

Diseño y validación de CritiTest, un instrumento para evaluar el pensamiento crítico en estudiantes universitarios

Design and Validation of CritiTest: An Instrument for Evaluating Critical Thinking Skills among University Students

Mireia Vendrell-Morancho¹, María José Fernández-Díaz²

¹ Universidad Complutense de Madrid mvendrel@ucm.es

² Universidad Complutense de Madrid mjfdiaz@ucm.es

Recibido: 8/8/2023

Aceptado: 28/10/2024

Copyright ©

Facultad de CC. de la Educación y Deporte.
Universidad de Vigo



Dirección de contacto:
Mireia Vendrell Morancho
Facultad de Educación – Centro de
Formación del Profesorado
Edificio La Alameda
C/ Rector Royo Villanova, 1
Ciudad Universitaria
28040 Madrid

Resumen

La limitada disponibilidad de herramientas especializadas para evaluar el pensamiento crítico en el estudiantado universitario hispanohablante supone un obstáculo para su desarrollo. En este estudio, se diseñó y validó *CritiTest*, un instrumento de evaluación del pensamiento crítico para estudiantes universitarios y universitarias de habla hispana. Para ello, se elaboró una definición integral de pensamiento crítico mediante una revisión bibliográfica y hermenéutica, que fue respaldada por una exhaustiva fundamentación teórica, la valoración de personas expertas y los análisis realizados en la prueba piloto. Los resultados indican que el pensamiento crítico es un proceso intelectual que implica la habilidad para analizar y evaluar argumentos y resolver problemas, así como la actitud y el conocimiento necesarios para hacerlo. *CritiTest* demostró buena-excelente fiabilidad a nivel global ($\alpha = 0,9$) y por dimensiones ($\alpha = 0,8$), posicionándose como uno de los instrumentos más consistentes a nivel internacional. En resumen, se concluye que *CritiTest* es una herramienta valiosa para optimizar el desarrollo y evaluación del pensamiento crítico en la Educación Superior hispanohablante. Asimismo, se sugieren posibles líneas de investigación a fin de seguir mejorando la evaluación del pensamiento crítico y, por ende, contribuir al fortalecimiento de la formación académica y personal del estudiantado.

Palabras clave

Pensamiento Crítico, Educación Superior, Construcción de Pruebas, Evaluación Educativa, Evaluación de Estudiantes

Abstract

The limited availability of specialised tools for assessing critical thinking in Spanish-speaking university students poses a significant challenge to its development. In this study, we designed and validated *CritiTest*, an assessment tool specifically aimed at evaluating critical thinking among Spanish-speaking university students. This was achieved by developing a comprehensive definition of critical thinking through a rigorous bibliographic and hermeneutic review, underpinned by a solid theoretical foundation, expert evaluations, and analysis from a pilot test. Our findings indicate that critical thinking is an intellectual process that encompasses the ability to analyse

and evaluate arguments, solve problems, and integrate the necessary attitude and knowledge to do so effectively. *CritiTest* demonstrated strong to excellent reliability at both the global level ($\alpha=0,9$) and across its dimensions ($\alpha=0,8$), positioning itself as one of the most robust instruments internationally. In conclusion, *CritiTest* is a valuable tool for enhancing the development and evaluation of critical thinking in Spanish-speaking higher education. We also suggest potential avenues for research to further enhance critical thinking evaluation, thus contributing to students' academic and personal growth.

Key Words

Critical Thinking, Higher Education, Test Construction, Educational Assessment, Student Evaluation

1. INTRODUCCIÓN

El pensamiento crítico es un componente fundamental en la educación contemporánea. Se destaca en discursos políticos, es una demanda creciente en la industria y constituye un objetivo clave que las instituciones educativas buscan desarrollar. En este contexto, su enseñanza es especialmente relevante en el ámbito universitario, donde se espera que las y los jóvenes líderes emergentes desarrollen las habilidades, actitudes y conocimientos necesarios para enfrentar los objetivos de desarrollo sostenible y liderar el progreso social.

No obstante, el desarrollo efectivo del pensamiento crítico en el entorno universitario enfrenta barreras significativas, como la carencia de claridad en su definición (Ericson, 2022; Plummer et al., 2022), la insuficiencia de recursos y la limitada capacitación docente (Archila et al., 2022; Frykholm, 2021), además de la falta de instrumentos de evaluación específicos para estudiantes universitarios (O'Leary et al., 2020). En este sentido, consideramos que el desarrollo y validación de un instrumento de evaluación contribuiría a minimizar estas barreras, permitiendo medir de manera precisa y objetiva el pensamiento crítico, adaptar estrategias de enseñanza, identificar deficiencias tempranamente y evaluar la efectividad de los programas educativos diseñados para mejorar esta competencia.

Por ello, el objetivo de este estudio es diseñar y validar un instrumento que permita evaluar el pensamiento crítico del estudiantado universitario hispanohablante. Para alcanzar este propósito, se ha realizado un análisis exhaustivo de las diferentes definiciones de pensamiento crítico propuestas en la literatura académica, con el fin de elaborar una definición integral que combine enfoques filosóficos, psicológico-cognitivos y educativos. Es importante destacar que la justificación teórica del instrumento de evaluación es exhaustiva y se enmarca en un estudio más amplio; sin embargo, este artículo, por limitaciones de espacio, se centra en los aspectos más relevantes de dicha fundamentación y en la descripción del diseño y validación del instrumento.

El pensamiento crítico ha sido objeto de debate a lo largo del tiempo, y su definición ha sido considerada problemática por diversos autores. Para abordar su enseñanza y evaluación, resulta esencial contar con una definición clara y precisa. A partir del análisis de la literatura académica, se han identificado dos dimensiones clave del pensamiento crítico: el *Análisis y Evaluación de Argumentos*, y la *Resolución de Problemas*. Estas dimensiones incluyen subdimensiones que permiten una comprensión más detallada del concepto, las cuales han servido de base para el diseño del instrumento de medición.

El pensamiento crítico es un proceso cognitivo complejo, holístico y de orden superior (Kageyama et al., 2022; Kleemola et al., 2022) que tiene como objetivo explorar declaraciones o problemas para llegar a una conclusión válida o a la elección de la alternativa con más probabilidad de éxito (Dwyer, 2017; Halpern, 2014). Para lograr este objetivo, se requiere aplicar habilidades cognitivas de manera organizada, disciplinada y cuidadosa, siguiendo estándares racionales y operando de manera deliberada, exhaustiva y sistemática (Gul y Akcay, 2020; Rauscher y Badenhorst, 2021).

Por otro lado, el pensamiento crítico es una habilidad educable y mejorable (Sutiani et al., 2021) que se compone de actitudes como el deseo de estar bien informado, la humildad intelectual y la integridad moral, y que se entrelaza con el sentido y razonamiento ético (Ennis, 2015; Paul y Elder, 2019). En este sentido, el pensamiento crítico implica una evaluación cuidadosa y constructiva de las ideas y argumentos, y un fomento del escepticismo amable, es decir, una inclinación a cuestionar sin caer en la parálisis por análisis (Danczak et al., 2020; Goodsett, 2020). A este respecto, el adjetivo “crítico” de pensamiento crítico no se refiere a ser crítico en el sentido negativo, sino más bien a ser crítico en el sentido de realizar una evaluación cuidadosa y constructiva.

