

Inventario para autoevaluar el pensamiento crítico en estudiantes de educación superior

César Augusto Cardena Ojeda

<https://orcid.org/0000-0001-6170-2223>
Escuela Normal de Ticul
cesarcardenaojeda@enty.edu.mx

Resumen

El pensamiento crítico constituye uno de los rasgos del perfil de egreso más valorados en los programas de formación de todas las áreas y niveles educativos, debido a su incidencia en la calidad del pensamiento en los ámbitos escolar, profesional y personal. A partir del reconocimiento de la importancia de esta capacidad cognitiva, se planteó como objetivo diseñar un instrumento válido que permitiera autoevaluar, de manera práctica, el nivel de pensamiento crítico en estudiantes de educación superior. Para ello, se realizó un estudio cuantitativo de diseño instrumental, siguiendo los lineamientos de la Comisión Internacional del Test (CIT) y otros referentes especializados. Los resultados evidenciaron, en las pruebas de hipótesis, valores psicométricos satisfactorios en términos de confiabilidad y validez, además de confirmar la ausencia de sesgos asociados a segmentos específicos de la población evaluada. En consecuencia, el instrumento puede considerarse adecuado para su aplicación en poblaciones más amplias de estudiantes de educación superior.

Palabras clave

Educación superior; inventario; pensamiento crítico; psicometría; validación.

Recibido: 20/01/2026 | Enviado a pares: 06/02/2026 | Aceptado por pares: 10/02/2026 | Aprobado: 08/05/2026

DOI: 10.5294/edu.2026.29.2.8

Para citar este artículo / To reference this article / Para citar este artigo

Cardena, C. (2026). Inventario para autoevaluar el pensamiento crítico en estudiantes de educación superior. *Educación y Educadores*, 29(2), e2928. <https://doi.org/10.5294/edu.2026.29.2.8>

Self-Assessment Inventory for Critical Thinking in Higher Education Students

Abstract

Critical thinking constitutes one of the most highly valued attributes in graduate profiles across educational programs at all disciplines and levels, due to its impact on the quality of thinking in academic, professional, and personal contexts. Recognizing the importance of this cognitive capacity, the objective of this study was to design a valid instrument that would enable the practical self-assessment of critical thinking levels among higher education students. To achieve this, a quantitative study with an instrumental design was conducted, following the guidelines established by the International Test Commission (ITC) and other specialized references. The results demonstrated satisfactory psychometric properties in hypothesis testing, particularly in terms of reliability and validity, while also confirming the absence of bias associated with specific segments of the assessed population. Consequently, the instrument can be considered suitable for application in broader populations of higher education students.

Keywords

Higher education; inventory; critical thinking; psychometrics; validation.

Inventário para autoavaliar o pensamento crítico em estudantes do ensino superior

Resumo

O pensamento crítico é uma das características mais valorizadas do perfil do egresso de formação de todas as áreas e níveis de ensino, devido ao seu impacto na qualidade do pensamento nas esferas acadêmica, profissional e pessoal. Com base no reconhecimento da importância dessa capacidade cognitiva, o objetivo deste estudo foi desenvolver um instrumento válido que permitisse autoavaliar, de forma prática, o nível de pensamento crítico em estudantes do ensino superior. Para isso, foi realizado um estudo quantitativo de delineamento instrumental, seguindo as diretrizes da Comissão Internacional de Testes e outros referenciais especializados. Os resultados indicaram, nos testes de hipóteses, valores psicométricos satisfatórios quanto à confiabilidade e à validade, além de confirmarem a ausência de vieses associados a segmentos específicos da população avaliada. Consequentemente, o instrumento pode ser considerado adequado para a aplicação em populações mais amplas de estudantes do ensino superior.

Palavras-chave

Ensino superior; inventário; pensamento crítico; psicometria; validação.

Introducción

En una institución pública de educación superior se identificó de manera sistemática que los estudiantes presentaban deficiencias evidentes en su pensamiento crítico. En consonancia con ello, se observó que mostraban limitaciones para formular juicios elementales sobre los temas abordados en clase, aun cuando estos presentaban un bajo nivel de complejidad. Asimismo, evidenciaban dificultades para analizar información textual para extraer de ella inferencias que trascendieran la lectura literal. De igual manera, tendían a aceptar sin cuestionamientos afirmaciones procedentes de fuentes teóricas.

También, mostraban dificultades para reformular o reconstruir ideas conocidas con el propósito de proponer nuevas vías de análisis o evaluación. En consecuencia, su interacción intelectual con los contenidos curriculares tendía a limitarse a la apropiación acrítica de las posturas presentadas en los textos académicos o en el discurso de los profesores.

A partir de la evidencia observada, se concluyó que existía una limitación importante en el pensamiento crítico de los estudiantes. Sin embargo, esta valoración tenía un carácter general y carecía de una base formal o técnica que la sustentara, pues se fundamentaba en la observación y en la reflexión sobre los comportamientos descritos. Por ello, se consideró de interés institucional contar con un instrumento práctico que permitiera estimar la magnitud y gravedad del problema identificado.

Para ello, se realizó una revisión exhaustiva de fuentes académicas en los ámbitos de la psicología y la educación con el fin de localizar un instrumento adecuado para esta tarea. Sin embargo, los instrumentos identificados carecían de alguna de las condiciones necesarias para su aplicación directa en la población de interés, tales como vigencia de sus

contenidos, la adecuación del lenguaje a la población objetivo e, incluso, un costo asequible.

Al respecto, puede mencionarse el trabajo de revisión sistemática realizado por Ossa *et al.* (2017), quienes analizaron de manera descriptiva y comparativa 97 instrumentos de medición del pensamiento crítico. Los autores reportaron fundamentos teóricos, evidencias de confiabilidad y validez, dimensiones y otros indicadores métricos relacionados con el lenguaje, la estructura y el tipo de contenido de cada instrumento. Este análisis permitió identificar divergencias, convergencias, potencialidades y limitaciones metodológicas entre los materiales examinados.

Entre los hallazgos de este trabajo destacó que la mayoría de los instrumentos habían sido diseñados en el idioma inglés y no contaban con una adaptación válida al español para la población mexicana. Asimismo, una proporción considerable de estos materiales estaban dirigidos a poblaciones distintas de la educación superior, como niños de entre nueve y doce años, por lo que su utilización no resultaba viable con estudiantes universitarios.

A partir de esta información, se realizó una revisión de instrumentos orientados específicamente a la medición del pensamiento crítico en educación superior, con el propósito de identificar sus elementos constitutivos y enfoques métricos, así como verificar si existía algún instrumento que reuniera las características y propiedades métricas necesarias para su utilización en el diagnóstico de la población de interés.

Esta revisión permitió identificar instrumentos desarrollados desde 1968, así como diversas actualizaciones de una misma prueba, lo que evidencia el interés sostenido que el estudio del pensamiento crítico ha suscitado durante varias décadas.

En la tabla 1, se presentan los resultados de esta revisión, organizados en orden cronológico.

Tabla 1. Principales instrumentos de medición del pensamiento crítico en educación superior

Instrumento	Autor(es)	Contenido	Dimensiones del pensamiento crítico
Watson-Glaser Critical Thinking Appraisal (WGCTA).	Watson y Glaser (1980).	80 reactivos de opción múltiple.	Inferencia, deducción, interpretación y evaluación de argumentos.
Cornell Critical Thinking Test (CCTT).	Ennis y Millman (1985).	71 reactivos de opción múltiple en la forma X.	Inducción, deducción, credibilidad, identificación de supuestos.
California Critical Thinking Disposition Inventory (CCTDI).	Facione (2000).	75 reactivos tipo Likert.	Curiosidad, mente abierta, sistematicidad, búsqueda de la verdad.
Halpern Critical Thinking Assessment (HCTA).	Halpern (2003).	25 situaciones problemáticas con opciones de respuesta mixta: abierta y de opción múltiple.	Razonamiento verbal, análisis de argumentos, toma de decisiones.
Halpern Critical Thinking Assessment using Every Day Situation (HCTAES).	Halpern (2006).	25 situaciones problemáticas con opciones de respuesta dual: abierta y cerrada.	Razonamiento verbal, análisis de argumentos, toma de decisiones y resolución de problemas, y razonamiento probabilístico en contextos de incertidumbre.
Test de Pensamiento Crítico (TPC).	Miranda (2003).	40 reactivos con opciones de respuesta múltiple.	Análisis, interpretación, evaluación.
Prueba de Pensamiento Crítico PENCRI-SAL.	Saiz y Rivas (2008).	35 reactivos de respuesta abierta.	Razonamiento, toma de decisiones, resolución de problemas.
HElghten Critical Thinking Assessment.	Lui <i>et al.</i> (2016).	40 reactivos de opción múltiple en diversos escenarios.	Evaluación de evidencia, análisis, razonamiento.

Fuente: elaboración propia.

La tabla 1 permite observar la variación metodológica en la recopilación de datos y en las interpretaciones sobre las dimensiones constituyentes del objeto de medida a lo largo del tiempo. En las primeras obras predominó el uso de reactivos de opción múltiple, orientados a facilitar una aplicación estandarizada. Posteriormente, se incorporaron progresivamente reactivos más abiertos, centrados en la solución de situaciones, con el fin de rastrear procesos cognitivos más complejos y contextualizados. Este cambio implicó también un incremento en las exigencias de análisis para la revisión de las respuestas de los evaluados.

Poco después se introdujeron formatos de respuesta más elaborados, como reactivos abiertos y

situaciones que buscan aproximarse a escenarios reales, lo que favoreció la evaluación de la transferencia del pensamiento crítico a contextos prácticos. También, se observa una ampliación en la comprensión del constructo de medida, al incorporar no solo habilidades, sino también disposiciones o actitudes hacia el pensamiento reflexivo, lo cual se evidencia, por ejemplo, en el cuestionario *California Critical Thinking Disposition Inventory* (CCTDI).

Esta adición evidencia un avance significativo en la conceptualización del pensamiento crítico, al reconocer que este no depende exclusivamente de habilidades cognitivas, sino también de factores motivacionales y disposicionales. En términos

generales, la medición del pensamiento crítico ha transitado desde enfoques centrados en rasgos aislados hacia modelos más integrados, aunque las dimensiones constitutivas del constructo han mantenido como eje central el análisis, el cuestionamiento y la argumentación.

