CC BY

ESTUDIOS

Revista **Complutense de Educación**
ISSNe: 1988-2793

EDICIONES COMPLUTENSE

https://dx.doi.org/10.5209/rced.70748

Disposición, habilidades del pensamiento crítico y éxito académico en estudiantes universitarios: metaanálisis

Yordanis Enríquez Canto[1]; Esteban Zapater Ferrer[2]; Giovani Martín Diaz Gervasi[3]

Recibido: Septiembre 2020 / Evaluado: Abril 2021 / Aceptado: Abril 2021

Resumen. Introducción: el pensamiento crítico se halla entre las principales habilidades blandas relacionadas al éxito académico y de alta demanda en el mundo laboral. El objetivo fue hacer un metaanálisis de estudios observacionales sobre la relación entre la disposición, habilidades del pensamiento crítico y el éxito académico en universitarios. Métodos: revisión sistemática con metaanálisis de efectos aleatorios siguiendo las pautas para estudios observacionales. Se realizaron búsquedas utilizando MEDLINE, EMBASE, Scopus y Cochrane Library. Se evaluó la calidad metodológica con la versión de la Escala de Newcastle-Ottawa. Se estimó la heterogeneidad, se utilizó la prueba Q de Cochran, el sesgo de publicación con el gráfico de embudo y la prueba de Begg. Se realizaron análisis de sensibilidad y diagramas de bosque. Resultados: se identificaron 6756 estudios, 32 correspondieron a los criterios de inclusión abarcando un total de 4962 participantes, 30 fueron meta analizados. El pensamiento crítico y el éxito académico estuvieron positivamente correlacionados: r=.26 (IC:95%; .18– .34) (p=.00), con una heterogeneidad elevada (I2=86.5%). En el análisis de subgrupos se observaron diferencias significativas para los siguientes moderadores: especialidad Q=42.86 (p=.00), número de dimensiones del instrumento Q=31.83 (p=.00), instrumentos Q=56.01 (p=.00), dimensiones evaluadas Q=25.09 (p=.00). Discusión: la magnitud de la correlación resultó débil y afectada por moderadores como la especialidad de estudio, el instrumento empleado, la evaluación de habilidades versus disposiciones y la cantidad de dimensiones consideradas.
Palabras clave: revisión sistemática; pensamiento; sentido crítico; rendimiento; estudiante.

[en] Disposition, critical thinking skills and academic success in university students: meta-analysis

Abstract. Introduction: critical thinking is among the main soft skills related to academic success and highly demanded in the workplace. The objective was to make a meta-analysis of observational studies on the relationship between disposition, critical thinking skills and academic success in university students. Methods: a systematic review with meta-analysis of random effects following the guidelines for observational studies. Searches were conducted using MEDLINE, EMBASE, Scopus and Cochrane Library. The methodological quality was evaluated with the Newcastle-Ottawa Scale version. Heterogeneity was estimated, Cochran's Q test was used, the publication bias with the funnel plot and Begg's test. Were performed sensitivity analyzes and forest plot diagrams. Results: 6756 studies were identified, 32 corresponded to the inclusion criteria with 4962 participants, 30 were meta-analyzed. Critical thinking and academic success were positively correlated: r=.26 (CI: 95%; .18-.34) (p=.00), with a high heterogeneity (I2=86.5%). In the subgroup analysis, significant differences were observed for the following moderators: specialty Q=42.86 (p=.00), number of dimensions Q=31.83 (p=.00), instruments Q=56.01 (p=.00), dimensions evaluated Q=25.09 (p=.00). Discussion: the magnitude of the correlation was weak and affected by moderators such as the study specialty, the instrument used, the evaluation of skills versus dispositions and the number of dimensions evaluated.
Keywords: systematic review; thinking; critical sense; achievement; students.

Sumario. 1. Introducción. 2. Método. 3. Resultados. 4. Discusión. 5. Referencias bibliográficas.

Cómo citar: Enríquez Canto, Y.; Zapater Ferrer, E.; Diaz Gervasi, G. M. (2021). Disposición, habilidades del pensamiento crítico y éxito académico en estudiantes universitarios: metaanálisis. *Revista Complutense de Educación, 32*(4), 525-536.

[1] Facultad de Ciencias de la Salud. Universidad Católica Sedes Sapientiae (Perú). E-mail: yenriquez@ucss.edu.pe ORCID: orcid.org/0000-0003-0449-5015
[2] Facultad de Ciencias de la Salud. Universidad Católica Sedes Sapientiae (Perú). E-mail: researchpoint21@gmail.com ORCID: orcid.org/0000-0003-3641-1146
[3] Facultad de Ciencias de la Salud. Universidad Católica Sedes Sapientiae (Perú). E-mail: gdiaz@ucss.edu.pe ORCID: orcid.org/0000-0002-2505-2689

1. Introducción

Entre los actuales objetivos de aprendizaje en la educación universitaria se encuentra el desarrollo del pensamiento crítico (PC) (Comer et al., 2019; Partido y Soto, 2019). Este se halla entre las principales habilidades blandas con alta demanda en un mundo laboral marcado por el requerimiento de personas con capacidades en el procesamiento de información, evaluación de evidencias y razonamiento argumentativo (Rebele y Pierre, 2019).