En síntesis, el desarrollo del pensamiento crítico está vinculado con el juicio reflexivo y la toma de decisiones razonables y éticas, siendo una habilidad relevante en múltiples ámbitos, desde la educación hasta el mundo laboral y personal. Esta competencia no solo requiere de habilidades cognitivas, sino también de actitudes y valores éticos, lo que evidencia su complejidad y potencial para ser enseñada y mejorada.

A continuación, se describe el proceso de diseño del instrumento de evaluación, basado en la definición integral del pensamiento crítico, y se examinan sus características técnicas, incluyendo la fiabilidad y la validez. Finalmente, se presentan las conclusiones del estudio, las limitaciones y las posibles líneas de investigación futuras para profundizar en la evaluación del pensamiento crítico en estudiantes universitarios hispanohablantes.

2. MÉTODO

2.1. Definición operativa

A partir de lo previamente expuesto, se puede inferir que el pensamiento crítico se compone de un conjunto de habilidades cognitivas y actitudes de disposición interrelacionadas que resultan fundamentales para la toma de decisiones informadas. En este trabajo, se propone un modelo de pensamiento crítico enfocado en las habilidades cognitivas necesarias para tomar decisiones reflexivas sobre qué creer y qué hacer. Dicho modelo se basa en la conceptualización de Ennis, quien definió el pensamiento crítico como “un pensamiento reflexivo y razonable que se centra en decidir qué creer o qué hacer” (Ennis, 1985, p. 45). En particular, el modelo propuesto se divide en dos habilidades intelectuales de orden superior: *Análisis y Evaluación de Argumentos*, y *Resolución de Problemas*.

La dimensión de *Análisis y Evaluación de Argumentos* se ha operativizado en los reactivos que componen el instrumento y se basa en dos componentes clave. En primer lugar, la capacidad de analizar argumentos, lo que implica detectar, examinar e identificar las diferentes partes de un argumento, sus relaciones y el principio integrador. En segundo

lugar, se encuentra la capacidad de recopilar e interpretar la información comprendida en el argumento (Tabla 1).

Subdimensión I	Indicadores
Identificación y análisis de argumentos	Identificación de argumentos
	1. Detección de argumentos explícitos e implícitos
	2. Reconstrucción de argumentos a partir de las premisas y la conclusión
	Identificación de las partes constitutivas básicas
	3. Identificación de conclusiones (declaradas y no declaradas, principales y secundarias)
Identificación y análisis del contenido	4. Identificación de premisas: implícitas y explícitas
	5. Identificación de argumentos circulares y tautología
	Descodificación de la información. Identificación de:
	6. Información fáctica (basada en hechos) y opiniones o conjeturas
	7. Claridad en el lenguaje: precisión léxica y gramatical; repetición y detalle
	8. Definición: condiciones necesarias y suficientes; inclusión o exclusión excesiva y equívoco
	9. Neutralidad: lenguaje emotivo, efectos visuales y auditivos
	Interpretación del significado. Identificación de:
	10. Tesis que se quiere defender
	11. Asunciones, implicaciones, presupuestos, sobreentendidos, condiciones, valores, creencias, prejuicios, estereotipos
Identificación y análisis de las relaciones	12. Intención, propósito, objetivo, meta o función
	13. Implicaciones, efectos o consecuencias de las conclusiones
	Intraargumentativas
	14. Identificación de la relación entre las premisas y la conclusión
	15. Identificación de la relación entre premisas
	Interargumentativas
	16. Identificación de concatenación
17. Identificación de la argumentación doble	
18. Identificación e interpretación de la objeción	
19. Identificación e interpretación de la recusación	
20. Identificación de la refutación	

Tabla 1. Indicadores de Subdimensión de Análisis de Argumentos

La evaluación de los argumentos es otro componente clave de la dimensión de *Análisis y Evaluación de Argumentos*, que implica estimar la fuerza o debilidad del argumento, independientemente del nivel de acuerdo con la conclusión (Tabla 2).

Subdimensión II	Indicadores
Evaluación pasiva	Criterio 1: Credibilidad (veracidad y aceptabilidad)
	21. Discriminación entre premisas verdaderas, probables y engañosas
	22. Reconocimiento de falacias comunes relativas a la credibilidad
	23. Estima el nivel de confianza a la probabilidad o verdad de las premisas
	24. Examina las posibilidades de sesgo en la evidencia ofrecida
	Criterio 2: Relevancia
	25. Identificación de falacias comunes relativas a la relevancia: Considerar la fuente del argumento en lugar del argumento
	26. Identificación de falacias comunes relativas a la relevancia: Apelación a un estándar inapropiado
	27. Identificación de falacias comunes relativas a la relevancia: Irse del tema y no avanzar en la argumentación
	Criterio 3: Suficiencia
	28. Identificación de errores comunes relativos a la suficiencia: Generalización
	29. Identificación de errores comunes relativos a la suficiencia: Analogía
	30. Identificación de errores comunes relativos a la suficiencia: Falsa Causalidad
31. Apelación a la ignorancia	
Criterio 4: Ética	
32. Rechazo de actos no éticos	
Evaluación activa	Fortalecimiento
	33. Identificación de información complementaria que apoya la conclusión
	Debilitación
	34. Reconocimiento de la necesidad de buscar pruebas contrarias
	35. Identificación de interpretaciones/explicaciones alternativas plausibles
Desarrollo de cuestionamientos e identificación de información adicional	
36. Identificación de información adicional	
37. Identificación de la responsabilidad carga de la prueba	

Tabla 2. Indicadores de Subdimensión de Evaluación de Argumentos

Por otro lado, la dimensión de *Resolución de Problemas* se ha desglosado en la Tabla 3. Esta dimensión se refiere a la capacidad de reconocer y resolver problemas de manera lógica y sistemática. Los reactivos que componen esta dimensión se han diseñado para evaluar la capacidad de identificar y definir el problema, la capacidad de generar y evaluar diferentes soluciones y la capacidad de elegir la solución más adecuada.

Subdimensión	Indicadores
[Fase 1] Identificación y análisis del problema	38. Identificación de los elementos básicos del problema
	39. Representación y formulación del problema
	40. Identificación de factores relevantes
	41. Identificación de requisitos de conocimiento
	42. Identificación de información relevante
[Fase 2] Identificación de la estrategia y la alternativa	43. Búsqueda de entendimiento y percepción
	44. Integración de información
	45. Identificación de posibles estrategias para la resolución del problema
	46. Elección de la mejor estrategia
	47. Identificación de múltiples criterios
[Fase 3] Actuación guiada por estrategias	48. Priorización de los criterios
	49. Evaluación de alternativas
[Fase 4] Evaluación final	50. Selección de la mejor alternativa
	51. Justificación de la alternativa elegida
	52. Planificación
	53. Ejecución y adopción de acciones correctoras
	54. Evaluación crítico-constructiva del procedimiento
	55. Interpretación y evaluación crítico-constructiva de los resultados

Tabla 3. Indicadores de Subdimensión de Resolución de Problemas

2.2. Elaboración del instrumento

Para el diseño del instrumento se siguieron los trece pasos que se describen a continuación (Figura 1). En primer lugar, se estableció la base teórica para el instrumento (Paso 1), lo que implicó una revisión exhaustiva de la literatura relevante y la construcción de un modelo teórico que sustentara la validez del constructo que se pretendía medir. En segundo lugar, se seleccionaron los atributos que reflejaban el constructo (Paso 2), lo que implicó la elaboración de una propuesta de sistema de indicadores basada en la fundamentación teórica. Esta propuesta fue evaluada por nueve expertos nacionales e internacionales en pensamiento crítico y medición, y se incorporaron las modificaciones sugeridas.

