabetes mellitus."</i> Muy amplia, no especifica el foco de la respuesta esperada ni el nivel taxonómico.
Inadecuado	<i>"¿Qué opina sobre el ayuno intermitente?"</i> Personal, no queda claro el nivel taxonómico al cual apunta.
Adecuado	<i>"Mencione tres criterios diagnósticos de DM tipo II según la ADA."</i> Delimita claramente el número de elementos esperados y la fuente referencial.

ÍTEMS DE DESARROLLO EXTENSO

● **Nivel Taxonómico:** Marzano (Según C.E y R.A.) (Nivel 3+)

● **Extensión esperada:** Respuesta integrada, argumentativa o comparativa

● **Instrumento para corrección de la pregunta:** Rúbrica (Obligatoria).

1. ¿Cuándo utilizarlo?

El ítem de desarrollo extenso está diseñado para medir habilidades de orden superior: integración de contenidos, pensamiento crítico, comparación de conceptos o resolución de casos complejos.

Requiere tiempo protegido para su aplicación y corrección, y es indispensable contar con una rúbrica analítica para garantizar la validez y confiabilidad.

2. Recomendaciones de construcción

Delimitación clara	Especifique el número de elementos, criterios o argumentos esperados.
Nivel taxonómico	Use verbos como: compare, evalúe, analice, diseñe, proponga.
Contexto auténtico	Presente casos o situaciones clínicas realistas que exijan integración.
Tiempo suficiente	Calcule el tiempo real de respuesta (resuelva usted y multiplique por 3).
Rúbrica	Diseñe la rúbrica antes de aplicar la evaluación. Defina niveles de logro con descripciones específicas.
Ponderación escalonada	Estratifique el puntaje según el nivel taxonómico: mayor puntaje a mayor complejidad.

✓ Fortalezas

- Promueve integración y pensamiento crítico.
- Revela el razonamiento clínico del estudiante.

✗ Limitaciones

- Alta carga de trabajo en la corrección.
- Requiere más tiempo de aplicación.

3. Ejemplo adecuado

"Compare los efectos fisiológicos del ayuno intermitente y la restricción calórica continua en usuarios con sobrepeso. Para la comparación indique 3 efectos de cada intervención y señale cuál sería más recomendable según la evidencia actual. Fundamente su respuesta."

12 PASOS PARA LA ELABORACIÓN DE ÍTEMS DE PRUEBAS ESCRITAS

1 Alinee los ítems en función al programa de asignatura

Use el programa de asignatura, identifique las competencias declaradas, los resultados de aprendizaje y criterios de evaluación por unidad, como a su vez, los contenidos como brújula o carta de navegación. Determine el nivel taxonómico esperado (Aplicar, analizar, entre otros) y redacte cada ítem en coherencia a aquello.

2 Utilice tabla de especificaciones

Distribuya proporcionalmente los ítems según los contenidos declarados por unidad y que estos posean diversos niveles taxonómicos según lo esperado. De esa forma asegurará validez de contenido y cobertura equilibrada.

3 Redacte ítems claros y significativos

El enunciado (stem) debe contener solo la información necesaria para contextualizar al estudiante. Use situaciones clínicas o contextuales auténticas, evitando el “window dressing” o sobrecarga de información relevante. También puede destacar en negrita información clave.

4 Incluya un lead-in preciso

La pregunta debe formularse en forma clara, sin ambigüedades y anticipar lo que se espera. Evite preguntas negativas o formulaciones vagas. **Recuerde que el enunciado debe contener información suficiente para que el ítem se responda por sí solo, sin tener la necesidad de revisar las alternativas**

5 Diseñe distractores plausibles

Los distractores deben ser creíbles, evitando opciones absurdas, pistas gramaticales o lógicas. Procure que la respuesta correcta no sea la más extensa, ya que esto puede convertirse en una pista involuntaria. Errores de concordancia de género, número o tiempo verbal pueden delatar la opción correcta. Del mismo modo, evite patrones repetitivos (por ejemplo, que muchas respuestas correctas estén en la alternativa “C”) o pistas lógicas que simplifiquen la resolución del ítem sin comprender el contenido.

6 Evite patrones y trampas

Prefiera omitir alternativas como “todas las anteriores”, “ninguna de las anteriores” ni recurrir al absolutismo. Los ítems formato K: “solo I y II”, que poseen múltiples alternativas verdaderas tienden a ser confusos y poco válidos.

7 Redacción de ítems de desarrollo

Evite consignas amplias como: “explique todo lo que sabe sobre...”. Apunte a la claridad en lo esperado: número de elementos, ejemplos requeridos o estructura de la respuesta.