La educación universitaria es un proceso multidimensional y desafiante tanto para los docentes como para los estudiantes. A fin de formar profesionales de elevada calidad y a la altura de los retos de su futura profesión los educandos deben desarrollar competencias como recopilar y manejar datos, ponderar argumentos y sustentar sus tesis mediante razones adecuadas (Afshar y Movassagh, 2014; AkbariLakeh et al., 2018; Shirazi y Heidari, 2019). Algunos de los desafíos educativos en el contexto universitario incluyen el pronóstico e identificación de los estudiantes con propensión a presentar dificultades académicas (D'Alessio et al. 2019). Son numerosos los estudios que han evaluado entre sus predictores las características socioeconómicas (Adamcik et al., 1996; Baberi y Bozorgi, 2016; Denial y Pitcher, 2007), el sexo (Bayat, 2014; Emir, 2009; Gadzella et al., 2002) o el grupo étnico (Ip et al., 2000; Nornoo et al., 2017; Whitney et al., 2016). No obstante, la investigación existente muestra algunos vacíos puesto que el éxito académico estaría condicionado por factores de diversa índole.

El PC se perfila entonces como un factor clave tanto para el logro académico como para un futuro desempeño profesional de calidad una vez culminado el recorrido universitario. Asimismo, la diversidad de definiciones hace que no sea de fácil delimitación conceptual (Kuhn, 2019; Poce et al., 2019). Para efectos del estudio se consideró la propuesta de Cottrell (2017), es decir, la capacidad de identificar un problema, seleccionar y evaluar la información pertinente, reconocer supuestos, formular hipótesis apropiadas y sacar conclusiones válidas e inferencias críticas. El PC ha sido tematizado a partir de dos dimensiones: disposiciones y habilidades. Las primeras son relativas a propensiones y características de la personalidad vinculadas con la curiosidad, la confianza en el razonamiento, la flexibilidad y apertura mentales (Facione et al., 2017). Por su parte, las habilidades del PC involucran un conjunto de capacidades básicas, entre ellas el análisis, evaluación, interpretación, explicación y autorregulación (Comer et al., 2019).

Entre los instrumentos de mayor uso para evaluar las disposiciones y habilidades del PC se encuentran: el *California Critical Thinking Disposition Inventory* (CCTDI), *California Critical Thinking Skills Test* (CCTST), *Watson-Glaser Critical Thinking Appraisal* (WGCTA) y el *Health Science Reasoning Test* (HSRT) por solo mencionar algunos. Son escalas que resultan fácilmente administrables empleando tiempos mínimos. En la evaluación de cinco habilidades el WGCTA propone 80 ítems generando un puntaje único. Por su parte, el CCTST es un cuestionario de respuesta múltiple de 34 ítems que abarcan nueve sub-escalas, considerando las habilidades adecuadas a contextos donde es importante resolver problemas y tomar decisiones. En la misma línea, el HSRT ha sido diseñado en modo específico para estudiantes de Ciencias de la Salud mediante preguntas de respuesta múltiple (33 ítems) con un puntaje único y valores para cinco sub-escalas. Mientras que en la evaluación de las disposiciones del PC el CCTDI evalúa las siguientes: búsqueda de la verdad, mentalidad abierta, tendencias analíticas, tendencias sistemáticas, curiosidad, madurez cognitiva, entre otras. El cuestionario propone una puntuación global y siete sub-escalas a través de 75 ítems.

En la educación superior universitaria tanto el desarrollo del PC como el éxito académico constituyen resultados educativos de alto valor (Rebele y Pierre, 2019). La relación entre ambos puede estar influenciada por diversos factores. Algunas características sociodemográficas como el sexo (Emir, 2009; Gadzella et al., 2002) y el origen étnico (Nornoo et al., 2017; Whitney et al., 2016) han sido analizadas en la literatura, así como la carrera de especialización o área de estudio con resultados discrepantes (Fong et al., 2017; Ross et al., 2013). Estudios evidenciaron que estudiantes de Física, Matemática e Ingeniería reportaron niveles más altos de PC respecto a estudiantes de Humanidades (Fong et al., 2017; Franco et al., 2015). Manalo y colaboradores (2013) a partir de una clasificación en disciplinas prácticas como Educación, Enfermería y Negocios y no prácticas como Historia, Idiomas y Psicología hallaron que los estudiantes del segundo grupo reportaron mayores niveles en las disposiciones del PC.

Por otro lado, en cuanto al avance dentro del recorrido universitario igualmente se han encontrado divergencias. Huhn y Parrott (2017) reportaron ausencia de diferencias significativas entre los estudiantes de los últimos años. Mientras que Ghanizadeh (2017) señaló diferencias entre los estudiantes del primer año universitario comparados con los que se encuentran en sus últimos años de estudio.

La situación descrita evidencia resultados contradictorios entre los factores que pudieran modular la relación entre el PC y el éxito académico. De acuerdo con esto se propuso realizar una revisión sistemática de la literatura sobre este tópico. Asimismo, el objetivo propuesto fue desarrollar un metaanálisis acerca de la relación entre la disposición, habilidades del pensamiento crítico y el éxito académico en estudiantes universitarios. Considerando este objetivo las preguntas de investigación que orientaron el metaanálisis fueron: ¿cuál es la magnitud general del tamaño del efecto promedio de la correlación entre las medidas de disposición, las habilidades de pensamiento crítico con el rendimiento académico en estudiantes universitarios, a partir de estudios observacionales?; ¿qué variables moduladoras (especialidad de estudio, instrumento estandarizado, dimensión evaluada, medida de desempeño) tienen mayor influencia en el tamaño del efecto promedio?

2. Método

2.1. Diseño y criterios para la inclusión de los estudios

Revisión sistemática y metaanálisis realizados siguiendo las pautas para metaanálisis de estudios observacionales (MOOSE, por sus siglas en inglés) (Stroup et al., 2009) considerando, igualmente, los elementos de informe preferidos para protocolos de revisiones sistemáticas y metaanálisis (PRISMA, por sus siglas en inglés) (Moher et al., 2015).