La redacción de los reactivos para evaluar la primera dimensión se realizó mediante la identificación de temas actuales y polémicos, utilizando diversas herramientas como Facebook, Twitter, verificadores de bulos y Google Trends (Paso 3). Algunos ejemplos de estos temas son la legalización del aborto, la marihuana y el impacto de la inmigración. En cuanto a la segunda parte del instrumento, que evalúa las habilidades de resolución de problemas, se seleccionaron ámbitos comunes en los que el estudiantado universitario suele tomar decisiones, como la familia, los estudios, las amistades y la organización de eventos. Cabe destacar que estos temas se utilizaron únicamente como medio para evaluar el pensamiento crítico y no como fin en sí mismos. Además, se realizó una asignación aleatoria entre los temas y los indicadores para minimizar cualquier posible sesgo.

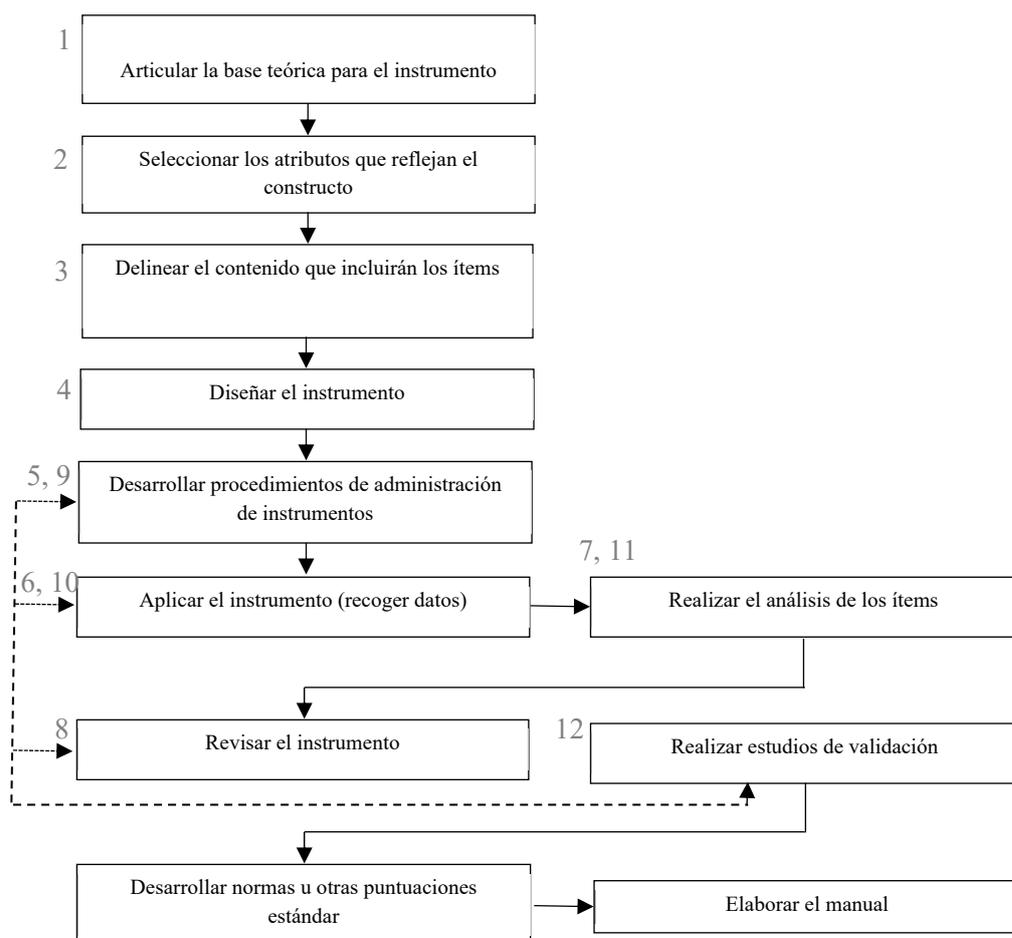


Figura 1. Fases en el diseño y elaboración del instrumento

En relación con el diseño del instrumento (Paso 4), se realizó en dos fases. La primera consistió en el diseño del instrumento preliminar, el cual incluía ítems de respuesta abierta y escrita.

Debido a la extensión del instrumento preliminar, se dividió en dos partes para evaluar la Dimensión 1 (*Análisis y Evaluación de Argumentos*) y la Dimensión 2 (*Resolución de Problemas*). A continuación, se desarrollaron los procedimientos de administración y aplicación del instrumento (Pasos 5 y 6). La muestra estuvo compuesta por 99 estudiantes universitarios (Paso 6). Se contactó directamente con los y las estudiantes para participar en el estudio y se les proporcionó información sobre el mismo.

Posteriormente, se realizó un análisis de las respuestas de las personas participantes utilizando los programas NVivo y Python, aplicando específicamente técnicas de Procesamiento de Lenguaje Natural (NLP) para el análisis del contenido (Paso 7). A partir de este análisis, se realizó una revisión exhaustiva del instrumento preliminar, lo que permitió ajustar y perfeccionar el diseño del instrumento final (Paso 8). Este instrumento final incluyó ítems de respuesta cerrada, desarrollados en función de los resultados obtenidos en la fase preliminar. Al basar el diseño final en el análisis de las respuestas anteriores, se lograron minimizar las limitaciones asociadas a los ítems de respuesta abierta.

El instrumento final, denominado *CritiTest*, se detalla en el siguiente apartado (descripción del instrumento definitivo).

Acto seguido, se repitieron los pasos 5, 6 y 7, que consistieron en el desarrollo de los procedimientos de administración del instrumento final (Paso 9), la aplicación del instrumento (Paso 10) y el análisis de los ítems finales (Paso 11), respectivamente. En este estudio, participaron 5.238 estudiantes de grado de diversas universidades españolas. Finalmente, se realizaron los estudios de validación (Paso 12), los cuales se presentarán en la sección de resultados, dejando los siguientes pasos para futuros trabajos.

2.3. Descripción del instrumento definitivo

CritiTest consta de dos partes para evaluar el pensamiento crítico. La primera parte, titulada “Situaciones Cotidianas”, evalúa la primera dimensión del pensamiento crítico, es decir, el *Análisis y Evaluación de Argumentos*. Con este fin, se presentan 15 situaciones cotidianas con afirmaciones que los encuestados deben evaluar en una escala de 1 a 5, donde 1 representa “Nada o Muy poco de acuerdo” y 5 representa “Totalmente o Muy de acuerdo”.

Cada una de las 15 situaciones se enfoca en una temática actual y socialmente polémica para medir el pensamiento crítico. Estas temáticas incluyen el modelo de estado, aborto, experimentación con seres humanos, legalización de la marihuana, inmigración, economía, derechos de los animales, pena de muerte, alimentos ecológicos, información nutricional, fenómenos paranormales, Dios, corrupción política, bien y mal, y prostitución.