Entre estos materiales, se consideró importante realizar una revisión más profunda al trabajo de Liu *et al.* (2016), quienes desarrollaron la prueba *HElghten Critical Thinking Test* (HCTT) para el Educational Testing Services de los EE. UU. Este es uno de los últimos instrumentos estandarizados más recientes orientados a evaluar el pensamiento crítico en educación superior, por lo que ha convertido en un referente obligado para la adaptación o el diseño de nuevos instrumentos dirigidos a una población del mismo nivel educativo, pero en contextos diferentes.

En este trabajo se estructuró una prueba con dos dimensiones: la analítica y la sintética, y fue integrada por reactivos de distintos formatos, como selección múltiple, respuesta múltiple y de tareas basadas en situaciones. El modelo preliminar fue pilotado en 3036 estudiantes de 35 instituciones de educación superior con programas de dos años y cuatro años de duración.

Los autores reportaron diversos análisis psicométricos sobre los datos obtenidos del pilotaje de la prueba, destacando que la estructura interna mostró una configuración unidimensional, ya que un único factor explicó hasta el 67 % de la varianza en las respuestas a los reactivos. Sin embargo, se mantuvo la consideración teórica de dos dimensiones (analítica y sintética) propuesta en el marco teórico del estudio.

El coeficiente de confiabilidad también mostró niveles satisfactorios, con valores elevados de estabilidad entre .90 y .99. Asimismo, se incorporaron modelos basados en la Teoría de Respuesta al Ítem (TRI) para analizar las variaciones en la dificultad y la ca-

pacidad discriminativa de los reactivos, lo que aportó mayor precisión a las mediciones en el continuo.

Además, se detectó sensibilidad al progreso del estudiante a lo largo de los distintos grados escolares de su programa educativo, aunque con puntajes discretos que no se consideraron determinantes. Por ello, se determinó que el instrumento no requería ajustes en los baremos de calificación en relación con este criterio.

Al final de su estudio, los autores concluyeron que el instrumento construido presenta evidencias de validez y confiabilidad para contribuir a la mejora de la acreditación y el desempeño de los estudiantes. No obstante, señalan que las evidencias de validez obtenidas son de carácter preliminar, por lo que recomiendan fortalecer su análisis mediante la ampliación de su aplicación en estudios longitudinales.

En este trabajo se presentan mecanismos de análisis psicométrico que pueden ser considerados en proyectos orientados a replicar la construcción de instrumentos con objetivos o variables similares. Por otra parte, a partir de las ideas expuestas en su revisión de literatura y en sus conclusiones, se corrobora la importancia del abordaje del pensamiento crítico en el acompañamiento integral de estudiantes de educación superior, así como la posibilidad de diseñar instrumentos técnicamente sólidos para su evaluación. No obstante, los autores advierten que estas tareas deben complementarse con acciones institucionales que permitan contrastar los resultados obtenidos mediante recursos cuantitativos (Liu *et al.*, 2014).

Sin embargo, las opciones de instrumento disponibles, incluido el caso más reciente revisado, carecían de al menos un atributo técnico básico para su aplicación, ya fuera el idioma, el nivel educativo de uso o el formato de presentación. En algunos casos, además, el instrumento consistía en una serie de situaciones por resolver, lo cual representaba una dificultad importante frente a la intención de

contar con un instrumento práctico que pudiera ser administrado de manera censal y, al mismo tiempo, revisado de forma eficiente y con mínima carga interpretativa.

A partir de esta revisión documental, se concluyó la conveniencia de abordar la medición del pensamiento crítico desde la autoevaluación, es decir, desde el autorreporte de la percepción del estudiante sobre su dominio de esta competencia, con el fin de propiciar un ejercicio reflexivo inicial que permita posteriormente desarrollar otras estrategias e instrumentos complementarios, integrados en un sistema de evaluación que mantenga su precisión y, al mismo tiempo, su carácter práctico.

Ante este panorama, se formuló la siguiente pregunta de investigación: ¿cómo medir la percepción de los estudiantes sobre su nivel de pensamiento crítico mediante un instrumento práctico, válido y confiable?

Justificación de la investigación

Este tema se consideró relevante de investigar, dado que el pensamiento crítico constituye un elemento distintivo del enfoque oficial de la licenciatura en educación impartida por la institución en la que se identificó el problema. Esta orientación se encuentra plasmada en el plan de estudios vigente, que expresa la intención de fomentar en las aulas “pensamiento crítico, científico y creativo para innovar la intervención pedagógica” (Secretaría de Educación Pública [SEP], 2022a, p. 1).

Asimismo, en el mismo documento oficial citado, específicamente en el apartado de rasgos generales del perfil de egreso, se establece que el estudiante que concluya su formación y obtenga su titulación deberá haber desarrollado un “pensamiento reflexivo, crítico, creativo y sistémico”. Esta formulación se reproduce de manera íntegra en el noveno rasgo de los “dominios del saber: saber ser y estar, saber conocer y saber y hacer” (SEP, 2022a, pp. 8-9) del mismo perfil general.

Por su parte, en los rasgos profesionales del perfil de egreso se establece que, al concluir su formación, el estudiante deberá ser capaz de realizar “un análisis crítico y transformador de la propia práctica para generar una docencia reflexiva que replantea el avance y progreso de los aprendizajes de niños y niñas” (SEP, 2022b, p. 14).

El desarrollo del pensamiento crítico ocupa un lugar destacado en los intereses formativos de la carrera, lo que se evidencia en su mención tanto en los rasgos generales como en los profesionales del perfil de egreso, situación que no ocurre con otras aspiraciones curriculares.

Ante ello, el desarrollo del pensamiento crítico puede constituirse en un eje específico de los programas de tutoría académica institucionales, especialmente en las escuelas normales, donde este servicio de acompañamiento aún carece de directrices normadas y de experiencia suficientes para definir las a partir de diagnósticos pertinentes (Ponce *et al.*, 2017).

Por otra parte, desde una perspectiva pedagógica, el pensamiento crítico se considera una capacidad fundamental para valorar la calidad de los aprendizajes construidos en los entornos escolares, así como la capacidad de autorregulación del aprendizaje. Su relevancia resulta especialmente evidente en la educación superior, donde la demanda de estas facultades es permanente y aumenta en complejidad a medida que avanza la trayectoria formativa (Facione, 1990; Vendrell y Rodríguez, 2020).

En la educación superior, el pensamiento crítico se asocia estrechamente con el desarrollo de la actividad investigadora, ya que se considera una habilidad cognitiva indispensable para comprender y aplicar los principios de rigor y racionalidad que la caracterizan. Esta exigencia resulta especialmente relevante si se considera que la investigación constituye, por sí misma, un desafío para los estudiantes (Benzanilla *et al.*, 2021; Mackay *et al.*, 2018; Vendrell y Rodríguez, 2020).

Al respecto, Lincoln y Kearney (2019) reflexionan sobre cómo la actividad pedagógica en la educación superior se encuentra ineludiblemente comprometida con la promoción del pensamiento crítico, así como sobre el papel fundamental que desempeñan las instituciones en esta tarea. Aunque reconozcan que se trata de una capacidad transversal que no requiera necesariamente un espacio curricular específico para su desarrollo, enfatizan la necesidad de promoverla de manera constante en todos los espacios curriculares. Asimismo, señalan que, debido a su complejidad cognitiva, su desarrollo solo puede alcanzarse mediante estrategias intencionadas y sistematizadas que privilegian el análisis, el cuestionamiento y el juicio argumentado.

Finalmente, se reconoce que este tipo de pensamiento constituye una herramienta necesaria para realizar ejercicios analíticos que posibiliten una comprensión más objetiva del mundo y sus fenómenos, en contraste con aquella que se obtiene mediante la aplicación del dogma, el mantenimiento del pensamiento supersticioso o la debilidad crítica (De la Borbolla, 2017, 2019; Schick y Vaughn, 1994/2024).

Objetivo

Para delimitar el alcance del problema identificado, se definió el siguiente objetivo: diseñar un instrumento válido para la autoevaluación el nivel de pensamiento crítico de educación superior.

Revisión de la literatura

Profundización a la noción de pensamiento crítico

En opinión de Vendrell (2025), profesora investigadora de la Universidad Complutense de Madrid, hasta la fecha no existe consenso sobre una definición unívoca del pensamiento crítico y la mayoría de las definiciones en el campo educativo resultan difusas o excesivamente abstractas para generar indicadores objetivos de reconocimiento, lo que se dificulta su orientación hacia un enfoque práctico (Vendrell, 2025).

Grosso modo, la investigadora ha identificado dos grandes grupos de conceptualización del pensamiento crítico en educación. Por un lado, se encuentran los especificistas, quienes sostienen que el pensamiento crítico no es una habilidad transferible de manera automática entre campos disciplinares, sino una capacidad situada que se configura según la disciplina en la que se desarrolle, adecuándose al lenguaje, métodos y criterios. En este sentido, su ejercicio se define a partir de marcos normativos propios de cada área.

Por otro lado, los generalistas consideran el pensamiento crítico como una competencia transversal, con manifestaciones transferibles como el análisis de ideas y argumentos y la evaluación de evidencias, por lo que puede aplicarse en distintos campos disciplinares.

Además de esta bifurcación conceptual y operacional, en los últimos años los sistemas escolares han incorporado dimensiones éticas, culturales y cívicas en sus enfoques sobre el pensamiento crítico, lo que ha generado nuevas divergencias entre quienes sostienen que debe mantenerse su carácter de habilidad cognitiva y quienes argumentan que requiere un componente actitudinal para completarse, especialmente en el ámbito escolar, donde la formación no se orienta únicamente a la futura inserción laboral, sino también a la convivencia en sociedad.

Así, al reflexionar sobre sus consideraciones éticas desde una postura sociocrítica, este tipo de pensamiento adquiere un potencia transformador, al orientar al sujeto hacia el cuestionamiento, la problematización y el posicionamiento frente a la realidad, configurando una posibilidad emancipadora que trasciende el entorno escolar y la focalización exclusiva en el desempeño académico.

En efecto, quien ejerce una crítica sistemática y rigurosa de la realidad, rompe con la aquiescencia y desarrolla disposición a denunciar o transformar aquellas condiciones con las que se siente inconforme o insatisfecho. En este caso, la pulsión transfor-

madura no proviene de un rasgo individual azaroso de la personalidad, sino que emerge de una exigencia ética a la búsqueda de alternativas frente al contexto.