Para la revisión fueron considerados elegibles:

a) Estudios observacionales que indagaron en la asociación entre la disposición y/o las habilidades del pensamiento crítico y el éxito académico de estudiantes universitarios de pregrado.
b) Los estudios podían incluir estudiantes universitarios de una o varias especialidades.
c) Con relación a la medición de la disposición al pensamiento crítico, fueron considerados elegibles estudios que hubiesen evaluado mediante el CCTDI. Mientras que para las habilidades del pensamiento crítico fueron candidatos aquellos que las midieron a través del CCTST, CCTT, HSRT, WGCTA o el PENCRISAL.
d) El éxito académico pudo haber sido cuantificado usando calificaciones en pruebas o cursos específicos, o con el promedio ponderado de calificaciones (GPA, por sus siglas en inglés).

2.2. Método de búsqueda bibliográfica

Se realizaron búsquedas sistemáticas y automatizadas en MEDLINE (desde el inicio de la base de datos), EMBASE (desde el inicio de la base de datos), Scopus, Cochrane Library. Igualmente, se realizó una búsqueda complementaria en Google Scholar. Asimismo, la búsqueda fue ampliada teniendo en cuenta la lista de referencias de los estudios identificados que midieron la relación entre la disposición y/o las habilidades del pensamiento crítico y el éxito académico en universitarios, incluidas las revisiones sistemáticas publicadas previamente sobre el tema. La búsqueda se realizó en idioma español e inglés sin limitaciones temporales, siendo la última fecha de búsqueda sistemática el 30 de mayo del 2020.

La metodología de búsqueda consideró el uso de un filtro para identificar los estudios utilizando los siguientes descriptores en español: *CCTDI, Inventario de disposiciones del pensamiento crítico de California, CCTST, Prueba de habilidades de pensamiento crítico de California, WGCTA, Prueba de pensamiento crítico de Watson y Glaser, CCTT, Prueba de pensamiento crítico de Cornell, HSRT, Prueba de razonamiento en Ciencias de la Salud, PENCRISAL, Prueba de pensamiento crítico Salamanca y rendimiento académico, éxito académico, logro académico*.

Así como en inglés bajo los siguientes términos: *CCTDI, California Critical Thinking Disposition Inventory, CCTST, California Critical Thinking Skills Test, WGCTA, Watson-Glaser Critical Thinking Appraisal, CCTT, Cornell Critical Thinking Test, HSRT, Health Science Reasoning Test, PENCRISAL, Critical Thinking of Salamanca* y *academic performance, academic success, academic achievement*. Para discriminar en la selección de las fuentes documentales se utilizaron los operadores booleanos: AND, NOT, OR, XOR.

2.3. Procedimiento

Para el análisis estadístico, la determinación del riesgo de sesgo de publicación y la elaboración de los diagramas de bosque (*forest plot*), fue utilizado el *software* Stata 15®.

Fueron realizadas búsquedas de los artículos por dos investigadores en forma paralela e independiente. Se excluyeron los estudios en los que el título y el resumen no resultaban relevantes al objetivo del metaanálisis. En los casos en que no hubo acuerdo se consensuaron con el tercer investigador.

Posteriormente, se realizó un análisis de los artículos seleccionados, lo que permitió extraer la información más relevante para identificar los datos esenciales de cada artículo. Se consideró para ello: las características del estudio (por ej. diseño); instrumentos con que fueron medidas las disposiciones, habilidades del pensamiento crítico; modalidad de cuantificación del éxito académico; número de estudiantes universitarios que participaron en el estudio, edad, sexo, carrera de estudio. Así como la información necesaria para evaluar el riesgo de sesgo.

De forma independiente los datos de cada artículo fueron extraídos por dos revisores utilizando una ficha estructurada de datos. Las discrepancias fueron solucionadas con la intervención del tercer investigador.

2.4. Análisis de la calidad metodológica

Dos revisores evaluaron los estudios que cumplieron con los criterios de inclusión considerando el riesgo de sesgo de forma independiente. Se empleó la versión de la Escala Newcastle-Ottawa (NOS, por sus siglas en inglés) (Wells et al., 2016) para evaluar la calidad de los estudios. La versión modificada incluyó la respuesta a una pregunta en la valoración de cada elemento: “Sí” (bajo riesgo de sesgo); “No” (alto riesgo de sesgo) y “Poco claro” (incertidumbre sobre el posible sesgo) respecto a cinco dominios. Estos fueron: tasa de respuesta, tipo de

muestreo, modo de administración de la prueba, uso de instrumentos validados y análisis estadístico. Además, a cada componente del dominio se le asignó 1 por bajo riesgo de sesgo y 0 por alto riesgo de sesgo o poco claro. La puntuación de calidad total para cada estudio pudo variar de 0 a 5: bajo riesgo de sesgo (≥3 puntos) o alto riesgo de sesgo (<3 puntos). Ver en el Anexo 1 mayores detalles sobre la versión modificada de la NOS y de la evaluación de la calidad metodológica.

2.5. Síntesis cuantitativa de datos

Después de la realización de la síntesis narrativa se procedió a la síntesis cuantitativa. El índice del tamaño del efecto analizado consideró el coeficiente *r* (Pearson o Spearman) que indicó la magnitud y la dirección de la correlación entre la disposición al pensamiento crítico y/o sus habilidades con el éxito académico. Bajo el presupuesto de que en un metaanálisis de correlaciones es más preciso el uso en la escala de Fisher *z* que las correlaciones reales (Deeks et al., 2019). Los datos fueron primero convertidos a la escala Fisher *z*, y la varianza de *z* se estimó como V=1/(n-3). Posteriormente, los valores agrupados de Fisher *z* con sus intervalos de confianza (IC) al 95% se transformaron nuevamente en valores de correlación (*r*).

Se utilizó el modelo de efectos aleatorios DerSimonian-Laird (Deeks et al., 2019) para obtener una estimación de síntesis del valor promedio de Fisher *z* entre los estudios. Para interpretar la magnitud del efecto se asumió que el coeficiente de correlación en el orden de .10 es débil, .30 es moderada y de .50 es fuerte (Cohen, 2013).