La segunda parte se enfoca en evaluar la capacidad del estudiantado universitario para *Resolver Problemas*. Para ello, se presentan 6 situaciones hipotéticas relacionadas con drogas, salud, trabajo, viajes, familia y organización de eventos, y se pide a los encuestados que evalúen una serie de afirmaciones utilizando la escala de 1 a 5 explicada previamente. En el Anexo puede encontrarse un ejemplo de pregunta de ambas secciones.

En este punto, es importante señalar dos cuestiones relevantes relacionadas con la evaluación del pensamiento crítico. En primer lugar, se recomienda que la formación y evaluación del pensamiento crítico se realice a través de temáticas actuales y socialmente polémicas (Velez y Power, 2020), ya que estas situaciones desafían la capacidad de analizar y evaluar información de manera objetiva y reflexiva, así como de tomar decisiones informadas en contextos complejos.

En segundo lugar, para medir el pensamiento crítico se utilizan ítems de respuesta con escala Likert de 5 grados. Esta escala permite establecer clasificaciones binarias con niveles de convicción, identificando el acierto y el nivel de convicción en la respuesta. Las escalas binarias con niveles de convicción son útiles para evaluar de manera precisa y objetiva la capacidad de una persona para analizar y evaluar información de manera crítica y reflexiva. Estas escalas se basan en preguntas de respuesta sí/no, lo que facilita medir claramente si una persona posee o no una habilidad o característica relacionada con el pensamiento crítico. Además, los niveles de convicción permiten establecer una medida de certeza en cuanto a la precisión de la evaluación, lo que garantiza resultados fiables y válidos.

2.4. Recogida de datos

Tras finalizar el diseño del instrumento, se estableció contacto con los y las líderes de todas las universidades españolas, incluyendo rectores/as, vicerrectores/as, decanos/as,

vicedecanos/as, profesores/as y representantes estudiantiles, con el objetivo de solicitar su colaboración en la recopilación de datos. Quienes expresaron interés en participar recibieron una versión virtual del instrumento para distribuir entre su estudiantado.

2.5. Participantes

En total participaron 5.238 estudiantes, de los cuales el 60,15% correspondió al género femenino, mientras que el 3,41% se identificó como no binario. La mayoría de los y las estudiantes se encontraba en el rango de edad de 17 a 24 años (85,53%), siendo el subgrupo más amplio el de 17 a 20 años (50,17%). En cuanto al año académico, el 47,83% del estudiantado se ubicó en el primer o segundo año, el 47,73% en el tercero o cuarto año y el 4,41% en el quinto o sexto año. Cabe destacar el alto porcentaje de estudiantes en el primer (30,11%) y cuarto (35,28%) año.

Respecto a las calificaciones del expediente académico, se encontró que la nota más frecuente fue un Notable (60,7%), seguido de Suficiente (27,5%), Excelente (9,1%), Sobresaliente (2,3%) e Insuficiente (0,4%). En cuanto al campo de conocimiento, Ciencias Sociales y Jurídicas fue el campo más representado (31,54%), seguido de Ciencias (22,04%), Ingeniería y Arquitectura (16,7%) y Ciencias de la Salud. Las Artes y Humanidades tuvieron la menor representación (14,32%). Por lo que a la zona geográfica se refiere, el estudio contó con la participación de estudiantes de todas las comunidades autónomas de España, siendo las tasas de participación más elevadas en Madrid (21,92%), Andalucía (16,06%), Valencia (11,78%) y Cataluña (9,41%). Estas comunidades autónomas se caracterizan por tener las poblaciones más grandes y el mayor número de universidades.

3. RESULTADOS

La fiabilidad y la validez son aspectos fundamentales en el análisis técnico de un instrumento de medida, pues determinan la calidad de los datos y de las inferencias que se hacen a partir de ellos. La fiabilidad de un instrumento se refiere a su capacidad para medir el constructo de forma estable, precisa o consistente. Por su parte, la validez se refiere al grado en que un instrumento mide lo que pretende medir. Por tanto, la fiabilidad de un instrumento es una condición necesaria pero no suficiente para establecer la validez de sus puntajes. Se expone a continuación el análisis de estas propiedades psicométricas a partir de los datos obtenidos en este estudio.

3.1. Fiabilidad total y por dimensiones

Con el objetivo de evaluar la fiabilidad del instrumento utilizado en este estudio, se calculó el coeficiente Alpha de Cronbach, una medida ampliamente utilizada para evaluar la consistencia interna de los instrumentos de medición debido a su eficacia y aplicabilidad. En concreto, se obtuvo un valor de alpha de 0,9 para escala total, y de 0,8 para las dimensiones de *Análisis y Evaluación de Argumentos* y *Resolución de Problemas* (Tabla 4). Según autores como George y Mallery (2003), Kaplan y Saccuzzo (2017) y Nunnally (1978), los coeficientes obtenidos son buenos a nivel de subdimensión y buenos-excelentes a nivel global. Es importante señalar que, aunque algunos ítems

presentan índices negativos, su inclusión se justifica teóricamente. Además, la eliminación de estos ítems no aumenta significativamente la fiabilidad, ni a nivel global ni de subdimensión, por lo que se decidió mantener la totalidad de los ítems.

	alpha de Cronbach	N ítems
Análisis y Evaluación de Argumentos	0,8	73
Resolución de Problemas	0,8	35
Total	0,9	108

Tabla 4. Coeficiente de fiabilidad total y por dimensiones

3.2. Validez

En cuanto a la validez, los procedimientos utilizados para justificar la validez del instrumento utilizado en este estudio son la fundamentación teórica, el juicio de personas expertas y el análisis de los resultados del estudio piloto.

En primer lugar, se justifica la validez del contenido por la fundamentación teórica en que se sustenta el *Modelo de Pensamiento Crítico* que se presenta en este trabajo. Esta revisión de la literatura justifica la inclusión de las dimensiones utilizadas. En segundo lugar, se realiza un juicio de personas expertas. En concreto, se cuenta con nueve personas expertas, nacionales e internacionales y de diferente dominio: medición, procesamiento del lenguaje natural, pensamiento crítico, análisis y evaluación de argumentos y resolución de problemas. Finalmente, se realiza un estudio piloto y se analizan los resultados a través de métodos de análisis sofisticados que incluyen el análisis estadístico y de contenido a través de procesamientos del lenguaje natural; esto es, a través de una tecnología de *machine learning* que permite interpretar, manipular y comprender el lenguaje humano.

En cuanto a la fundamentación teórica, se analizaron las diferentes definiciones del pensamiento crítico localizadas en la literatura académica, se identificaron aquellas estrechamente relacionadas con habilidades o comportamientos cognitivos específicos susceptibles de ser identificados y observados empíricamente, y se utilizaron para construir el modelo de pensamiento crítico incluido en el estudio.

Este último paso fue especialmente complejo, ya que las definiciones de pensamiento crítico son en su mayoría teóricas y no permiten identificar indicadores claramente medibles u observables. Por esta razón, se consultó documentación específica sobre cada una de las dos dimensiones principales identificadas (*Análisis y Evaluación de Argumentos* y *Resolución de Problemas*) para poder definir las y medirlas de manera más precisa.