Rancière (1987/2022) plantea esta idea al señalar que la pedagogía tradicional, centrada en la explicación, funciona como mecanismo de subordinación al docente experto. Frente a ello, propone una educación emancipadora en la que el sujeto descubre que puede aprender por sí mismo al verificar su propia capacidad. Aunque reconoce la diversidad entre los individuos de un grupo, sostiene que esta no responde a diferencias de inteligencia, sino a distintos momentos de un mismo trayecto formativo, por el cual todos pueden avanzar si se establecen condiciones de igualdad y estímulo.

De conformidad con esta proposición, se reconoce que no todo pensamiento sofisticado es en sí mismo emancipador si no integra una crítica orientada por valores de responsabilidad social. En este sentido, se evidencia la necesidad de promover disposiciones como la apertura al diálogo y el análisis de posturas contrarias, no con el propósito de detectar debilidades argumentativas en los demás, sino con una disposición genuina al cambio y a la acción.

Al respecto, exponentes de la psicología cognitiva como Halpern (2014) y Redhana *et al.* (2017), señalan de manera enfática que el pensamiento crítico emerge del “querer” encontrar información precisa sobre la realidad, la cual suele estar oculta entre la vasta información disponible. Este proceso implica un sobreesfuerzo y una dedicación que solo se sostienen a partir del convencimiento de que el resultado será valioso por sí mismo. Si bien se reconoce que debe prevalecer la reflexión racional, se enfatiza la importancia de su base actitudinal. Así, el pensamiento crítico implica diversos elementos que funcionan simultáneamente como componentes funcionales y como propiciadores, es decir, como motivadores.

Por su parte, Robles (2019) considera que el pensamiento crítico es una habilidad cognitiva superior

compuesta por dimensiones racionales y actitudinales que, en términos generales, permite analizar y formular ideas con base en criterios argumentativos.

En tanto habilidad, no se desarrolla de manera natural o espontánea ni aparece automáticamente con el paso del tiempo o con la maduración cognitiva cronológica, como ocurre con otros procesos cognitivos tales como la memoria, la percepción de profundidad, el autorreconocimiento, el sentido de permanencia del objeto o la coordinación visomotora, los cuales se desarrollan progresivamente en distintas etapas del desarrollo.

Por el contrario, el pensamiento crítico debe ser aprendido y ejercitado, especialmente en los tiempos actuales, caracterizados por la abundancia de distracciones que interfieren con el estado de concentración necesario para la elaboración de razonamientos, así como por la disponibilidad de herramientas que permiten acceder de manera inmediata y sencilla a información que satisface cualquier cuestionamiento.

Al respecto, Bouygues (s.f.), presidente de la Fundación Reboot¹, organización internacional sin fines de lucro con sede en París, Francia, dedicada a la promoción del pensamiento crítico en profesores y padres de familia, enfatiza que el pensamiento crítico es una habilidad cada vez más infrecuente. Esto se debe a que el individuo contemporáneo, independientemente de su edad u ocupación, no se ve obligado a recurrir a sus propios recursos intelectuales para resolver dudas o problemas, ya que dispone de dispositivos tecnológicos que ofrecen respuestas inmediatas.

La autora añade que, en la actualidad, predomina el llamado pensamiento selectivo, consistente en el reforzamiento de las propias opiniones, es decir, una forma de sesgo de confirmación que resulta cómodo y gratificante, pero que limita la diversificación del pensamiento (Harvard Business Review [HBR], 2023).

¹ <https://reboot-foundation.org>

Este tipo de pensamiento tiene una base cognitiva formal que parte del cuestionamiento, es decir, de la formulación de preguntas que desafían la conformidad con las ideas, por lógicas, convincentes o complacientes que estas parezcan. En su nivel más concreto, la crítica se origina en dicho cuestionamiento. En este sentido, Pohlmann y Neethi (s.f.), directivos de Mu Sigma², una de las empresas más importantes a nivel mundial en analítica de datos y toma de decisiones, señalan que el pensamiento crítico se construye a partir de la elaboración de cuatro tipos de preguntas ante problemas complejos.

El primer tipo de estas preguntas, las clarificadoras, permite corroborar que lo leído, escuchado o formulado inicialmente ha sido comprendido con claridad. Por su parte, las preguntas adyacentes exploran aspectos relacionados con el tema de reflexión, con el fin de ampliar su comprensión o distinguirlo claramente de otros temas cercanos, pero distintos del objeto de interés.

Las preguntas “de embudo” revisan detalles minuciosos que puedan sustentar explicaciones causales, mientras que las preguntas “de elevación” revisan de manera general las conclusiones, con el fin de propiciar la corrección de los resultados, la revisión de su coherencia y de su adecuación al panorama general (HBR, 2023).

Otro aspecto fundamental por considerar es el incremento de la eficacia y de la autonomía crítica. Así, se espera que el ejercicio fortalezca la calidad de la respuesta cognitiva ante los problemas que se enfrentan, permitiendo resolverlos con mayor precisión y rapidez, o abordar situaciones progresivamente más complejas (Vincent, 2021).

Algunos autores, como los citados por García y López (2020) en el texto *Pensamiento crítico para el aprendizaje*, consideran que este pensamiento consiste en la aplicación integrada de diversas operaciones cognitivas para examinar planteamientos concretos; en este caso. En este enfoque de carácter

pedagógico, se propone una noción pragmática y centrada en los desempeños del estudiante, entre los que destacan el involucramiento en la tarea de análisis, el diálogo con fuentes de distintas perspectivas, la experimentación y la conceptualización.

Estas conductas, en suma, son las que evidencian un comportamiento crítico para los siete ensayistas que colaboraron en la obra. Por otra parte, se debe reconocer que las discrepancias, aunque escasas, tienen implicaciones importantes en la definición de la propuesta.

Al realizar una revisión comparativa de las concepciones del tema, se identifican como componentes comunes las habilidades cognitivas, la fundamentación en la evidencia y la reflexión o inferencia racional. A continuación, se presenta la tabla 2 en la que se incluye, en una de sus entradas, a los autores revisados y, en la columna de contraste, los elementos identificados en sus principales publicaciones sobre pensamiento crítico, con el fin de reconocer coincidencias y discrepancias.

Según lo revisado, el aspecto con mayor discrepancia entre las fuentes corresponde a una disposición, es decir, a un componente actitudinal. Al respecto, se observa con claridad que la incorporación de aspectos actitudinales es un desarrollo más reciente, en contraste con el énfasis en las habilidades cognitivas.

Por otra parte, resulta interesante observar que la resolución de problemas se considera prescindible únicamente por De la Borbolla (2019), cuya formación filosófica le permite abordar construcciones teóricas sin la necesidad de traducirlas en operaciones aplicadas o producciones concretas.

Profundizando en los estudios antecedentes sobre la evaluación del pensamiento crítico, Betancourt *et al.* (2017) desarrollaron una investigación cuantitativa con el objetivo de medir las habilidades de 141 estudiantes universitarios seleccionados intencionalmente de la carrera de Derecho de una universidad de Chile. Para abordar su objetivo, se

2 <https://www.mu-sigma.com>

Tabla 2. Revisión de elementos del pensamiento crítico en autores principales

Autor o autores	Elemento					
	Análisis	Inferencia	Disposición a la autocorrección	Argumentación	Evaluación o duda	Solución de problemas
Ennis <i>et al.</i> (1964).	x	x		x		
Stein <i>et al.</i> (2003).	x	x		x	x	x
Da Dalt y Difabio (2007).	x	x		x	x	x
Hatcher (2011).	x	x		x	x	x
Ossa <i>et al.</i> (2017).	x			x	x	x
Haber (2020).	x	x	x	x	x	x
De la Borbolla (2019).	x	x	x	x	x	
Vendrell y Rodríguez (2020).	x	x	x	x	x	x
Escobar (2022).	x	x	x	x	x	x
Arancibia <i>et al.</i> (2022).	x	x		x	x	x

Fuente: elaboración propia.

empleó la prueba HCTAES, desarrollada por Halpern (2006), en una versión en español validada específicamente para el contexto chileno y para este estudio.

Este instrumento se compone de reactivos que evalúan el razonamiento verbal, la argumentación, la formulación de hipótesis y la toma de decisiones, a través de 25 situaciones de doble formato. Es estas, además de construir una respuesta abierta para cada caso, el evaluado debe seleccionar otra respuesta entre opciones cerradas.

A partir de ello, se busca reconocer la *memoria de recuerdo libre*, que permite construir de manera autónoma y espontánea una respuesta que pone en juego habilidades críticas, de la *memoria de reconocimiento*, que se activa al seleccionar una opción entre las alternativas previamente estructuradas (Halpern, 2006).

En los resultados se encontró que la población evaluada presentó bajos niveles de pensamiento crítico, con un mejor desempeño en las respuestas

cerradas, en comparación con las preguntas de respuesta abierta.

En cuanto a las dimensiones, el análisis básico y la interpretación mostraron los mejores puntajes, mientras que la inferencia y la autorregulación, consideradas habilidades de alta demanda cognitiva, fueron las de menor desempeño. En términos generales, se observó que los estudiantes presentaron dificultades para fundamentar sus argumentaciones, evaluar fuentes y tomar decisiones basadas en evidencias.

A partir de estos hallazgos, los autores concluyeron que existe un problema relevante en la capacidad de pensamiento crítico de los estudiantes de educación superior participantes en el estudio. En la discusión teórica, se estableció que este tipo de pensamiento no se desarrolla de manera automática conforme se avanza en el trayecto escolar, sino que requiere una intencionalidad pedagógica que, en muchos casos, no está presente en las aulas. En este sentido, el estudio evidenció una brecha entre las

aspiraciones formativas del currículo de educación superior y las habilidades efectivamente demostradas por los estudiantes en su desempeño cotidiano.

Método

Enfoque de la investigación

El enfoque de investigación de este estudio fue cuantitativo, en tanto el análisis de los datos, así como la toma de decisiones y la construcción de conclusiones, se basaron en criterios estadísticos preestablecidos por consenso de expertos o en fórmulas derivadas del proceso analizado (Maksimovic y Evtimov, 2023).

Tipo de investigación

De manera más específica, este trabajo se puede clasificar como de tipo instrumental, ya que se orienta al desarrollo de pruebas, lo que implica desde su diseño hasta la evaluación de sus propiedades psicométricas, tales como la validez y la confiabilidad (Maksimovic y Evtimov, 2023).