Se tuvo en consideración para la determinación de la heterogeneidad el valor de I^2. Para el análisis de subgrupos se consideró como criterio para su realización la presencia de al menos tres estudios por subgrupo. Se compararon: 1) grupos de especialidades consideradas (Ciencias de la Salud, otras carreras y grupos mixtos de estudiantes); 2) estudios que medían una o dos dimensiones del pensamiento crítico; 3) instrumentos que miden las dimensiones del pensamiento crítico; 4) medición de las habilidades *versus* la disposición independientemente del instrumento empleado; 5) estudios de diseño transversal *versus* estudios longitudinales; 6) medición del éxito académico a través del GPA *versus* otros indicadores de rendimiento (por ej. notas de cursos). Se estimaron diferencias estadísticas significativas entre subgrupos utilizando la prueba *Q* de Cochran (*Chi² test*). Los valores estadísticamente significativos de *Q* indicaron que el promedio de los tamaños de los efectos variaba entre los subgrupos de las variables moderadoras más de lo previsto solo por el error de muestreo (Deeks et al., 2019). A fines del análisis cuantitativo se consideraron por separado los datos de los estudios que evaluaron tanto las disposiciones como las habilidades del PC.

Se determinó el sesgo de publicación con la prueba de Egger y la inspección visual del correspondiente gráfico de embudo. Igualmente, se aplicó la prueba de Begg. Finalmente, se realizó un análisis de sensibilidad a fin de estudiar la influencia de cada uno de los estudios en la estimación global del efecto y, por lo tanto, la estabilidad de la medida obtenida.

3. Resultados

Los resultados de la búsqueda y del proceso de selección de los estudios se evidencian en la figura 1. La estrategia de búsqueda localizó 6756 estudios después de eliminar los duplicados. Mientras que fueron 32 los que cumplieron con los criterios de inclusión abarcando publicaciones desde 1987 al 2020. El acuerdo entre los revisores para su selección fue del 94% y del 96% para la extracción de datos.

La tabla 1 muestra las características principales de los 32 estudios incluidos con un total de 4962 participantes. El 72% (23) de los estudios se realizó en los Estados Unidos, mientras que otros cinco se realizaron en Irán (16%). Fueron incluidos sujetos de ambos sexos entre los 18 y 54 años, de ellos el 56% fueron mujeres. En su mayoría las investigaciones consideraron estudiantes de las especialidades de Ciencias de la Salud (Kennison, 2006; Miller, 1992; Miller et al., 1993; Pardamean, 2012; Scott y Markert, 1994; Tiessen, 1987; Vendrely, 2007; Williams et al., 2006). Mientras que, en la medida del pensamiento crítico principalmente se evaluaron sus habilidades. Tres estudios (AkbariLakeh et al., 2018; Emir, 2009; Ip et. al, 2000) evaluaron exclusivamente las disposiciones. Por último, en los estudios incluidos, los resultados para el rendimiento académico fueron mayormente promedios ponderados de calificaciones de los cursos. Mientras que en un grupo más pequeño se utilizaron exámenes individuales (Emir, 2009; Bayat, 2014; Nornoo et al., 2017; Whitney et al., 2016).