8 Elaboración de ítems de desarrollo breve y extenso según propósito

Los ítems de desarrollo breve apuntan a medir algún aspecto específico, centrado en hechos, procedimientos concretos, cálculos, etc. Los ítems de desarrollo extenso buscan medir integración, comparación, promueven el pensamiento crítico. Ambos exigen desarrollar instrumentos de observación directa para su corrección, como pauta de cotejo o rúbrica, para anexarlas a la evaluación y con ello asegurar confiabilidad.

9 Redacte ítems de desarrollo considerando el tiempo real de respuesta

Ajuste la complejidad, extensión y formato de la pregunta al tiempo disponible. Una buena pregunta en la que se calculó menos tiempo del requerido genera resultados que no permiten pesquisar el nivel de conocimiento del estudiante respecto a la interrogante trazada.
Sugerencia práctica: Como experto, resuelva todo el instrumento y estime el tiempo destinado para luego multiplicarlo por 3. Es una buena regla para ajustar la duración

10 Valide y revise sus ítems

En lo posible aplique un checklist técnico, para asegurar la presencia de: claridad, alineación, pertinencia, plausibilidad, taxonomía, etc. (Ver en página 9: Pauta de revisión técnica y pedagógica)

11 Revise el desempeño de cada ítem luego de aplicarlo

Analice dificultad, discriminación y frecuencia de errores. Así se mejora el banco de preguntas en un futuro y permite retroalimentación oportuna.

12 Optimice sus energías

Imagine sus esfuerzos como un limonero, del cual puede aprovechar la pulpa, la cáscara y sus semillas. En este caso, los ítems desarrollados podría emplearlos en sus clases como aprendizaje activo para interactuar con la audiencia, desarrollar actividades de repaso e incluso como apresto previo a la primera evaluación, para que los estudiantes conozcan las formas, estilos y tipos de ítem que usted desarrolla.

PAUTA DE REVISIÓN TÉCNICA Y PEDAGÓGICA DE PRUEBAS ESCRITAS

Este instrumento tiene como propósito **apoyar el proceso de validación externa de los instrumentos de evaluación tipo prueba escrita** con preguntas de **selección múltiple, desarrollo breve y/o extenso**, asegurando su coherencia curricular, claridad lingüística y pertinencia evaluativa.

Se recomienda revisar el instrumento de evaluación en su totalidad. Para cada criterio, **marque SI o NO** y utilice el campo de observaciones para justificar su decisión, sugerir mejoras o señalar aspectos relevantes que sean objetivos, específicos y concretos.

Este proceso tiene un carácter formativo y busca **fortalecer la calidad de los instrumentos evaluativos**, garantizando que se ajusten a los estándares institucionales y contribuyan efectivamente al aprendizaje significativo de los estudiantes.



SECCIÓN 1: REVISIÓN GENERAL DEL INSTRUMENTO

CRITERIO	SI	NO	OBSERVACIONES / RECOMENDACIONES
Está alineado con los criterios de evaluación.			
El nivel cognitivo de las preguntas corresponde al nivel del curso.			
Se indica la ponderación o puntaje por ítem o sección.			
La prueba explicita instrucciones generales y específicas por sección.			
El tiempo calculado es adecuado para la extensión del instrumento.			

SECCIÓN 2: REVISIÓN POR TIPO DE PREGUNTA

A. Ítems de selección múltiple

CRITERIO	SI	NO	OBSERVACIONES / RECOMENDACIONES
Cada pregunta presenta una sola respuesta correcta.			
No se repiten patrones de respuesta.			
Las extensiones de las alternativas son similares entre ellas.			
Existe concordancia gramatical entre enunciado y opciones.			
Considera 4 opciones.			
La pregunta se responde sin necesidad de mirar las alternativas.			
*Existen preguntas redactadas en negativo.			
*Existen preguntas donde se pregunta por lo falso.			
*Existen alternativas como: Todas las anteriores, Ninguna de las anteriores.			
*Las preguntas y opciones contienen términos ambiguos como: Casi siempre, Generalmente, Rara vez, Nunca, etc.			
Las opciones que contienen cifras están ordenadas en forma ascendente o descendente.			
Se incluye pauta de corrección y/o clave de respuestas.			

Completar sólo los ítems que contenga el instrumento de evaluación.

Nota: Los criterios marcados con * se debe evitar, pues inducen al error por parte de los estudiantes.

B. Ítems de desarrollo breve

CRITERIO	SI	NO	OBSERVACIONES / RECOMENDACIONES
Las preguntas son claras, delimitadas y precisas. Orientan el razonamiento del estudiante.			
Se espera una respuesta concreta, no interpretativa (frase, palabra, cifra, etc.).			
Las preguntas apuntan a un nivel taxonómico desde el nivel 3 en adelante, según taxonomía de Marzano y Kendall.			
Se incluye Pauta de cotejo o Escala de apreciación para revisión de respuestas.			