Por otra parte, desde una categoría más amplia, corresponde a una investigación descriptiva, ya que se orienta a la identificación y análisis sin manipulación ni experimentación. En su lugar, se centra en el examen observacional e informativo de la estructura del contenido del instrumento y en el análisis del comportamiento de los datos obtenidos de su aplicación (Namakforoosh, 2014; Pimienta, De la Orden y Estrada, 2018).

Población

La población se conformó por el total de estudiantes matriculados en cuatro instituciones de una misma entidad federativa. Tres de ellas eran de sostenimiento público e impartían el mismo programa oficial de estudios. Por su parte, la institución de sostenimiento particular incluyó estudiantes de tres programas académicos: uno correspondía al mismo programa ofrecido por las instituciones públicas; otro era análogo y compartía el 50 % de su contenido curricular con dicho programa y, el tercero, pertenecía a un área de estudio o denominación temática diferente.

Este último caso se incorporó con el propósito de realizar análisis diferenciales del comportamiento del instrumento en función del tipo de institución, y así sustentar su posible utilización en poblaciones similares a la considerada en la investigación. Asimismo, permitió efectuar análisis estadísticos orientados a respaldar la recomendación de generalizar su uso.

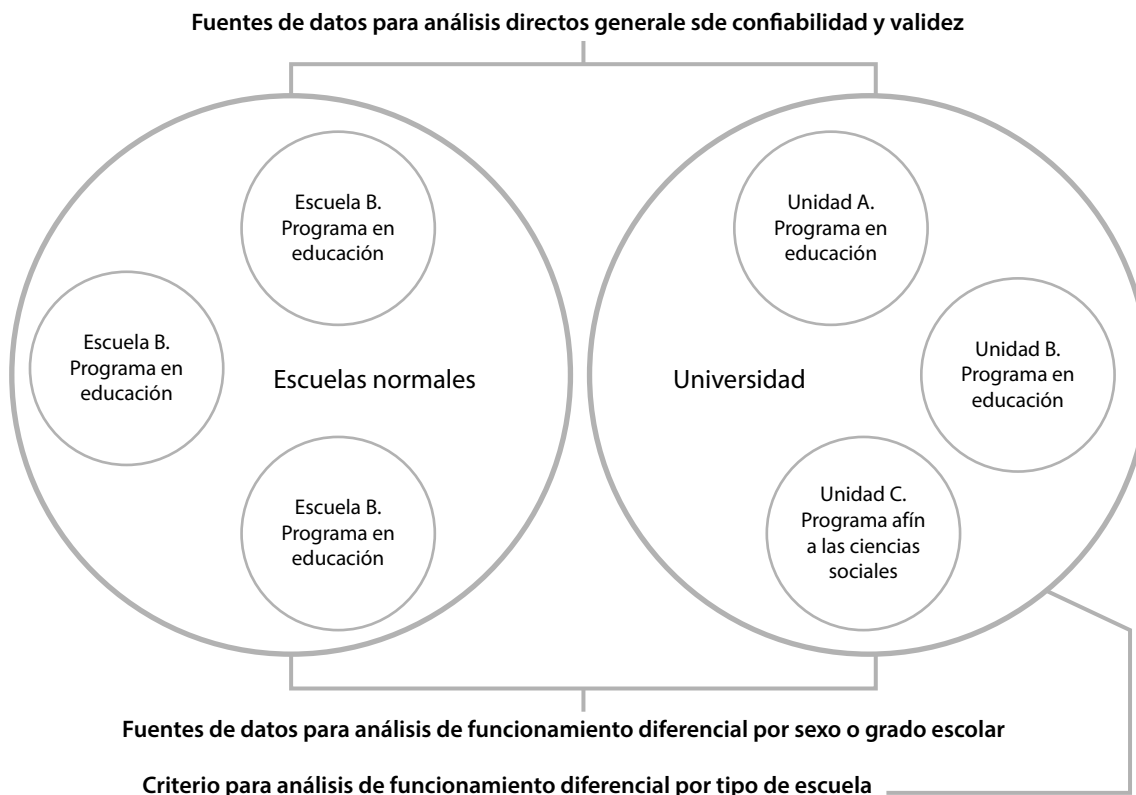
La figura 1 muestra la composición de la población estudiada, con el fin de identificar sus unidades diferenciadas y los análisis que fue posible realizar a partir de los datos obtenidos en cada caso

Para facilitar la comprensión de la composición de la población, se presenta la tabla 3.

Tabla 3. Composición de la población por número de sujetos integrantes

Escuela	Sostenimiento	Programa educativo impartido	Sujetos encuestados	% de los encuestados respecto del total
A	Público	Licenciatura en Educación Primaria	390	99
B	Público	Licenciatura en Educación Primaria	302	97
C	Público	Licenciatura en Educación Primaria	70	87
D	Particular	Licenciatura en Educación Primaria	103	91
		Licenciatura en Educación Preescolar	230	93
		Otro (educación especial, psicología educativa)	363	96
		Total:	1 458	

Fuente: elaboración propia.

Figura 1. Composición de la población por unidades de análisis

Fuente: elaboración propia.

Como puede observarse, en todas las unidades de análisis se logró encuestar al menos el 87 % del total de los sujetos elegibles, lo que permite sustentar el carácter censal de la población estudiada.

Procedimiento para el acopio de datos

Para la recolección de los datos, se contactó personalmente a los directores de las instituciones en las que se pretendía aplicar una versión preliminar del instrumento, previamente sometido a revisión crítica por cinco expertos. En todos los casos, se explicó el propósito del estudio y se aseguró el cumplimiento de las normas éticas relacionadas con el tratamiento y la difusión de los resultados.

Una vez obtenida la autorización para la aplicación del instrumento, se realizó una entrevista

telefónica individual con cada uno de los coordinadores académicos de las instituciones participantes, con el fin de acordar los detalles relacionados con la distribución del enlace de acceso al formulario y el plazo previsto para su respuesta, de acuerdo con las disposiciones organizativas de cada institución. Asimismo, los coordinadores se comprometieron a mantener comunicación permanente con el investigador responsable para monitorear el avance y resolver dudas que pudieran surgir.

Plan de análisis de datos

Una vez aprobado por los jueces expertos que lo revisaron, el instrumento se aplicó mediante un formulario en línea durante un periodo de cuatro semanas. Los datos obtenidos durante su aplicación

fueron analizados mediante diversos procedimientos, siguiendo principalmente las orientaciones de Tornimbeni *et al.* (2013) y de las directrices de la Comisión Internacional del Test (Leong *et al.*, 2020).

El primer análisis consistió en revisar la concordancia entre las calificaciones independientes emitidas por jueces expertos durante la evaluación de la validez de contenido del instrumento. Para ello, se utilizó una guía de evaluación que solicitaba la asignación de una calificación numérica de cada reactivo, así como un comentario justificativo para recomendar su conservación, modificación o eliminación (Klaus, 2023; Medrano y Pérez, 2023, p. 76).

Las puntuaciones asignadas por los jueces se utilizaron para calcular el coeficiente Kappa de Fleiss, que estima el grado de concordancia entre las evaluaciones de tres o más jueces cuando se emplean escalas de medición nominal u ordinal. La tabla 4 presenta los criterios de interpretación propuestos por Altman (1991, como se cita en Torres y Perera, 2009).

Tabla 4. Interpretación del coeficiente de concordancia de Fleiss

Valor	Fuerza de concordancia
0,00-0,20	Nula
0,21-0,40	Débil
0,41-0,60	Moderada o suficiente
0,61-0,80	Buena o fuerte
0,81-1,00	Muy buena o fuerte

Fuente: Cálculo de fiabilidad y concordancia entre codificadores de un sistema de categorías para el estudio del foro online en *e-learning* (Torres y Perera, 2009).

Para complementar la prueba de concordancia, se empleó la prueba de Lawshe, también denominada Razón de Validez de Contenido (Content Validity Ratio [CVR]), cuyo objetivo es determinar de manera sistemática si los reactivos de un instrumento son considerados esenciales para la medición de un

constructo teórico. Por sí mismo, esto permite valorar la pertinencia del contenido de un instrumento y, además, indicar la dirección del acuerdo entre jueces. Este aspecto resulta relevante, ya que un grupo de jueces puede coincidir en considerar ineficaces o deficientes los reactivos de un instrumento, lo que generaría un nivel de acuerdo elevado, pero poco deseable desde la perspectiva de la validación de los atributos de sus propiedades métricas (Romero *et al.*, 2023).

Esta prueba se calcula mediante la siguiente fórmula:

Figura 2. Fórmula del índice de Lawshe

$$CVR = \frac{n_e - (N/2)}{N/2}$$

Donde n_e representa el número de jueces que calificaron el reactivo como “esencial”; mientras que N corresponde al número total de jueces participantes. Los valores críticos superiores a .75 se consideran satisfactorios (Lawshe, 1975, como se cita en Polit *et al.*, 2007).

En este punto, resulta importante señalar que los jueces participantes en el proceso de validación de contenido del instrumento eran profesionales con una destacada trayectoria académica. De conformidad con lo sugerido por Galindo-Domínguez (2020), se procuró que los participantes contaran con un sólido dominio del tema objeto de evaluación, experiencia en su abordaje, independencia entre sí y disposición para involucrarse en la tarea con el nivel de dedicación requerido por el modelo de validación empleado.

Uno de las juezas participantes fue la Dra. Gallardo, directora del Doctorado en Educación del Instituto Tecnológico y de Estudios Superiores de Monterrey (ITESM). Cuenta con una amplia trayectoria en investigación y producción científica, y ha desarrollado proyectos orientados a la promoción del pensamiento crítico en dicha institución.

El segundo juez participante en esta actividad fue de La Borbolla, doctor en Filosofía por la Universidad Complutense de Madrid, catedrático emérito de la Facultad de Filosofía y Letras de la Universidad Autónoma de México, conductor de programas culturales en televisión abierta y autor de numerosas columnas de opinión y obras de amplia difusión.

La tercera jueza fue la Dra. Vendrell, graduada y docente investigadora del programa de doctorado de la Universidad Complutense de Madrid, España. Es esta institución forma parte del departamento de investigación y psicología sobre pensamiento crítico, y colabora con el Centro de Psicometría de la Universidad de Cambridge. Adicionalmente, ha realizado estancias de investigación relacionadas con la evaluación y el desarrollo del pensamiento crítico en la Universidad de Oxford, Reino Unido, y en la Universidad de Barcelona, España.