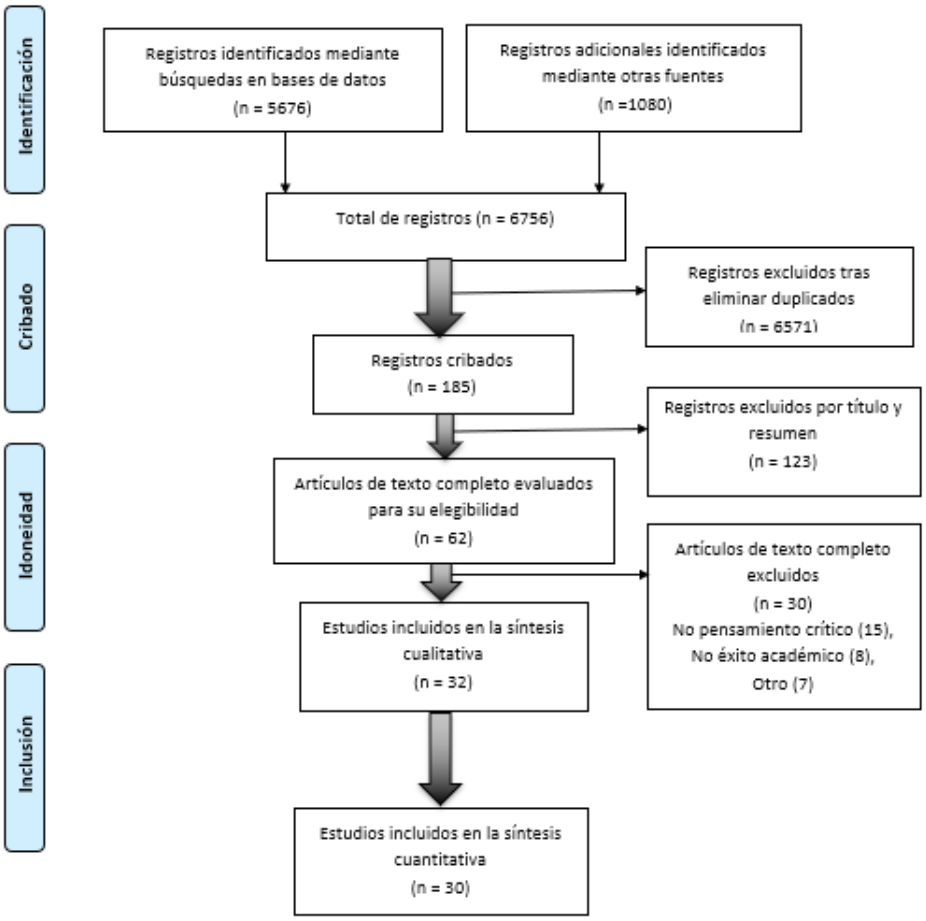


Figura 1. Diagrama de flujo para la selección de artículos

Tabla 1. Información detallada para cada estudio primario

Autor, año y país	Número de participantes	Edad*	Carrera (s)	Criterios de inclusión
Tiessen[a] (1987) – E.E.U.U.	150	21.29	Enfermería	Estudiantes de pregrado del programa de bachillerato en enfermería, que hayan culminado su cuarto año de estudios.
Miller[a] (1992) – E.E.U.U.	137	-	Enfermería	Estudiantes de pregrado del programa de enfermería.
Miller et al.[a] (1993) – E.E.U.U.	196	-	Medicina	Estudiantes de pregrado del programa de medicina.
Scott y Markert[a] (1994) – E.E.U.U.	92	-	Medicina	Estudiantes de pregrado del programa de medicina.
Adamcik et al.[a] (1996) – E.E.U.U.	78	-	Farmacia	Estudiantes de pregrado del último año de farmacia.
Jenkins [a] (1998) – E.E.U.U.	95	-	Auditoría	Estudiantes de pregrado de dos semestres y cuatro secciones de auditoría.
Ip et al.[b] (2000) – China	122	20	Enfermería	Estudiantes que hayan culminado su 1er, 2do, o 3er año del programa de pregrado.
Gadzella et al.[a] (2002) – E.E.U.U.	114	19 – 54	Educación	Estudiantes de educación matriculados en un curso de psicología educativa.
Kidd y Latif [a, b] (2003) – E.E.U.U.	80	-	Farmacia	Estudiantes de pregrado del programa de farmacia.
Miller [a, b] (2003) – E.E.U.U.	102	21.9±4.3	Farmacia	Estudiantes del cuarto año de pregrado del programa de farmacia.
Williams et al. [a, b] (2003) – E.E.U.U.	207	24.0±5.1	Odontología	Estudiantes de primer año de pregrado del programa de odontología.
Kennison [a] (2006) – E.E.U.U.	57	25	Enfermería	Estudiantes de pregrado del último año del programa de enfermería.
Williams et al.[a,b] (2006) – E.E.U.U.	78	25.4±6	Odontología	Estudiantes de pregrado del programa de odontología de tres universidades.
Denial y Pitcher [a,b] (2007) – E.E.U.U.	78	-	Optometría	Estudiantes del tercer año de pregrado del programa de optometría.
El Hassan y Madhum [a] (2007) – Líbano	273	-	Artes Ciencias Ingeniería Negocios Enfermería	Estudiantes de pregrado de las diferentes facultades y semestres.

Autor, año y país	Número de participantes	Edad*	Carrera (s)	Criterios de inclusión
Turner y Williams [a] (2007) – E.E.U.U.	418	-	Educación	Estudiantes de pregrado pertenecientes a alguna de las 12 secciones de un curso de desarrollo humano. Estudiantes que desarrollaron el examen de vocabulario académico pre y post test.
Vendrely [a] (2007) – E.E.U.U.	42	28.77±3.89	Terapia Física	Estudiantes de pregrado del programa de terapia física.
Emir [b] (2009) – Turquía	279	21 – 24	Ed. primaria Ed. de ciencias naturales Ed. de matemáticas Ed. de estudios sociales Ed. de idioma turco	Estudiantes de pregrado matriculados en el semestre académico del estudio.
Pardamean [a] (2012) – E.E.U.U.	98	-	Odontología	Estudiantes de pregrado matriculados en el programa de odontología.
Afshar et al.[a] (2014) – Irán	76	18 – 25	Lengua inglesa	Estudiantes de pregrado del programa de Lengua inglesa.
Bayat [a] (2014) – Turquía	181	-	Ed. de ciencias naturales Ed. de idioma turco Ed. de estudios sociales Ed. de lengua inglesa Ed. preescolar Ed. primaria	Estudiantes de pregrado de segundo año que estudien en los diferentes departamentos de la facultad de educación en el periodo de primavera 2010-2011.
Tsai [a] (2014) – E.E.U.U.	439	-	Odontología	Estudiantes de tercer y cuarto año de pregrado del programa de odontología. Estudiantes que aceptaron participar y se inscribieron en una de las diez escuelas de odontología acreditadas incluidas en el estudio.
Schwanz y McIlreavy [a] (2015) – E.E.U.U.	132	19.30±2.12	-	Estudiantes de pregrado matriculados en el curso de introducción a la psicología.
Baberi y Bozorgi [a] (2016) – Irán	265	-	Psicología	Estudiantes de pregrado matriculadas en el programa de psicología.
Whitney et al.[a,b] (2016) – E.E.U.U.	71	-	Odontología	Estudiantes de pregrado del programa de odontología de las universidades seleccionadas.
Ghanizadeh [a] (2017) – Irán	196	22.5±3.06	Literatura inglesa Ed. de lengua inglesa Traducción inglesa	Estudiantes de pregrado matriculados en las carreras y alguna de las tres universidades del estudio.
Huhn y Parrot [a] (2017) – E.E.U.U.	178	-	Terapia física	Estudiantes que autorizaron a la Federación de Juntas Estatales de Terapia Física para publicar sus puntajes NPTE.
Nornoo et al.[a] (2017) – E.E.U.U.	234	H:27±6 M: 28±5	Farmacia	Estudiantes de pregrado matriculados en el programa de farmacia.
AkbariLakeh et al.[b] (2018) – Irán	50	-	Enfermería	Estudiantes de pregrado del último año del programa de enfermería.
Comer et al.[a] (2019) – E.E.U.U.	260	24.5 (3.9)	Farmacia	Estudiantes de pregrado matriculados en el programa de farmacia.
Partido y Soto [a,b] (2019) – E.E.U.U.	45	**2018:** 22.03±3.65 **2017:** 22.00±0.63	Odontología	Estudiantes de pregrado del año 2017 o 2018. Capacidad de cumplir con el protocolo de estudio.
Shirazi y Heidari [a] (2019) – Irán	139	21.88±2.09	Enfermería	Estudiantes de pregrado matriculados en el programa de enfermería.
a: Estudio que evaluó habilidades del pensamiento crítico, b: Estudio que evaluó disposiciones del pensamiento crítico. * Medias y/o desviación estándar (±). Última fecha de búsqueda sistemática: 30 de mayo del 2020.				