Una vez elaborado el modelo de pensamiento crítico, se procedió a realizar un juicio de expertos para validar su contenido. Se seleccionaron 9 expertos de diferentes dominios, como se indicó anteriormente. A cada uno de ellos y ellas se les proporcionó una descripción detallada del modelo de pensamiento crítico y se les pidió que realizaran una evaluación del mismo, aportando sugerencias y comentarios para mejorarlo. Las personas expertas evaluaron el modelo de pensamiento crítico en cuanto a su claridad, relevancia y coherencia. Tras recibir las evaluaciones de las personas expertas, se analizaron sus comentarios y se realizaron las modificaciones pertinentes en el modelo.

A continuación, una vez elaborados los instrumentos preliminares se llevó a cabo un estudio piloto con una muestra de 99 estudiantes de grado de varias universidades españolas. La muestra se seleccionó al azar y se utilizó el programa NVivo 12 y Python 3.10.6 para analizar los ítems preliminares y evaluar el comportamiento de los reactivos en dos fases. En la primera fase, se analizó la frecuencia de respuesta de cada ítem y se interpretaron los resultados. En la segunda fase, se exploraron los datos mediante técnicas de NLP para evaluar el volumen de respuestas que incluyeron términos clave y términos representantes de razonamiento. Se identificó que el 56% de las respuestas aluden a los sustantivos del enunciado, mientras que solo el 27% de las respuestas llevan un término de razonamiento. Además, se aplicó un modelo lingüístico de análisis de sentimientos para identificar la polarización de los sujetos y se determinó que el 66% de las respuestas fueron clasificadas como neutrales, en comparación con un 32% negativas y casi un 1% positivas. Las conclusiones indican que los ítems con poca variabilidad deben ser eliminados y que los ítems de preguntas cerradas deben incluirse en el instrumento debido a su longitud.

4. DISCUSIÓN Y CONCLUSIONES

El pensamiento crítico es un pilar esencial de la educación universitaria, ya que permite al estudiantado enfrentar de manera efectiva desafíos personales, académicos, laborales y sociales (Prat-Sala y van Duuren, 2022). Esto contribuye a la formación de una ciudadanía capaz de tomar decisiones informadas y llegar a conclusiones bien fundamentadas (Aktoprak y Hursen, 2022). En este sentido, el presente estudio ofrece una aportación significativa tanto a la comunidad científica como educativa, al presentar un instrumento diseñado específicamente para evaluar el pensamiento crítico en estudiantes universitarios hispanohablantes. Este instrumento se basa en una definición integral que abarca enfoques filosóficos, psicológicos y educativos, y que además incorpora la dimensión ética del razonamiento. Para su evaluación, se abordan temáticas socialmente controvertidas, asegurando una medición lo más objetiva posible, de acuerdo con el marco teórico adoptado y justificado.

En particular, se define el pensamiento crítico como un proceso intelectual que abarca dos habilidades fundamentales: el *Análisis y la Evaluación de Argumentos*, y la *Resolución de Problemas*. Desde un enfoque lógico-dialéctico, el análisis de argumentos está configurado por las habilidades requeridas para detectar, identificar y examinar las diferentes partes de un argumento, sus relaciones y el principio integrador. Por tanto, su finalidad es reconocer el contenido y estructura del argumento. El contenido se refiere a las proposiciones que expresan las premisas y la conclusión, mientras que la estructura alude a la forma en que las premisas se vinculan entre sí y de manera modelizada con la conclusión. La detección del argumento (Archila et al., 2022), la discriminación entre hechos y opiniones (Heard et al., 2020) o la identificación de la relación lógica entre la premisa y la conclusión (Eemeren y Henkemans, 2016) son ejemplos de habilidades propias de esta categoría. En tal sentido, el análisis del argumento permite identificar y descodificar la información compuesta en el argumento, por lo que facilita la evaluación ético-racional de este (Chatfield, 2022; Hatcher y Possin, 2020).

Por su parte, la evaluación del argumento tiene como objetivo estimar, independientemente del nivel de acuerdo con la conclusión, la fuerza o debilidad con la

que la premisa apoya a la conclusión (Dwyer, 2017; Gul y Akcay, 2020). Por ende, se relaciona con la detección de errores de razonamiento (Heard et al., 2020), la construcción de contraargumentos e hipótesis alternativas (Liu y Stapleton, 2014) y el cuestionamiento e identificación de información adicional necesaria (Wu, 2021). Asimismo, dada la creciente relevancia de la dimensión ética del pensamiento crítico (Dwyer, 2017), la evaluación del argumento contempla también el rechazo de declaraciones que promueven la vulneración de los derechos humanos, como la humillación, la discriminación o la ofensa. La evaluación de la credibilidad de la información (Marttunen et al., 2021), la identificación de falacias relativas a la relevancia, como apelar a una fuente no apropiada (Da San Martino et al., 2020) y la identificación de falsas relaciones causales (Cottrell, 2017) son ejemplos concretos de habilidades propias de esta categoría.

Por lo que a la *Resolución de Problemas* se refiere, esta se relaciona con el reconocimiento y la resolución de problemáticas de manera lógica y sistemática (Aktoprak y Hursen, 2022; Braun et al., 2020; Shavelson et al., 2019). En concreto, en este trabajo y desde una perspectiva psicológica cognitiva, se define como un proceso mental que, para alcanzar su objetivo, sigue cuatro fases. Estas son: (a) identificación y análisis del problema, (b) identificación de la estrategia y la alternativa, (c) actuación guiada por estrategias, y (d) evaluación final. En coherencia, ejemplos de habilidades propias de la resolución de problemas son: identificar los elementos básicos del problema (Dwyer, 2017), sus características (Shanta y Wells, 2020) y requisitos de conocimiento necesarios para su adecuada solución (García Ruiz et al., 2020), reconocer y elegir la mejor alternativa de solución (Halpern, 2014), ejecutar y adoptar acciones correctoras cuando proceda (OECD, 2017), evaluar de forma crítico-constructiva tanto el resultado como el procedimiento (Schoenfeld, 1980; Shavelson et al., 2019), entre otros.

En función de los elementos anteriormente destacados se configuró el sistema de dimensiones, subdimensiones e indicadores del instrumento diseñado y utilizado para recoger los datos del estudio. Por lo que a la fiabilidad del instrumento se refiere, este presenta niveles buenos-excelentes, tanto a nivel global como por dimensiones. Concretamente, redondeando las décimas, el coeficiente de alfa de Cronbach es de 0,9 en la escala total y de 0,8 en sus subdimensiones. Por lo tanto, se posiciona entre los instrumentos de medida del pensamiento crítico del estudiantado universitario con mayor consistencia interna a nivel internacional. En el caso del contexto universitario hispanohablante, hasta la fecha de escritura de este artículo (marzo del 2023) se desconoce la existencia de instrumentos con el mismo propósito (medir el pensamiento crítico del estudiantado universitario español) con índices de consistencia interna igual o superiores al de la herramienta diseñada y aplicada en este estudio. En cuanto a la calidad de los ítems que componen el instrumento, todos ellos han mostrado niveles satisfactorios, pues su eliminación no repercutía positivamente en los índices de fiabilidad. Por ello, se concluye que los ítems que integran el instrumento contribuyen a la precisión con la que el instrumento mide el pensamiento crítico.