A estos tres jueces se sumaron dos profesoras con amplia experiencia docente. La primera fue Peraza, docente de educación básica y superior con más de treinta años de trayectoria profesional, entre los que se incluyen cinco años como directora de una institución de educación superior, además de un reconocido desempeño en el ámbito educativo.

También participó la maestra en Innovación Uicab, docente de educación superior con más de dieciocho años de experiencia y amplio reconocimiento profesional como docente, diseñadora curricular, investigadora y asesora de trabajos de investigación para la obtención de grados académicos. Estas profesionales aportaron su experiencia en el trabajo de aula dentro de la institución en la que se identificó originalmente el problema abordado en esta tesis, por lo que contribuyeron con conocimientos y experiencia relacionados con la observación, evaluación y promoción del pensamiento crítico en contextos educativos.

Una vez que el modelo del instrumento superó las pruebas de concordancia de Fleiss y de Lawshe,

se administró a la población piloto con el propósito de generar una base de datos suficientemente robusta para la realización de los análisis que se describen a continuación.

En primer lugar, se examinó la confiabilidad mediante el coeficiente alfa de Cronbach. La confiabilidad se refiere a la capacidad de un instrumento para producir puntuaciones consistentes (Tornimbeni *et al.*, 2013). Para este estudio, se consideró aceptable un coeficiente superior a .70 y “bueno” a partir de .80 (Celina y Campo, 2005).

Otro análisis fundamental realizado fue el análisis factorial exploratorio, procedimiento que permite identificar patrones de agrupación o comunalidades entre los reactivos de un instrumento cuando no se dispone de una estructura factorial previamente definida. Como parte de este análisis, se revisaron dos pruebas asociadas de especial importancia: la medida de adecuación muestral Kaiser-Meyer-Olkin (KMO) y la prueba de esfericidad de Bartlett (Pizarro y Martínez, 2020).

La primera prueba evalúa la adecuación de la muestra para la realización del análisis factorial, es decir, permite verificar que las correlaciones entre las variables sean suficientemente fuertes como para sustentar un modelo factorial útil. La tabla 5 presenta la guía de interpretación de este coeficiente muestral.

Tabla 5. Interpretación del coeficiente Kaiser-Meyer-Olkin

Valor	Nivel de consistencia
$\geq 0,80$	Excelente
0,700,79	Bueno
0,60 – 0,69	Aceptable
0,50-0,59	Pobre o marginal
< 49	Inaceptable

Fuente: *El análisis exploratorio de los ítems: una guía práctica, revisada y actualizada* (Lioret *et al.*, 2014).

La prueba de esfericidad de Bartlett permite determinar si la matriz de correlaciones difiere de manera estadísticamente significativa de una matriz de identidad, es decir, de una estructura en la que las variables no se encuentran correlacionadas entre sí. En este procedimiento, la esfericidad se evalúa mediante el valor de p : cuando este es inferior a .05, se rechaza la hipótesis nula de identidad, lo que indica la existencia de correlaciones suficientes entre las variables y, por tanto, justifica la realización del análisis factorial (Lioret *et al.*, 2014).

Otro procedimiento de análisis empleado fue el Análisis de Funcionamiento Diferencial de los ítems (DIF), representado comúnmente mediante el estadístico χ^2 de Mantel-Haenszel (χ^2_{HM}). Su objetivo es identificar si existen reactivos que funcionen de manera diferente para distintos grupos de personas en función de alguna característica específica, como el sexo, el nivel de estudios o el grado escolar, entre otras (Leong *et al.*, 2020).

En este caso, se empleó el estadístico de Mantel-Haenszel, considerado uno de los procedimientos más estables para este tipo de análisis. De manera general, el método compara proporciones de respuesta correctas e incorrectas entre dos grupos definidos a partir de una categoría específica relacionada teóricamente con el constructo de estudio. Asimismo, cuando no se identifica sesgos asociados a las categorías analizadas, este procedimiento aporta evidencia que respalda la generalización de los resultados.

En este caso, se revisó el análisis diferencial en función del sexo, el grado escolar de los participantes y su procedencia institucional, de conformidad con dos categorías: “escuelas normales” e “instituciones de educación superior”, es decir, instituciones distintas a las pertenecientes al subsistema de educación normal.

Para determinar la presencia o ausencia de un comportamiento diferencial, se empleó la prueba de significancia habitual: si $\chi^2_{HM} > \text{Sig.}$, se retiene

la hipótesis nula, la cual establece que los reactivos funcionan de manera indistinta según el sexo, el grado escolar u otra categoría analizada. Para realizar estos análisis y generar la mayoría de gráficas empleadas en la investigación, se utilizó el programa *Statistical Package for Social Sciences* (SPSS), desarrollado por la empresa International Business Machines (IBM).

Consideraciones éticas

En el importante plano de las consideraciones éticas involucradas en el presente trabajo de investigación, se realizó una revisión cuidadosa no solo de los lineamientos generales aplicables a la investigación en las ciencias sociales y de la conducta, sino también de las disposiciones específicas relativas al uso de encuestas.

En este sentido, Singer (1995, como se cita en De Leeuw *et al.*, 2008), advierte que los investigadores adquieren responsabilidad tanto frente al público usuario de los resultados como ante la comunidad científica, ya que las decisiones personales o profesionales pueden sustentarse en la información que se comunica. Por ello, resulta indispensable maximizar el rigor metodológico a lo largo de todo el proceso de investigación.

En términos generales, se siguió el principio establecido en el Informe Belmont, elaborado en EE. UU., en 1979, por la Comisión Nacional para la Protección de Sujetos Humanos en la Investigación Biomédica y del Comportamiento, el cual puede resumirse en la siguiente máxima del apartado Beneficencia: “Minimizar los daños posibles y maximizar los beneficios posibles” (p. 4).

Adicionalmente, se aseguró el cumplimiento de tres principios fundamentales de la investigación en ciencias sociales: evitar cualquier daño a los participantes, prevenir situaciones de estrés prolongado y abstenerse de ocultar o tergiversar los fines y actividades del estudio (American Psychological Association [APA], s.f.; American Sociological Asso-

ciation [ASA], 1999; Office for Human Research Protection [OHRP], 2000).

Resultados

Validez de contenido

Como parte de validación del contenido del instrumento, se realizó una prueba de concordancia de Fleiss. Los resultados obtenidos para la revisión preliminar se presentan en la tabla 6.

En este reporte, se observa un coeficiente de concordancia de .720, valor que corresponde a un nivel de concordancia bueno o fuerte. Asimismo, destaca un p valor de .000, lo que sugiere una probabilidad mínima de que los resultados observados se deban al azar (Pennsylvania Department of Education [PDE], 2025). Al revisar las categorías de calificación, se encontró una concordancia

fuerte para las evaluaciones de nivel 3 ($K = 0,750$) y una concordancia muy fuerte para las de nivel 4 ($K = 0,970$).

En la tabla 7 se muestra el reporte de la prueba de Fleiss, en los casos individuales donde se encontraron puntajes de los jueces.

En la tabla 7 no se presentan métricas correspondientes a las categorías 1 y 2 de calificación, debido a que los jueces no asignaron estos niveles a ninguno de los reactivos evaluados. Esta situación constituye una evidencia relevante de la pertinencia identificada en el modelo del instrumento.

Razón de validez de contenido

En la prueba de Lawshe para la razón de validez de contenido por reactivo, se obtuvieron los resultados de la tabla 8.

Tabla 6. Prueba general de concordancia de Fleiss

Overall Kappa						
	Kappa	Asymptotic Standard Error	Z	P Value	Lower 95 % Asymptotic CI Bound	Upper 95 % Asymptotic CI Bound
Overall	.720	.082	8,820	.000	.560	.880

Fuente: elaboración propia.

Tabla 7. Prueba individual de concordancia de Fleiss

Kappas for Individual Categories							
Rating Category	Conditional Probability	Kappa	Asymptotic Standard Error	Z	P Value	Lower 95 % Asymptotic CI Bound	Upper 95 % Asymptotic CI Bound
3	.750	.720	.082	8,820	.000	.560	.880
4	.970	.720	.082	8,820	.000	.560	.880

Fuente: elaboración propia.

Tabla 8. Prueba de Lawshe por reactivo

Reactivo	Resultado	Reactivo	Resultado
1	$5-(5/2) / 5/2 = 1$	9	$5-(5/2) / 5/2 = 1$
2	$5-(5/2) / 5/2 = 1$	10	$5-(5/2) / 5/2 = 1$
3	$5-(5/2) / 5/2 = 1$	11	$4-(5/2) / 5/2 = 0.6$
4	$5-(5/2) / 5/2 = 1$	12	$5-(5/2) / 5/2 = 1$
5	$5-(5/2) / 5/2 = 1$	13	$5-(5/2) / 5/2 = 1$
6	$5-(5/2) / 5/2 = 1$	14	$5-(5/2) / 5/2 = 1$
7	$4-(5/2) / 5/2 = 0,6$	15	$5-(5/2) / 5/2 = 1$
8	$5-(5/2) / 5/2 = 1$	Total	$14.2/15 = 0,94$

Fuente: elaboración propia.

En este caso, se observa que los jueces alcanzaron un nivel fuerte de concordaron. Una vez obtenidos resultados favorables en relación con la validez de contenido del instrumento, se procedió a su aplicación piloto. Con la base de datos generada a partir de esta aplicación, se realizó la revisión de los atributos psicométricos del instrumento (Roebianto *et al.*, 2023).

Confiabilidad

La prueba de estabilidad mediante el coeficiente alfa de Cronbach realizada con los datos del pilotaje arrojó resultados satisfactorios. De acuerdo con los criterios establecidos por Oviedo y Campos (2005), el coeficiente obtenido en este análisis se consideró “bueno” o “fuerte”.

Tabla 9. Prueba de Confiabilidad por alfa de Cronbach

Alfa de Cronbach	N.º de elementos
.857	15

Fuente: elaboración propia.

Validez de constructo por análisis factorial

Para revisar la validez de constructo del instrumento, se realizó un análisis factorial exploratorio, dado que no se contaba con una estructura prelimi-

nar claramente definida que pudiera contrastarse mediante un análisis factorial confirmatorio. En su lugar, se partió de un modelo teórico sobre el cual se llevó a cabo un proceso de reducción de dimensiones (Méndez y Rondón, 2012).