Por su parte, en la evaluación de la calidad metodológica de los 32 estudios incluidos hubo ocho (25%) (Adamcik et al., 1996; Bayat, 2014; El Hassan y Madhum, 2007; Gadzella et al., 2002; Schwanz y McIlreavy, 2015; Scott y Markert, 1994; Williams et al., 2006; Whitney et al., 2016) con alto riesgo de sesgo para el dominio tasa de respuesta, siete estudios (21.8%) (Afshar y Movassagh, 2014; Gadzella et al., 2002; Kidd y Latif, 2003; Miller, 2003; Pardamean, 2012; Williams et al., 2003; Williams et al., 2006) en el tipo de muestreo y seis estudios (18.7%) (Adamcik et al., 1996; Facione et al., 2017; Ghanizadeh, 2017; Huhn y Parrott, 2017; Shirazi y Heidari, 2019; Whitney et al.,

2016) en el modo de administración de la prueba de pensamiento crítico. Mientras que para el análisis estadístico resultaron tres (9.3%) con alto riesgo de sesgo (Denial y Pitcher, 2007; Miller et al., 1993; Turner y Williams, 2007) y para el uso de instrumentos validados un solo estudio (3.1%) (Comer et al., 2019). Se encontró un riesgo de sesgo alto o poco claro en cuatro áreas o más para dos estudios (6.2%) (Denial y Pitcher, 2007; Turner y Williams, 2007). Fueron siete (21.8%) (AkbariLakeh et al., 2018; Baberi y Bozorgi, 2016; Denial y Pitcher, 2007; Emir, 2009; Jenkins, 1998; Tsai, 2014; Turner y Williams, 2007) los estudios evaluados con al menos dos dominios con riesgo de sesgo poco claro.

La figura 2 evidencia que los dominios con mayor presencia de riesgo de sesgo fueron la tasa de respuesta y el tipo de muestreo. Dos estudios (Denial y Pitcher, 2007; Turner y Williams, 2007) no fueron considerados en el metaanálisis, ya que no alcanzaron el mínimo de puntuación requerida después de la evaluación de la calidad metodológica.

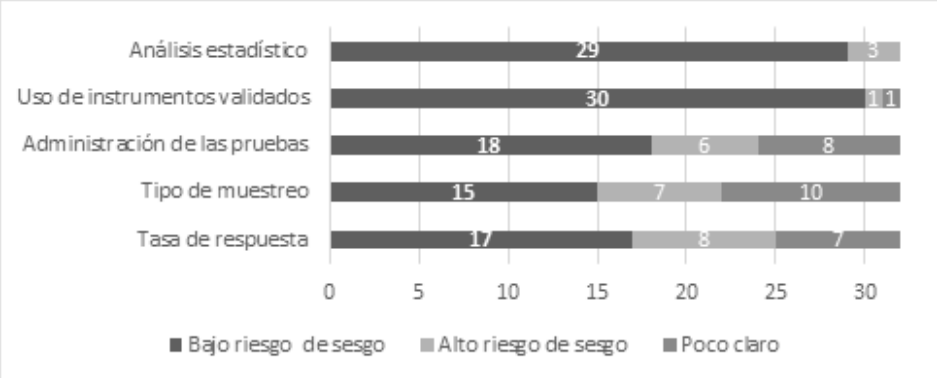


Figura 2. Resumen de la calidad metodológica de la evidencia

En la figura 3 se muestra, con un diagrama de bosque el metaanálisis de 30 estudios, a partir del modelo de efectos aleatorios. El pensamiento crítico tuvo una correlación positiva con el éxito académico, r=.26 (IC:95% .18– .34) (p=.00), con una heterogeneidad elevada (I^2=86.5%) entre los estudios.