En relación con la validez, esta queda avalada por la exhaustividad de la fundamentación teórica, el respaldo de personas expertas en la temática que la apoyan y los análisis realizados. En primer lugar, como se ha indicado previamente, la revisión de la literatura ha dejado en evidencia la ambigüedad en relación con qué se entiende por pensamiento crítico y qué elementos lo componen. Estas declaraciones han sido reiteradamente señaladas en anteriores estudios tales como en los de Ericson (2022), Hatcher y Possin (2020) y Plummer et al. (2022). Esta falta de claridad se debe, como

también ha quedado de manifiesto previamente en la literatura académica, a la complejidad inherente del pensamiento crítico (Archila et al., 2022; Hyytinen et al., 2019). Efectivamente, el pensamiento crítico es un constructo de orden superior de alta complejidad que, a su vez, recibe influencia de otros constructos de difícil conceptualización y medición como son el pensamiento creativo o la metacognición.

En este trabajo, el modelo de pensamiento crítico que se presenta se justifica racionalmente por ser una aportación que no entra en conflicto con el análisis de la naturaleza y elementos de pensamiento crítico destacados por los principales autores en el área. Más aún, es una propuesta que profundiza en las áreas de acuerdo relacionadas con el constructo. Además, esta profundización está en coherencia con los planteamientos tanto de los tres movimientos de pensamiento crítico como de las principales disciplinas que lo estudian; esto es, la filosófica, la psicológica y la educativa. Igualmente, concuerda con los planteamientos de la teoría de la argumentación desde una perspectiva lógico-dialéctica y la resolución de problemas desde una perspectiva eminentemente psicológica-cognitiva.

En cuanto al juicio de personas expertas, la posibilidad de haber involucrado a nueve especialistas, tanto nacionales como internacionales, cuyas áreas de experiencia abarcan la construcción de instrumentos de medición, el análisis de información a través de inteligencia artificial –específicamente en el procesamiento del lenguaje natural para el análisis de información– y el pensamiento crítico, así como en la argumentación y resolución de problemas, ha resaltado dos cuestiones importantes. En primer lugar, se ha evidenciado la influencia que la formación y el área de experiencia tienen en la comprensión del pensamiento crítico. Por ejemplo, las personas expertas en pensamiento crítico, argumentación o resolución de problemas tienden a entenderlo como un constructo genérico, aunque reconocen que existen habilidades cognitivamente más exigentes que otras. Destacan la necesidad de elaborar una definición lo suficientemente completa para abarcar la totalidad (o la mayoría) de las habilidades que caracterizan al pensamiento crítico. Por otro lado, los comentarios de las personas expertas en psicometría y análisis de datos estuvieron enfocados en la obtención y análisis de datos. En concreto, sus recomendaciones buscaban que la definición operativa fuera lo más clara y precisa posible y que se pudiera traducir en conductas concretas e ítems cuyas respuestas de las personas participantes permitieran detectar diferencias entre ellas.

En segundo lugar, esta experiencia ha demostrado la necesidad de diseñar instrumentos de medición en los que se involucre tanto a expertos y expertas en el constructo que se mide como en la medición y análisis de datos. En particular, al involucrar a personas expertas en el constructo que se mide, se asegura una comprensión profunda de los elementos y aspectos clave relacionados con el tema de estudio (Keith, 2010). Estas personas expertas en contenido son capaces de evaluar la relevancia y pertinencia de los elementos que componen el sistema de indicadores, garantizando así que se estén midiendo las variables adecuadas y que se capturen de manera precisa y completa. Asimismo, contar con especialistas en medición y análisis de datos resulta crucial para garantizar la validez y fiabilidad de los instrumentos de medición utilizados (Kline, 2016). Estas personas expertas se enfocan en aspectos de claridad, asegurándose de que los ítems del instrumento sean comprensibles y capaces de capturar las diferencias entre los sujetos evaluados de manera efectiva. Además, su experiencia en análisis de datos contribuye a la correcta interpretación y análisis de los resultados obtenidos. En este sentido, la colaboración entre personas expertas en el constructo y en la medición y análisis de datos

es fundamental, dada la complejidad y la importancia de contar con instrumentos de medición sólidos en la investigación académica.

Por último, el diseño y aplicación de un instrumento integrado por preguntas de respuesta cerrada elaboradas sobre la base de los análisis de respuestas abiertas ha permitido construir un instrumento basado en el razonamiento del propio estudiantado. A su vez, ha permitido reducir la varianza irrelevante del constructo propia de los instrumentos de opción abierta cuyo puntaje requiere de interpretación por parte del evaluador. En esta línea, cabe señalar que, si bien la aplicación de modelos de procesamiento de lenguaje natural es relativamente reciente en el ámbito de la psicometría, esta muestra grandes potencialidades a la hora de diseñar y aplicar instrumentos de medición de constructos complejos como el pensamiento crítico que son aplicados a grandes muestras. Algunas de estas ventajas detectadas en este ámbito son: un contraste más rápido de las hipótesis durante el análisis de datos (por medio de la automatización), la evaluación del lenguaje y su influencia en constructos complejos (mediante de la evaluación de la frecuencia de los términos), y la estandarización de los análisis para diferentes idiomas (mediante la traducción). A modo de ejemplo, si se quiere estudiar la posibilidad de evaluar la calidad de las argumentaciones primeramente es necesario comprobar que existen suficientes muestras de argumentación; esto es, que no se responde de forma monosilábica. Del mismo modo, posibilita análisis que softwares tradicionales no permiten, como el análisis de sentimientos o de riqueza verbal. El primero puede habilitar correlacionar la polaridad (sentimiento positivo o negativo) de un texto con una variable determinada. Por ejemplo, si hay una mayor probabilidad de estar de acuerdo con la afirmación de una persona con su nivel en el constructo. Por su parte, automatizar el análisis del repertorio del vocabulario de una persona puede permitir evaluar si las personas más críticas se expresan con mayor precisión o palabras más originales (palabras menos frecuentes). Asimismo, la traducción de texto basada en modelos de lenguaje permite la valoración de respuestas de diferentes idiomas. En este sentido, pueden estudiarse aspectos como la correlación entre un constructo y el idioma en el que se desarrolla.

En conclusión, el presente estudio ha sentado las bases para una evaluación rigurosa y confiable del pensamiento crítico en el estudiantado universitario hispanohablante. Sin embargo, aún queda mucho por explorar y descubrir en este apasionante campo. Esperamos que este trabajo inspire a quienes investiguen en el futuro a seguir explorando y perfeccionando técnicas y herramientas para medir y fomentar el pensamiento crítico en la educación, contribuyendo así a la formación de una ciudadanía más reflexiva, informada y comprometida con su entorno.