El método de extracción empleado fue el de componentes principales, debido a su capacidad para reducir el número de factores en un conjunto de datos, conservando la mayor claridad posible de información y permitiendo identificar los patrones más relevantes de variación entre los reactivos (Lio-*ret et al.*, 2014). Al tratarse de un proceso exploratorio, la extracción se basó en autovalores superiores a 1, es decir, no se estableció previamente un número fijo de factores.

Por otra parte, se recurrió al método de rotación ortogonal Varimax, con el fin de maximizar la varianza entre las variables y así mantener su independencia relativa para el análisis particular de cada componente, lo cual resulta útil para orientar intervenciones específicas (Lio-*ret et al.*, 2014). Por consenso de expertos, se estableció que un nivel mínimo de saturación de .40 para la identificación de los componentes (López y Gutiérrez, 2019).

A continuación, la tabla 10 presenta los resultados obtenidos mediante este procedimiento.

Tabla 10. Matriz de componente rotado en el análisis factorial

	Componente		
	1	2	3
R1	.698	.159	.100
R9	.647	.300	.049
R10	.640	.186	.246
R11	.637	.173	.138
R12	.626	.154	.315
R2	.605	.261	.093
R13	.526	-.002	.345
R5	.177	.655	.011
R4	.349	.614	.105
R7	-.187	.593	.455
R3	.399	.564	.102
R8	.367	.450	.140
R15	.223	.091	.806
R14	.327	.084	.760
R6	.153	.394	.474

Fuente: elaboración propia.

Según esta matriz, el primer componente quedó conformado por siete reactivos con cargas factoriales entre .698 y .526; el segundo componente se integró por cinco reactivos, con cargas entre .655 y .450 y, el último, por tres reactivos con cargas de .806 a .474. Ninguno de los reactivos presentó una carga factorial inferior al criterio establecido, por lo que se mantuvieron los 15 reactivos

del modelo original. En cuanto a la varianza explicada por este modelo, se obtuvo un valor total de 50,389, lo que puede considerarse satisfactorio, ya que permite atribuir más de la mitad de la variabilidad de las respuestas a la estructura propuesta de tres dimensiones subyacentes. La tabla 11 presenta el reporte correspondiente a este análisis de varianza explicada.

Tabla 11. Varianza total explicada

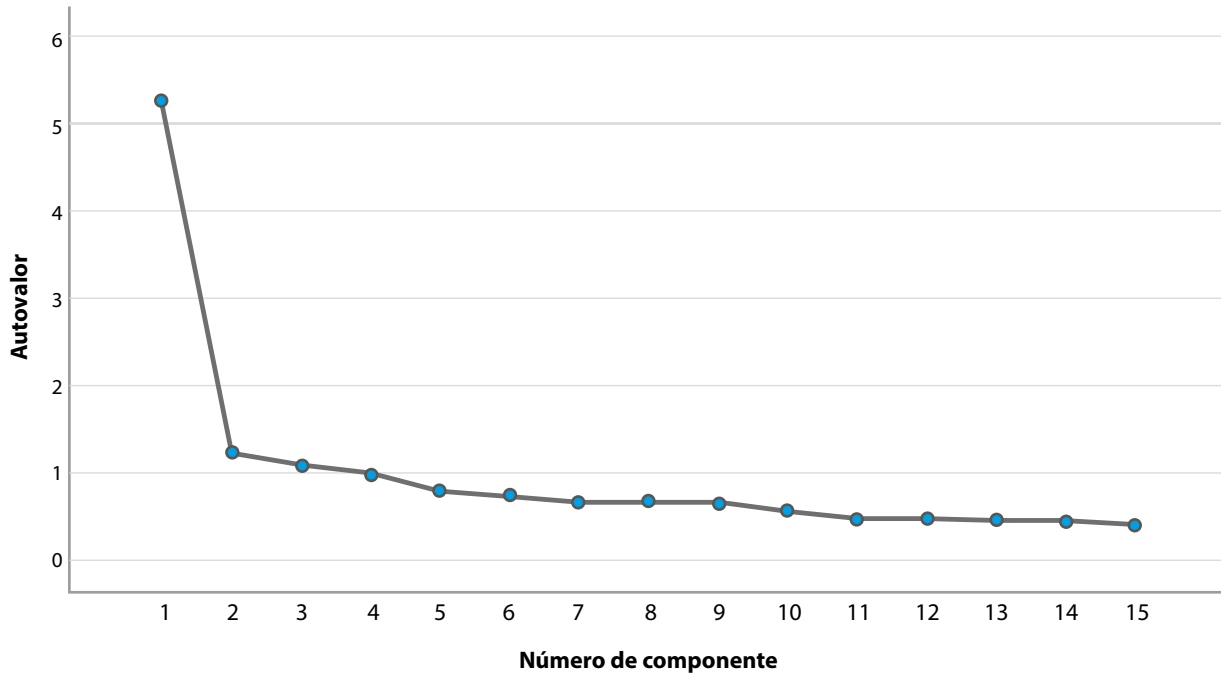
Componente	Autovalores iniciales			Sumas de cargas al cuadrado de la extracción			Sumas de cargas al cuadrado de la rotación		
	Total	% de varianza	% acumulado	Total	% de varianza	% acumulado	Total	% de varianza	% acumulado
1	5,241	34,942	34,942	5,241	34,942	34,942	3,418	22,786	22,786
2	1,228	8,183	43,125	1,228	8,183	43,125	2,121	14,140	36,926
3	1,090	7,264	50,389	1,090	7,264	50,389	2,019	13,462	50,389
4	.969	6,457	56,846						
5	.803	5,354	62,200						
6	.753	5,017	67,217						
7	.691	4,610	71,827						
8	.676	4,507	76,333						
9	.623	4,152	80,485						
10	.580	3,866	84,351						
11	.518	3,453	87,804						
12	.515	3,431	91,236						
13	.473	3,156	94,391						
14	.442	2,949	97,340						
15	.399	2,660	100,000						

Fuente: elaboración propia.

* Método de extracción: análisis de componentes principales.

De manera particular, el primer componente, con un autovalor de 5,241, explica el 34,942 % de la variabilidad de la prueba, lo que indica que constituye el elemento principal del constructo. Por su parte, el segundo componente, con un autovalor de 1,228, explica el 8,18 % de la variabilidad y, en conjunto con el componente anterior, acumula el 43,125 % de

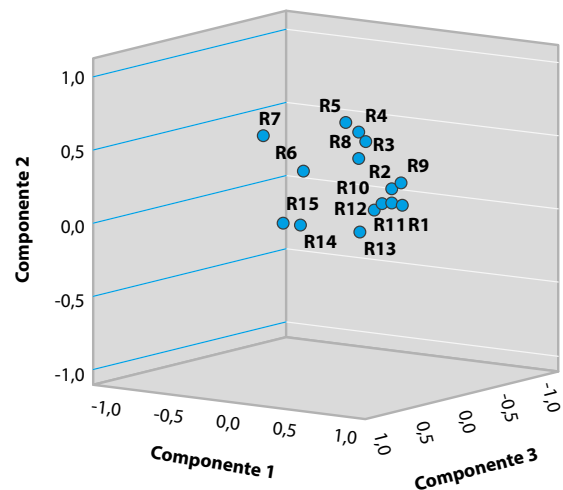
la varianza explicada. Finalmente, el tercer componente explica el 7,264 % de la variabilidad y eleva la varianza acumulada al 50,389 %, correspondiente al conjunto de componentes que superaron el criterio de autovalor mayor que 1. La figura 3 muestra el comportamiento de los reactivos respecto al criterio de autovalor mayor que 1.

Figura 3. Gráfico de sedimentación del análisis factorial

Fuente: elaboración propia empleando el programa *Statistical Package for Social Sciences* (SPSS).

Se observa una inflexión pronunciada entre el primer y el segundo componente, seguida de una tendencia estable en los valores posteriores. Este comportamiento indica que los tres primeros factores concentran la mayor parte de la varianza explicada, lo cual coincide con la estructura teórica del constructo, mientras que los componentes restantes realizan una contribución marginal.

Para ofrecer una perspectiva complementaria del comportamiento de los datos procesados mediante el análisis factorial y de los componentes identificados en el modelo propuesto, puede revisarse la posición de cada reactivo en el espacio rotado. Esta representación visual permite corroborar o ajustar los modelos de agrupación a partir de sus cargas factoriales y, en su caso, identificar valores atípicos que deban considerarse con atención al decidir sobre su ubicación, mantenimiento o exclusión.

Figura 4. Gráfico de componentes en espacio rotado

Fuente: elaboración propia, empleando el programa SPSS.

La distribución de los reactivos en el espacio factorial mostró agrupaciones consistentes, lo que indica una adecuada relación entre estos y las dimensiones latentes que representan los procesos cognitivos asociados a la variable de estudio. Esta disposición constituye una evidencia visual de la validez estructural del instrumento, al mostrarlo como un conjunto integrado de componentes interrelacionadas y no un constructo unidimensional.

Pruebas de adecuación del análisis factorial

La medida de adecuación muestral, expresada mediante el estadístico Kaiser-Meyer-Olkin (KMO = .910), indicó un nivel muy alto de correlación entre los reactivos del instrumento, lo que corroboró la idoneidad de los datos para la realización del análisis factorial. La prueba de esfericidad ($\chi^2 = 6053,756$, $p < .000$) resultó estadísticamente significativa, lo que permite considerar que la matriz de correlaciones difiere de identidad.

En consecuencia, existen correlaciones suficientes entre los reactivos, con lo que ratifica la pertinencia del modelo factorial empleado.

El reporte de estas dos pruebas de pertinencia del análisis factorial se muestran en la tabla 12.

Tabla 12. Pruebas Kaiser-Meyer-Olkin y de Esfericidad de Bartlett

Medida Kaiser-Meyer-Olkin de adecuación de muestreo		.910
Prueba de esfericidad de Bartlett	Aprox. Chi-cuadrado	6053,756
	gl	105
	Sig.	.000

Fuente: elaboración propia.

En su conjunto, estos dos indicadores muestran que la matriz de datos cumple los supuestos básicos para considerar pertinente y robusto el análisis factorial realizado, lo que respalda la extracción de los tres factores identificados.

Análisis de funcionamiento diferencial

Al analizar el comportamiento de los datos en función de las categorías “hombre” y “mujer”, con el propósito de identificar posibles sesgos en las puntuaciones asociados al sexo de los participantes, se obtuvo el siguiente resultado: el valor de p (.06) es superior al nivel de significancia establecido del 5 %, por lo que se retiene la hipótesis nula de igualdad entre los grupos. En consecuencia, no se detectaron diferencias estadísticamente significativas en el funcionamiento de la prueba entre hombres y mujeres.