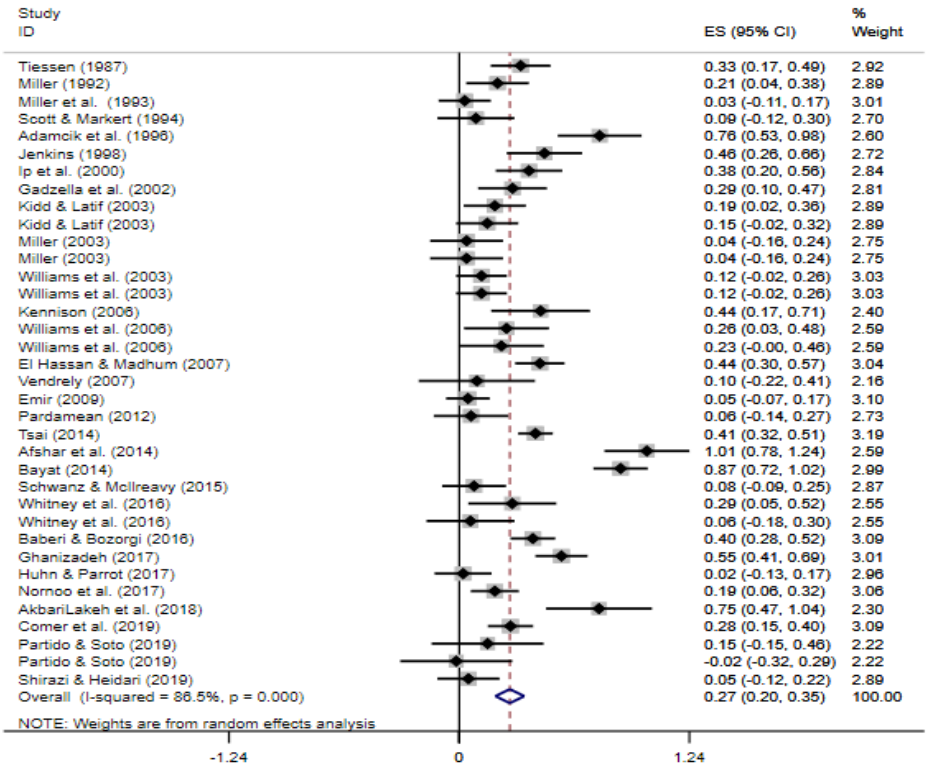


Figura 2. Resumen de la calidad de la evidencia mediante diagrama de bosque

Asimismo, la heterogeneidad fue τ2=.04; χ^2=260.22, df=35 (p= .00). Mientras que la prueba de efecto general resultó estadísticamente significativa z = 6.90 (p= .00).

Posteriormente, teniendo en consideración este resultado de heterogeneidad, se procedió a realizar el análisis de subgrupos. A tal fin, se consideraron separadamente los datos de los estudios que evaluaron tanto las disposiciones, como las habilidades del PC en su correlación con el éxito académico. En la tabla 2 se observan diferencias signi-

ficativas *Q*=42.86 (*p*=.00) con un tamaño del efecto promedio menor en los sujetos que provenían de facultades de Ciencias de la Salud (*r*=.21) respecto a los de otras facultades (*r*=.51). Igualmente, hubo diferencias significativas *Q*=31.83 (*p*=.00) con un tamaño del efecto promedio mayor en los estudios que midieron una dimensión del pensamiento crítico (*r*=.34) respecto a los que evaluaron ambas (*r*=.14). Del mismo modo, se verificaron diferencias significativas *Q*=56.01 (*p*=.00) con un tamaño del efecto promedio mayor en los que emplearon el WGCTA (*r*=.40) mientras que fue menor con el uso del HSRT (*r*=.13).

Tabla 2. Análisis de los moderadores categóricos de la relación entre pensamiento crítico y éxito académico

Subgrupos	k	ES	p	IC 95%	I^2	Q de Cochran
Especialidad						42.86 (*p*=.00)
Ciencias de la Salud	27	.21	.00	.14– .27	73.2%*	
Otras	6	.51	.00	.21– .79	95.1%*	
Mixtos	3	.36	.00	.10– .61	88.8%*	
Nº de dimensiones						31.83 (*p*=.00)
Uno	18	.34	.00	.23– .44	89.6%*	
Dos	12	.14	.00	.08– .19	.00%	
Instrumentos						56.01 (*p*=.00)
WGCTA	10	.40	.00	.23– .57	90.6%*	
CCTDI	8	.14	.00	.06– .22	39.2%	
CCTST	13	.31	.00	.17– .44	87.4%*	
HSRT	5	.13	.01	.02– .24	9.05%	
Diseño						10.23 (*p*=.00)
Transversal	21	.30	.00	.19– .40	88.8%*	
Longitudinal	15	.24	.00	.12– .34	80.0%*	
Dimensiones evaluadas						25.09 (*p*=.00)
Habilidades	28	.31	.00	.21– .40	87.9%*	
Disposición	8	.14	.00	.06– .22	39.2%	
Éxito académico						.20 (*p*=.65)
GPA	31	.27	.00	.19– .34	83.1%*	
Otros	5	.29	.07	-.30– .61	95.1%*	

K: número de estudios en el subgrupo; ES: tamaño del efecto; IC: intervalo de confianza al 95%; I^2: heterogeneidad; * $p \leq .05$; WGCTA: *Watson-Glaser Critical Thinking Appraisal*; CCTDI: *California Critical Thinking Disposition Inventory*; CCTST: *California Critical Thinking Skills Test*; HSRT: *Health Science Reasoning Test*; GPA: promedio ponderado de calificaciones.

Por otra parte, resultaron significativas las diferencias *Q*=25.09 (*p*=.00) en el tamaño del efecto promedio mayor en la medición de habilidades del pensamiento crítico (*r*=.31) comparadas con un tamaño menor del efecto promedio de las disposiciones del pensamiento crítico (*r*=.14). Finalmente, se compararon los efectos del pensamiento crítico en los puntajes obtenidos en los exámenes de cursos (k=5) con el GPA (k=31) hallándose diferencias estadísticamente no significativas (*Q*=.20, *p*=.65).

La prueba de regresión de Egger rechazó el sesgo de publicación (IC:95% –3.48-3.13) (*p*=.91). Se encontraron señales de asimetría en el gráfico de embudo debido a los estudios con estimaciones de tamaño del efecto más grandes en el lado derecho de la figura (Emir, 2009; Nornoo et al., 2017; Tsai, 2014). Se efectuó un análisis de sensibilidad para evaluar las consecuencias de su presencia. La estimación combinada del tamaño del efecto respecto a los otros 27 estudios fue igualmente sólida y significativa (*r*=.26, IC:95% .18-.34).