BIBLIOGRAFÍA

- Aktoprak, A. y Hursen, C. (2022). A bibliometric and content analysis of critical thinking in primary education. *Thinking Skills and Creativity*, 44, 101029. <https://doi.org/10.1016/J.TSC.2022.101029>
- Archila, P.A., Molina, J., Danies, G., Truscott De Mejía, A.-M. y Restrepo, S. (2022). Using the Controversy over Human Race to Introduce Students to the Identification and the Evaluation of Arguments. *Science & Education*, 31, 861-892. <https://doi.org/10.1007/s11191-021-00299-8>

- Braun, H.I., Shavelson, R.J., Zlatkin-Troitschanskaia, O. y Borowiec, K. (2020). Performance Assessment of Critical Thinking: Conceptualization, Design, and Implementation. *Frontiers in Education*, 5(September), 1-10. <https://doi.org/10.3389/educ.2020.00156>
- Chatfield, T. (2022). *Critical thinking: Your guide to effective argument, successful analysis and independent study* (J. Seaman (ed.)). Sage.
- Cottrell, S. (2017). *Critical Thinking Skills: Developing Effective Analysis and Argument* (Vol. 13, Issue 1). Palgrave Macmillan.
- Da San Martino, G., Cresci, S., Barrón-CedeñoCede, A., Yu, S., Di Pietro, R. y Nakov, P. (2020). *A Survey on Computational Propaganda Detection*. www.usnews.com/news/national-news/articles/2016-11-14/avoid-these-fake-
- Danczak, S.M., Thompson, C.D. y Overton, T.L. (2020). Development and validation of an instrument to measure undergraduate chemistry students' critical thinking skills. *Chemistry Education Research and Practice*, 21(1), 62-78. <https://doi.org/10.1039/c8rp00130h>
- Dellantonio, S. y Pastore, L. (2021). Ignorance, misconceptions and critical thinking. *Synthese*, 198(8), 7.473-7.501. <https://doi.org/10.1007/s11229-019-02529-7>
- Dwyer, C.P. (2017). *Critical thinking: Conceptual perspectives and practical guidelines*. Cambridge University Press.
- Eemeren, F.H. van y Henkemans, A.F.S. (2016). Argumentation: Analysis and Evaluation. En F.H. van Eemeren y A.F.S. Henkemans. *Argumentation*. Routledge.
- Ennis, R.H. (1985). A Logical Basis for Measuring Critical Thinking Skills. *Educational Leadership*, 43(2), 44-48. <https://pdfs.semanticscholar.org/80a7/c7d4a98987590751df4b1bd9adf747fd7aaa.pdf>
- Ennis, R.H. (2015). Critical Thinking: A Streamlined Conception. En M. Davies y R. Barnett (Eds.). *The Palgrave Handbook of Critical Thinking in Higher Education* (pp. 31-48). Palgrave Macmillan. <https://doi.org/10.1007/978-1-137-37805-7>
- Ericson, J.D. (2022). Mapping the Relationship Between Critical Thinking and Design Thinking. *Journal of the Knowledge Economy*, 13(1), 406-429. <https://doi.org/10.1007/s13132-021-00733-w>
- Frykholm, J. (2021). Critical thinking and the humanities: A case study of conceptualizations and teaching practices at the Section for Cinema Studies at Stockholm University. *Arts and Humanities in Higher Education*, 20(3), 253-273. <https://doi.org/10.1177/1474022220948798>
- García Ruiz, M., Gutiérrez Olivar, V., Rayas-Prince, J.G.R. y Vázquez-Alonso, Á. (2020). Los efectos de la temática socioambiental en las habilidades de pensamiento crítico del futuro profesorado de primaria. *Tecné, Episteme y Didaxis: TED*, 48, 75-90. <https://doi.org/10.17227/ted.num48-12382>
- George, D. y Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update*. Allyn & Bacon.
- Giri, V. y Paily, M.U. (2020). Effect of Scientific Argumentation on the Development of Critical Thinking. *Science & Education*, 29, 673-690. <https://doi.org/10.1007/s11191-020-00120-y>
- Goodsett, M. (2020). Best practices for teaching and assessing critical thinking in information literacy online learning objects. *The Journal of Academic Librarianship*, 46(5), 102163. <https://doi.org/10.1016/j.acalib.2020.102163>
- Gul, M.G. y Akcay, H. (2020). Structuring a new socioscientific issues (SSI) based instruction model: Impacts on pre-service science teachers' (PSTs) critical thinking skills and dispositions. *International Journal of Research in Education and Science*, 6(1), 141-159. <https://doi.org/10.46328/ijres.v6i1.785>

- Halpern, D.F. (2014). *Thought and Knowledge: An Introduction to Critical Thinking* (5th ed.). Psychology Press.
- Hatcher, K. y Possin, D. (2020). Chapter 16 Commentary. En F. Fasko y D. Fair (Ed.), *Critical Thinking and Reasoning. Theory, Development, Instruction, and Assessment* (pp. 298-322). Brill/ Sense Publishing. <https://doi.org/10.4135/9781412991445.d102>
- Heard, J., Scoular, C., Duckworth, D., Ramalingam, D. y Teo, I. (2020). *Critical Thinking: Skill Development Framework*. The Australian Council for Educational Research(ACER). https://www.researchgate.net/publication/345958573_Critical_thinking_Skill_development_framework
- Hyytinen, H., Toom, A. y Shavelson, R.J. (2019). Enhancing Scientific Thinking Through the Development of Critical Thinking in Higher Education. En M. Murtonen y K. Balloo (Eds.). *Redefining Scientific Thinking for Higher Education* (pp. 59-78). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-24215-2_3
- Kageyama, Y., Zamudio, S.Z. y Barton, M. (2022). Incorporation of simulation features to improve higher order thinking skills. *The International Journal of Management Education*, 20(2), 100628. <https://doi.org/10.1016/J.IJME.2022.100628>
- Kaplan, R.M. y Saccuzzo, D.P. (2017). *Psychological testing: Principles, Applications, and Issues* (9th ed.). Cengage Learning.
- Keith, C. (2010). *An Introduction to Psychological Assessment and Psychometrics*. SAGE Publications.
- Kleemola, K., Hyytinen, H. y Toom, A. (2022). Exploring internal structure of a performance-based critical thinking assessment for new students in higher education. *Assessment and Evaluation in Higher Education*, 47(4), 556-569. <https://doi.org/10.1080/02602938.2021.1946482>
- Kline, P. (2016). *A Handbook of Test Construction*. Routledge.
- Liu, F. y Stapleton, P. (2014). Counterargumentation and the cultivation of critical thinking in argumentative writing: Investigating washback from a high-stakes test. *System*, 45(1), 117-128. <https://doi.org/10.1016/J.SYSTEM.2014.05.005>
- Marttunen, M., Salminen, T. y Utraiainen, J. (2021). Student evaluations of the credibility and argumentation of online sources. *Journal of Educational Research*, 114(3), 294-305. <https://doi.org/10.1080/00220671.2021.1929052>
- Nunnally, J.C. (1978). *Psychometric Theory* (2nd ed.). McGraw-Hill.
- O'Leary, M., Reynolds, K., Ling, G., Liu, O.L., Belton, S., O'Reilly, N. y McKenna, J. (2020). Assessing Critical Thinking in Higher Education: Validity Evidence for the Use of the HEIghten™ Critical Thinking Test in Ireland. *Journal of Higher Education Theory and Practice*, 20(12), 115-130. <https://doi.org/10.33423/jhetp.v20i12.3783>
- OECD (2017). *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic, Financial Literacy and Collaborative Problem Solving, revised edition, PISA*. OECD Publishing. http://www.oecd-ilibrary.org/education/pisa-2015-assessment-and-analytical-framework_9789264255425en%0Apapers3://publication/doi/10.1787/9789264255425-en
- Paul, R. y Elder, L. (2019). *A Guide For Educators to Critical Thinking Competency Standards: Standards, Principles, Performance Indicators, and Outcomes With a Critical Thinking Master Rubric*. The Foundation for Critical Thinking. www.criticalthinking.org
- Plummer, K.J., Kebritchi, M., Leary, H.M. y Halverson, D.M. (2022). Enhancing Critical Thinking Skills through Decision-Based Learning Keywords Decision-based learning. *Critical thinking. Innovative Higher Education*, 47, 711-734. <https://doi.org/10.1007/s10755-022-09595-9>