El valor de la razón de momios (*Odds Ratio*, OR = 0,09) indica una ligera tendencia numérica a favor de las mujeres en las puntuaciones obtenidas, dado que el coeficiente es inferior a 1. No obstante, esta diferencia carece de significación estadística, por lo que puede interpretarse como una variación no sistemática.

El valor estadístico de la χ^2 (3,95) refuerza la conclusión de que no existe un comportamiento diferencial significativo, ya que no se alcanza el nivel de evidencia requerido para establecer la presencia de un sesgo real.

La tabla 13 reporta el detalle de esta prueba de análisis de diferencias sistemáticas:

Tabla 13. Prueba Mantel-Haenzel de funcionamiento diferencial por sexo

Medida Mantel-Haenzel	Valor
a (hombre-alto)	136,0
b (hombre-no alto)	316,0
c (mujer-alto)	249,0
d (mujer-no alto)	749,0
OR_MH (hombre contra mujer)	0,0945961059427600
CMH χ^2	3,950164583640220
gl	1,0
p-valor	0,064686678380784300

Fuente: elaboración propia.

Este mismo análisis se replicó en función del grado escolar de los estudiantes. Para este caso, la tabla 14 presenta los resultados obtenidos.

Tabla 14. Prueba Mantel-Haenzel de funcionamiento diferencial por grado escolar

Medida Mantel-Haenzel	Valor
a (Grado 1-2 alto)	219,0
b (Grado 1-2 no alto)	554,0
c (Grado 3-4 alto)	169,0
d (Grado 3-4 no alto)	516,0
OR_MH (Grado 1-2 contra 3-4)	1,2069724221904200
CMH Chi-2	2,3050751084515000
gl	1,0
p-valor	0,12895204723580600

Fuente: elaboración propia.

El *p* valor (0,128) indica que no existen una diferencias estadísticamente significativas entre los estudiantes de grados inferiores y superiores. Por su parte, el valor OR (1,20) sugiere que los alumnos de primero y segundo grado presentan una ligera ventaja (1,20 veces) en la probabilidad de alcanzar puntuaciones altas en comparación con los estudiantes de tercero y cuarto grado. Sin embargo, dado que esta diferencia no es estadísticamente significativa ($p > 0,05$), puede considerarse que existe una equivalencia funcional entre ambos grupos. De esto se reporta en la tabla 15:

Tabla 15. Prueba Mantel-Haenzel de funcionamiento diferencial por tipo de escuela

Medida Mantel-Haenzel	Valor
a (Normal alto)	269,0
b (Normal no alto)	504,0
c (IES alto)	399,0

Medida Mantel-Haenzel	Valor
d (IES no alto)	416,0
OR_MH (escuela 1-2 contra 3-4)	0,8759724223224200
CMH Chi-2	1,0141751184505100
gl	1,0
p-valor	0,06663214733620400

Fuente: elaboración propia.

En este caso, la medida OD presenta un valor cercano al 1, que constituye el valor de referencia. Este resultado puede interpretarse como una probabilidad prácticamente equivalente de obtener puntuaciones altas en pensamiento crítico tanto en estudiantes de escuelas normales como en estudiantes de otros tipos de instituciones.

No obstante, el hecho de que el valor sea ligeramente inferior a 1 permite advertir una leve desventaja para los estudiantes normalistas en la obtención de estas puntuaciones deseables. Sin embargo, dado que el rango establecido para considerar equivalencia comprende valores entre .85 a 1,15, puede concluirse que no existe un comportamiento no diferencial entre las poblaciones analizadas.

Asimismo, el valor de *p* de la prueba (0,06) es superior al nivel de significancia de .05, lo que refuerza la evidencia de igualdad entre los grupos comparados. Así, puede concluirse que los resultados indican que las puntuaciones del instrumento funcionan de manera homogénea en los grupos analizados, es decir, su funcionamiento es consistente entre distintos grados escolares, sexo y tipos de institución. Estos hallazgos respaldan la propuesta de generalizar su uso a poblaciones de educación superior distintas de las escuelas normales.

Finalmente, el instrumento se presenta en la tabla 16:

Tabla 16. Modelo final del instrumento para medir el pensamiento crítico

Factor o indicador	Reactivo
Análisis de información y argumentos	Elaboro ideas propias y puedo explicarlas o sustentarlas mediante argumentos lógicos.
	Busco los argumentos que sustentan las afirmaciones que me presentan.
	Cuando se me señala un error, procuro comprender los fundamentos de dicho señalamiento.
	Identifico conclusiones erróneas a partir de los argumentos y datos disponibles.
	Reconozco qué información necesito para formular la conclusión.
	Comprendo razonamientos relacionados con temas poco familiares.
	Cuando abordo un problema, distingo la información relevante de aquella que no lo es.
Disposición para la evaluación y corrección	Evalúo objetivamente las ideas de otras personas, basándome en la solidez de sus argumentaciones.
	Defiendo mi postura mediante argumentos lógicos, incluso cuando no sea compartida por la mayoría.
	Examino los argumentos que sustentan ideas contrarias a las mías.
	Cuestiono las ideas que carecen de evidencia o argumentos, incluso cuando provienen de una figura autoridad.
	Modifico mi postura cuando la evidencia disponible es suficiente y racionalmente convincente.
Exposición de argumentos concluyentes.	Elaboro conclusiones claras a partir del análisis de argumentos y datos.
	Expongo mis razonamientos de manera coherente al abordar problemas complejos.
	Verifico que las fuentes que consulto cuenten con credenciales académicas o científicas adecuadas.

Fuente: elaboración propia.

Conclusión

Se ratificó el problema de investigación planteado, consistente en la inexistencia de un instrumento práctico para medir el pensamiento crítico de estudiantes de educación superior en el contexto mexicano, que contara con adecuados atributos psicométricos de confiabilidad y validez, así como con acceso abierto para su utilización masiva en instituciones educativas. La revisión documental de fuentes académicas especializadas permitió constatar que la mayoría de los instrumentos y materiales disponibles sobre este tema se centran en contextos extranjeros y, principalmente, en el ámbito de la educación básica.

La revisión del marco teórico mostró que la noción de pensamiento crítico no ha experimentado variaciones sustanciales a lo largo del tiempo, según se desprenda de las obras consultadas de los principales exponentes del tema. Más bien, en los últimos años se ha enfatizado la importancia de su dimensión actitudinal, la cual siempre ha sido reconocida como un componente de esta capacidad, aunque históricamente quedó relegada frente a la dimensión cognitiva. Esta última, a su vez, se ha centrado de manera predominante en la habilidad argumentativa.

La hipótesis nula fue rechazada, puesto que el instrumento diseñado para medir el pensamiento crítico en estudiantes de educación superior pre-

sentó resultados satisfactorios en todas las pruebas técnicas a las que se fue sometido, entre ellas de lenguaje, estructura, contenido, confiabilidad y validez. En todos los casos, se comprobó que el instrumento construido cumplió con los criterios internacionales de la disciplina psicométrica. En consecuencia, se alcanzó el objetivo general de la investigación, orientado a la construcción de un instrumento con atributos de diseño satisfactorios.

El proceso de validación ciega de contenido mediante jueces expertos fue satisfactorio y contribuyó a que la versión del cuestionario aplicada en el pilotaje generara datos válidos y consistentes. Esto se reflejó en los resultados obtenidos en las distintas pruebas estadísticas empleadas durante el proceso de validación.

El instrumento resultante de esta investigación puede considerarse una contribución relevante, ya que atiende la medición de un tema de gran importancia para los sistemas de educación superior. Además, se presenta en un formato práctico y no mostró sesgos ni funcionamiento diferencial en los grupos analizados, lo que favorece su utilización en estudios de amplia cobertura e, incluso, en aplicaciones censales, cuando sea necesario.

Recomendaciones

Como resultado de conducir esta investigación, fue posible recoger experiencias valiosas de las cuales formular algunas recomendaciones para consideración de los lectores de este documento que estén interesados en abordar o profundizar en esta misma línea de estudio.

Conviene prestar atención a la creciente tendencia entre los investigadores del pensamiento crítico de incorporar un componente ético a esta capacidad. Si bien esta postura aún no se ha difundido

ampliamente, resulta pertinente seguir de cerca su desarrollo, pues, de consolidarse, podría representar una transformación relevante en la tradición académica sobre el tema.

Asimismo, puede resultar de gran interés explorar la relación entre el pensamiento crítico y la autorregulación del aprendizaje mediante estudios correlacionales. Esta otra constituye otro de los elementos centrales en las aspiraciones formativas de las instituciones de educación superior. En efecto, el discurso curricular presenta ambas capacidades como rasgos relevantes del perfil de egreso, asociados a la formación de profesionales altamente competentes y comprometidos con la mejora continua.

Desde el punto de vista técnico, también se recomienda realizar un análisis factorial confirmatorio en una población diversificada, con el fin de ratificar la estabilidad y el funcionamiento del modelo de tres dimensiones. Por otra parte, los estudios orientados a construir un modelo amplio de evaluación del pensamiento crítico requerirán incorporar fuentes de valoración de la variable, especialmente aquellas basadas en la observación directa del desempeño de los estudiantes en situaciones reales de intervención pedagógica.

Esto permitirá estimar la concordancia entre los resultados obtenidos mediante distintos instrumentos y estrategias de recolección de datos, lo que fortalezca la validez interna o convergente de la evaluación. Por otra parte, dado que este trabajo no tuvo un alcance experimental ni de intervención, puede proporcionarse como línea futura de investigación la evaluación de la eficacia de estrategias orientadas al fortalecimiento del pensamiento crítico en los estudiantes. De este modo, sería posible contribuir al desarrollo de estrategias que promuevan cambios positivos en los procesos formativos.