SECCIÓN 2: REVISIÓN POR TIPO DE PREGUNTA

C. Ítems de desarrollo extenso

CRITERIO	SI	NO	OBSERVACIONES / RECOMENDACIONES
El enunciado está redactado con claridad y proporciona la información suficiente para que el estudiante comprenda lo requerido y responda adecuadamente.			
Las preguntas apuntan a un nivel taxonómico desde el nivel 3 en adelante, según taxonomía de Marzano y Kendall.			
Se incluye Rúbrica para revisión de respuestas.			
Existen preguntas donde se pide la opinión del estudiante y se solicita que fundamente.			

SECCIÓN 3: ACCESIBILIDAD LINGÜÍSTICA Y VISUAL

CRITERIO	SI	NO	OBSERVACIONES / RECOMENDACIONES
El texto utiliza vocabulario claro y adecuado al nivel del curso.			
El texto evita jerga innecesaria o lenguaje excesivamente técnico.			
La prueba es visualmente legible: tamaño de letra, márgenes, espaciado.			
Las imágenes, gráficos o información visual son legibles y están bien referenciados.			
Hay espacio suficiente para que el/la estudiante responda.			
Las instrucciones están redactadas con claridad y sin ambigüedades.			
La estructura visual favorece la orientación del/la estudiante (numeración, títulos por sección).			

Independientemente del tipo de ítem utilizado, todo instrumento de evaluación escrito debe: alinearse con el programa de asignatura, especificar claramente el nivel taxonómico esperado, incluir instrucciones claras, contar con pauta de corrección o rúbrica, y ser validado técnica y pedagógicamente antes de su aplicación.



¿QUÉ TIPO DE ÍTEM USAR Y CUÁNDO?

CRITERIO DE DECISIÓN	SELECCIÓN MÚLTIPLE	DESARROLLO BREVE	DESARROLLO EXTENSO
Nivel taxonómico principal	Recordar / Comprender / Aplicar	Aplicar / Analizar	Analizar / Evaluar / Crear
Tiempo de corrección	Muy bajo (automatizable)	Moderado	Alto
Participación del estudiante	Seleccionar alternativa	Producir respuesta breve	Construir respuesta compleja
Instrumento de corrección	Clave de respuestas	Pauta de cotejo o escala	Rúbrica analítica
Amenaza principal a la validez	Pistas involuntarias	Ambigüedad del enunciado	Subjetividad sin rúbrica
Medición de pensamiento crítico	Limitado	Parcial	Alto potencial
Cobertura de contenidos	Amplia	Moderada	Reducida
Retroalimentación al estudiante	General	Específica	Muy específica

DECISIÓN RÁPIDA



¿Necesita cobertura amplia con corrección rápida?

SELECCIÓN MÚLTIPLE

¿Quiere medir un concepto o cálculo específico?

DESARROLLO BREVE

¿Necesita evidenciar razonamiento, integración o pensamiento crítico?

DESARROLLO EXTENSO

Ningún tipo es universalmente superior: la elección depende del resultado de aprendizaje y del contexto.

BIBLIOGRAFÍA

- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). Taxonomy of educational objectives: Handbook I, cognitive domain. David McKay.
- Case, S. M., & Swanson, D. B. (2002). Constructing written test questions for the basic and clinical sciences (3rd ed.). National Board of Medical Examiners.
- Gottlieb, M., Chan, T. M., Sherbino, J., Yilmaz, Y., Trueger, N. S., & Krzyzaniak, S. (2023). Educator's blueprint: A how-to guide for developing high-quality multiple-choice questions. *AEM Education and Training*, 7(1), e10836. <https://doi.org/10.1002/aet2.10836>
- Marzano, R. J., & Kendall, J. S. (2007). The new taxonomy of educational objectives (2nd ed.). Corwin Press.
- Schuwirth, L. W. T., & van der Vleuten, C. P. M. (2011). Programmatic assessment: From assessment of learning to assessment for learning. *Medical Teacher*, 33(6), 478–485. <https://doi.org/10.3109/0142159X.2011.565828>
- Unidad de Educación en Ciencias de la Salud, Universidad Autónoma de Chile. (2026). Doce pasos para la elaboración de ítems de pruebas escritas. Universidad Autónoma de Chile.
- Unidad de Educación en Ciencias de la Salud, Universidad Autónoma de Chile. (2026). Pauta de revisión técnica y pedagógica de pruebas escritas. Universidad Autónoma de Chile.