- Prat-Sala, M y van Duuren, M. (2022). Critical Thinking Performance Increases in Psychology Undergraduates Measured Using a Workplace-Recognized Test. *Teaching of Psychology*, 49(2), 153-163. <https://doi.org/10.1177/0098628320957981>
- Rauscher, W. y Badenhorst, H. (2021). Thinking critically about critical thinking dispositions in technology education. *International Journal of Technology and Design Education*, 31(3), 465-488. <https://doi.org/10.1007/s10798-020-09564-3>
- Schoenfeld, A.H. (1980). Teaching Problem-Solving Skills. *The American Mathematical Monthly*, 87, 794-805. <https://doi.org/10.1080/00029890.1980.11995155>
- Shanta, S. y Wells, J.G. (2020). T/E design based learning: assessing student critical thinking and problem solving abilities. *International Journal of Technology and Design Education*, 32(1), 267-285. <https://doi.org/10.1007/s10798-020-09608-8>
- Shavelson, R.J., Zlatkin-Troitschanskaia, O., Beck, K., Schmidt, S. y Marino, J.P. (2019). Assessment of University Students' Critical Thinking: Next Generation Performance Assessment. *International Journal of Testing*, 19(4), 337-362. <https://doi.org/10.1080/15305058.2018.1543309>
- Southworth, J. (2022). Bridging critical thinking and transformative learning: The role of perspective-taking. *Theory and Research in Education*, 20(1), 44-63. <https://doi.org/10.1177/14778785221090853>
- Sutiani, A., Situmorang, M. y Silalahi, A. (2021). Implementation of an Inquiry Learning Model with Science Literacy to Improve Student Critical Thinking Skills. *International Journal of Instruction*, 14(2), 117-138. <https://doi.org/10.29333/iji.2021.1428a>
- Velez, G. y Power, S.A. (2020). Teaching students how to think, not what to think: Pedagogy and political psychology. *Journal of Social and Political Psychology*, 8(1), 388-403. <https://doi.org/10.5964/jspp.v8i1.1284>
- Wu, C. (2021). Training Teachers in China to Use the Philosophy for Children Approach and Its Impact on Critical Thinking Skills: A Pilot Study. *Education Sciences*, 11(5), 206. <https://doi.org/10.3390/educsci11050206>

ANEXO

En el ejemplo presentado en la Tabla 5, se evalúa la capacidad de identificar errores de suficiencia relacionados con la generalización. En concreto, se plantea la identificación de muestras no suficientes (ítem 1) y no representativas (ítem 2) en la encuesta mencionada. Además, se plantea la identificación de los efectos auditivos del debate como una posible fuente de influencia en las respuestas de la audiencia (ítem 3). Se espera que el estudiantado sea capaz de identificar estas posibles fuentes de error en la interpretación de los resultados de la encuesta y evaluar su impacto en la validez de las conclusiones extraídas. Es relevante mencionar que los ítems 1 y 2 son ítems inversos.

AMAIA: Después de este acalorado debate sobre la Monarquía y la República, el 54% de los telespectadores ha manifestado estar a favor de la Monarquía y el 46% restante, de la República. Pero lo más interesante es que el 97% del total coincide en la necesidad de celebrar un referéndum para que sea el pueblo quien decida el modelo de Estado. Por tanto, si hacemos caso a lo que realmente quiere el pueblo español, deberíamos celebrar un referéndum.

Indica tu grado de acuerdo con las siguientes afirmaciones, del 1 al 5, siendo 1 Nada o Muy poco de acuerdo y 5 Totalmente o Muy de acuerdo.

1. La mayoría de los españoles quiere una votación para determinar el modelo de Estado.
 2. Los votantes de la encuesta representan la población española.
 3. Los efectos auditivos, como el tono o el ritmo del habla de los miembros del debate, han podido condicionar la respuesta de la audiencia.
-

Tabla 5. Ejemplos de ítems de la dimensión “Situaciones cotidianas”

Por su parte, un ejemplo de pregunta de la dimensión de Resolución de Problemas se encuentra en la Tabla 6.

Tu mejor amigo está pasando por un momento complicado (...). Después de hablar con él, te reconoce que tiene una fuerte adicción a la cocaína y te pide que le ayudes a buscar un centro de desintoxicación. Después de una exhaustiva búsqueda, estáis entre dos centros:

- I. El primero lo dirige un francés ex-adicto que tiene un par de años más que tu amigo. El 90% de las personas que siguió el tratamiento durante un año logró desintoxicarse de la cocaína.
- II. El segundo lo dirige un terapeuta alemán de mediana edad que nunca ha probado la cocaína pero que tiene formación superior en psicobiología de la adicción a la cocaína. El 30% de las personas que se inscribieron al tratamiento logró desintoxicarse.

Indica tu grado de acuerdo con las siguientes afirmaciones, del 1 al 5, siendo 1 Nada o Muy poco de acuerdo y 5 Totalmente o Muy de acuerdo.

1. El porcentaje de éxito sería un criterio más importante que la nacionalidad a la hora de elegir el centro.
 2. Es más probable que se cure en el primer centro (terapeuta ex-adicto) (ítem inverso).
-

Tabla 6. Ejemplos de ítems de la dimensión “Resolución de Problemas”

En la primera afirmación se espera que los encuestados valoren criterios relevantes para la toma de decisiones. En este caso, el porcentaje de éxito del tratamiento es un criterio más importante que la nacionalidad del director del centro. El objetivo es elegir la alternativa con la mayor probabilidad de éxito para resolver el problema de la adicción a la cocaína. Por lo tanto, el hecho de que el centro sea dirigido por un exadicto francés o un terapeuta alemán es menos importante que la tasa de éxito que se puede obtener en cada uno de ellos.

En la segunda afirmación, que es un ítem inverso, se evalúa la selección de la mejor alternativa. Al respecto, se espera que el estudiante muestre un grado de acuerdo bajo. En primer lugar, porque los porcentajes mostrados no son equiparables. El 90% de éxito se refiere a la tasa de éxito del total de personas que llevan un año en tratamiento, mientras que el 30% se refiere a las personas que se han inscrito en el tratamiento. Por lo tanto, estos porcentajes no son comparables directamente. Además, la experiencia personal del director del primer centro, al ser un ex-adicto, puede estar sesgada y no necesariamente aplicable a otros pacientes. En cambio, el segundo centro, dirigido por un terapeuta con formación en psicobiología de la adicción a la cocaína, es más probable que tenga un enfoque más científico y efectivo para el tratamiento de la adicción, al estar basado en estudios científicos y en un enfoque más generalizado a diferentes pacientes. Por lo tanto, es más probable que el tratamiento en el segundo centro sea más efectivo.