Referencias

- American Psychological Association. (s.f.). *Ethical Principles of Psychologists and Code of Conduct*.
- Betancourt-Zambrano, S., Muñoz-Morán, K. T. y Rosas, T. J. (2017). Evaluación del pensamiento crítico en estudiantes de educación superior de la región de Atacama-Chile. *Prospectiva. Revista de Trabajo Social e Intervención social*, (23), 199-223. <https://doi.org/10.25100/prts.voi23.4594>
- Bezanilla, M. J., Galindo-Domínguez, H. y Poblete, M. (2021). Importance of teaching critical thinking in higher education and existing difficulties according to teacher's views. *Multidisciplinary Journal of Educational Research*, 11(1). <https://doi.org/10.17583/remie.o.6159>
- Celina-Oviedo, H. y Campo-Arias, A. (2005). Aproximación al uso del coeficiente Alpha de Cronbach. *Revista Colombiana de Psiquiatría*, 34(4), 572-580.
- de la Borbolla, Ó. (2017). *El arte de dudar*. Penguin Random House.
- de la Borbolla, Ó. (2019). *La rebeldía de pensar* (5.^a ed.). Fondo de Cultura Económica.
- de Leeuw, E., Hox, J. y Dilman, D. (2008). *International handbook of survey methodology*. European Association of Methodology.
- Ennis, R., Gardiner, W., Guzzeta, J., Morrow, R., Paulus, D. y Ringel, L. (1964). *The Cornell conditional-reasoning test, form X*.
- Ennis, R. H. y Millman, J. (1985). *Cornell critical thinking tests*. Illinois Critical Thinking Project, Department of Educational Policy Studies, University of Illinois at Urbana-Champaign.
- Escobar-Domínguez, M. G. (2022). Práctica pedagógica del pensamiento crítico desde la psicología cultural. *Revista Sophia*, (36), 301-326. <https://doi.org/10.17163/soph.n36.2024.10>
- Facione, P. (1990). *Critical Thinking: a statement of expert consensus for purposes of educational assesment and instruction*. American Philosophical Association.
- Facione, P. (2000). *Test California de Destrezas en Pensamiento Crítico CCTST-2000*. Insight Assessment – The California Academic Press.
- Galindo-Domínguez, H. (2020). *Estadística para no estadísticos. Una guía básica sobre la metodología cuantitativa de trabajos académicos*. Editorial Área de Innovación y Desarrollo, S. L. <https://doi.org/10.17993/EcoOrgyCso.2020.59>
- García, L. M. y López, J. (Coords). (2020). *El pensamiento crítico para el aprendizaje*. Universidad Nacional Autónoma de México.
- Haber, J. (2020). *Critical Thinking*. The MIT Press. <https://doi.org/10.7551/mitpress/12081.001.0001>

- Halpern, D. F. (2006). *Halpern Critical Thinking Assessment: Using every day situation: background and scoring standards* [Informe técnico]. Claremont Mckenna College.
- Halpern, D. F. (2003). The How and Why of critical thinking assessment. En D. Fasko (Ed.), *Critical thinking and reasoning. Current research, theory, and practice* (pp. 355-366). Hampton Press.
- Halpern, D. F. (2014). *Thought and Knowledge: An Introduction to Critical Thinking* (4.ª ed.). Psychology Press.
- Harvard Business Review. (2023). *HBR Guide to Critical Thinking. Ask the right questions. Analyze and learn. Solve complex problems.*
- Hatcher, D. (2011). Which test? Whose scores? Comparing standardized critical thinking tests. *New Directions for Institutional Research*, (149). <https://doi.org/10.1002/ir.378>
- Klaus, B. (2023). Ensuring content validity of psychological and educational tests: The role of experts. *Frontline Learning Research*, 8(6), 1-37. <https://doi.org/10.14786/flr.v8i6.517>
- Leong, F., Bartram, D., Cheung, F., Geisinger, K. y Iliescu, D. (2020). *Manual internacional de pruebas y evaluación del ITC*. Manual Moderno.
- Lincoln, D. y Kearney, M. L. (2019). Promoting critical thinking in higher education. *Studies in Higher Education*, 44(5), 799-800, <https://doi.org/10.1080/03075079.2019.1586322>
- Lioret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A. y Tomás-Miguel, I. (2014). Exploratory item factor analysis: A practical guide revised and up-dated. *Anales del Psicología*, 30(3), 1151-1169. <https://doi.org/10.6018/analeps.30.3.199361>
- Liu, O., Mao, L., Frankel, L. y Crotts-Roohr, K. (2014). Assessing critical thinking in higher education: Current state and directions for next generations assessments. *ETS Research Reports Series*, 2014: 1-23. <https://doi.org/10.1002/ets2.12009>
- Liu, O., Mao, L., Frankel, L. y Xu, J. (2016). Assessing critical thinking in higher education: The HEIghten approach and preliminary validity evidence. *Assessment & Evaluation in Higher Education*, 41(5). <https://doi.org/10.1080/02602938.2016.1168358>
- López-Aguado, M. y Gutiérrez-Provecho, L. (2019). Cómo realizar e interpretar un análisis factorial exploratorio utilizando SPSS. *REIRE Revista d'Innovació i Recerca en Educació*, 12(2), <https://doi.org/10.1344/reire2019.12.227057>
- Mackay-Castro, R., Franco-Cortázar, D. y Villacís-Pérez, P. (2018). El pensamiento crítico aplicado a la investigación. *Universidad y Sociedad*. 10(1). <https://rus.ucf.edu.cu/index.php/rus/article/view/803>
- Maksimovic, J. y Evtimov, J. (2023). Positivism and post-positivism as the basis of quantitative research in pedagogy. *Research in Pedagogy*, 13(1). <https://doi.org/10.5937/IstrPed2301208M>
- Medrano, L. y Pérez, E. (2023). *Manual de psicometría y evaluación psicológica*. Editorial Brujas.

- Méndez-Martínez, C. y Rondón-Sepúlveda, M. A. (2012). Introducción al análisis factorial exploratorio. *Revista Colombiana de Psiquiatría*, 41(1), 197-207. [https://doi.org/10.1016/S0034-7450\(14\)60077-9](https://doi.org/10.1016/S0034-7450(14)60077-9)
- Miranda-J., Ch. (2003). El pensamiento crítico en docentes de educación general básica en Chile: un estudio de impacto. *Estudios pedagógicos*, (29), 39-54. <https://doi.org/10.4067/S0718-07052003000100003>
- Namakforoosh, M. (2014). *Metodología de la investigación*. Limusa.
- Office for Human Research Protection. (2020). *International Compilation of Human Research Standards*.
- Ossa-Cornejo, C., Palma-Luengo, M., Lagos-San Martín, N., Quintana-Abello, I. y Díaz-Larenas, C. (2017). Análisis de instrumentos de medición del pensamiento crítico. *Ciencias Psicológicas*, 11(1), 19-28. <https://doi.org/10.22235/cp.v11i2.1343>
- Pennsylvania Department of Education. (2025). *Making sense of NCEs and standard errors*.
- Pimienta, J., De la Orden, A. y Estrada, R. (2018). *Metodología de la investigación*. Pearson.
- Pizarro-Romero, K. y Martínez-Mora, E. O. (2020). Análisis factorial exploratorio mediante el uso de las medidas de adecuación muestral KMO y esfericidad de Bartlett para determinar factores principales. *Journal of Science and Research*, (5), 903-924. <https://revistas.utb.edu.ec/index.php/sr/article/view/1046>
- Polit, D., Beck, C. y Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health*, 30(4). <https://doi.org/10.1002/nur.20199>
- Ponce-Ceballos, S., Aceves-Villanueva, Y. y Serna-Rodríguez, A. (2017). Prácticas de la tutoría académica en una escuela normal de México. *XVII Congreso Nacional de Investigación Educativa*. Consejo Mexicano de Investigación.
- Rancière, J. (2022 [1987]). *El maestro ignorante. Cinco lecciones sobre la emancipación intelectual*. Libros del Zorzal.
- Redhana, W., Karyasa, W. y Fitria-Atrisa, N. (2017). Development Of Critical Thinking Disposition Inventory. *Advances in Social Science, Education and Humanities Research*, 14(1). <https://doi.org/10.2991/icirad-17.2017.37>
- Roebianto, A., Savitri, S. I., Aulia, I., Suciya, A. y Mubarak, L. (2023). Content validity: definition and procedure of content validation in psychological research. *TPM*, 30(1), 5-18. <https://doi.org/10.4473/TPM30.1.1>
- Robles-Robles, Alfredo. (2019). La formación del pensamiento crítico: habilidades básicas, características y modelos de aplicación en contextos innovadores. *Revista de Ciencias Humanísticas y Sociales*, 4(2), 13-24. <https://doi.org/10.33936/rehuso.v4i2.2128>
- Romero-Jeldres, M., Díaz-Costa, E. y Faousi-Nadim, T. (2023). A review of Lawshe's method for calculating content validity in the social sciences. *Frontiers in Education*, 8. <https://doi.org/10.3389/educ.2023.1271335>
- Saiz, C. y Rivas, S. (2008). Evaluación en pensamiento crítico: una propuesta para diferenciar formas de pensar. *Ergo, Nueva Época*, (22-23), 22-66.

- Secretaría de Educación Pública. (2022a). *Anexo 5. Licenciatura en educación primaria* [Documento normativo oficial].
- Secretaría de Educación Pública. (2022b). *Disposiciones generales de los planes y programas de estudio de las licenciaturas para la formación de maestras y maestros de educación básica* [Documento normativo oficial]
- Schick, T. y Vaughn, L. (2024 [1994]). *Pensamiento Crítico. Cómo pensar acerca de cosas raras*. McGraw Hill.
- Tornimbeni, S., Pérez, E. y Olaz, F. (2013). *Introducción a la Psicometría*. Editorial Paidós.
- Torres-Gordillo, J. J. y Perera-Rodríguez V. H. (2009). Cálculo de fiabilidad y concordancia entre codificadores de un sistema de categorías para el estudio del foro online en e-learning. *Revista de Investigación Educativa*, 27(1), 89-103. <https://revistas.um.es/rie/article/view/94291>
- Vendrell-Morancho, M. (2024). Desarrollo del Pensamiento Crítico en la universidad: estrategias para superar desafíos y fomentar su crecimiento. *Revista de la Educación Superior*, 53(210), 1-21. <https://resu.anuies.mx/ojs/index.php/resu/article/view/2848>
- Vendrell, M. y Rodríguez-Mantilla, J. M. (2020). Pensamiento crítico: conceptualización y relevancia en el seno de la educación superior. *Revista de Educación Superior*, 49(194), 45-62. <https://resu.anuies.mx/ojs/index.php/resu/article/view/1124>
- Vincent-Lancrin, S. (2021). *Skills for life: fostering critical thinking*. IDB. <https://doi.org/10.18235/0003743>
- Watson, G. y Glaser, E. (1980). *Watson-Galser critical thinking appraisal*. Pearson.