4. Discusión

Se realizó una revisión sistemática de la literatura con metaanálisis con el objetivo de examinar la relación entre el pensamiento crítico y el rendimiento académico en estudiantes universitarios. El pensamiento crítico se correlacionó de manera positiva con el éxito académico en el grupo considerado. La magnitud de esta asociación fue débil. Esta relación se confirmó en el análisis de sensibilidad con un tamaño del efecto promedio invariado. A su vez, la correlación fue influenciada por moderadores como la especialidad de estudio, el instrumento empleado, la evaluación de habilidades *versus* disposiciones, así como la cantidad de dimensiones estimadas en cada estudio.

El pensamiento crítico se correlacionó positivamente con el éxito académico en magnitud débil (Cohen, 2013). El estudio reportó, comparativamente, un valor intermedio del tamaño del efecto (.26) en las correlaciones respecto

a metaanálisis previos. En este sentido, para Ross (2013), el tamaño del efecto promedio entre pensamiento crítico y el logro académico fue .31. Por el contrario, Fong y colaboradores (2017) reportaron un tamaño del efecto promedio de .24. Estas diferencias pudieran estar conectadas a las características sociodemográficas disímiles de sus muestras. Cabe señalar que el primero abarcó solamente internos de carreras en Ciencias de la Salud de varios países. Mientras que el segundo consideró estudiantes de *community colleges* estadounidenses. En el primer caso estos se encontraban finalizando su recorrido universitario, contrariamente el otro grupo se hallaba en los dos primeros años de la educación superior. Además, puede que el período de internado en salud suponga un incremento en las responsabilidades de no fácil compatibilidad con los estudios que pudiera afectar el PC (Anders et al., 2019; Stevenson y Clegg, 2013).

Por otra parte, entre los moderadores de esta relación destacó la especialidad de estudio. El tamaño del efecto promedio fue más del doble en los estudiantes que provenían de otras facultades respecto a los de Ciencias de la Salud. La evidencia de la influencia de la carrera de estudios sobre el éxito académico hasta el momento es contradictoria. Al respecto, la revisión con metaanálisis en internos de salud de Ross (2013) señaló coeficientes para estudiantes de enfermería de .33 y de medicina de .28. Sin embargo, la mitad de sus subgrupos de análisis estaban compuestos por un solo estudio lo que comporta alto riesgo de sesgo en la generalización de sus resultados. Otro metaanálisis (Fong et al., 2017) en población universitaria reportó valores para las carreras no vinculadas a las Ciencias de la Salud de .26, aunque esta diferencia no resultó significativa.

El instrumento empleado afectó la relación entre el PC y el éxito académico. El tamaño del efecto promedio reflejó ser más del doble en los universitarios evaluados con el WGCTA respecto al HSRT. El primer instrumento se desarrolló para valorar el pensamiento crítico en población general, abocándose a inferencias, reconocimiento de supuestos, razonamiento deductivo, interpretación y evaluación de argumentos (Anders et al., 2019). Mientras que el segundo evalúa universitarios de Ciencias de la Salud midiendo análisis, inferencias, evaluación, inducción y deducción (Comer et al., 2019; Partido y Soto, 2019). Las diferencias se relacionan a la especificidad de dominio y su transferibilidad en la medición del PC (Rear, 2019; Wang, 2017). En el sentido que para llevar a cabo su medición no se deberían utilizar conocimientos especializados de los sujetos evaluados como es en el caso del HSRT. Por el contrario, Rear (2019) propone que estos debieran ser genéricos. Es decir, si la aplicación se lleva a cabo dentro un contexto particular, no se justificaría del todo la evaluación de las habilidades generales del PC (Rear, 2019).

Por otra parte, el tamaño del efecto promedio fue mayor en la medición de las habilidades del PC comparado con las disposiciones en su relación con el éxito académico. Otro metaanálisis reportó diferencias similares (Ross et al., 2013). Una explicación de esta divergencia radicaría en que la relación entre disposiciones y logro académico se encontraría mediada por las habilidades del PC. Las disposiciones son motivaciones internas que permitirían responder habitualmente a una situación del entorno, aunque en modo potencialmente flexible (Taghinezhad et al., 2019). Estas son el presupuesto de las habilidades que derivan en comportamientos evidenciando niveles de PC variables entre sujetos (Taghinezhad et al., 2019).

Por último, la relación entre PC y éxito académico fue moderada por la cantidad de dimensiones estimadas en cada estudio. En aquellos que evaluaron una sola dimensión el tamaño del efecto promedio fue más del doble respecto a aquellos que evaluaron las dos. Este efecto moderador se debería a que 24 estudios indagaron en la medición de una dimensión del PC y solo seis consideraron ambas (Kidd y Latif, 2003; Partido y Soto, 2019; Miller, 2003; Whitney et al., 2016; Williams et al., 2003; Williams et al., 2006). En efecto, fueron 21 los que midieron habilidades, mientras que tres (AkbariLakeh et al., 2018; Emir, 2009; Ip et al., 2000) evaluaron solamente las disposiciones. Dados los pocos estudios que consideraron esta última se requiere más investigación en esta área.

Entre las limitaciones del metaanálisis, primeramente, se debe considerar la alta heterogeneidad entre los estudios. Aunque esta se espera en el diseño observacional (Deeks et al., 2019). Para explorar sus posibles causas se realizaron análisis de subgrupos. En segundo lugar, los hallazgos de la síntesis cuantitativa no deben interpretarse como apoyo de una relación causal. Es decir, no es posible inferir que en la población estudiada el pensamiento crítico condicione un mayor rendimiento debido a niveles elevados del primero.

En tercer lugar, como resultado de información incompleta hubo variables importantes que no se pudieron examinar. Entre estas figuran el tiempo dedicado al estudio, horas de trabajo o estrategias de aprendizaje. Por último, en la evaluación del PC tampoco fue posible deducir su temporalidad: su medición al comienzo o al final del recorrido académico pudiera actuar como variable de confusión.

Una fortaleza del estudio es que, en su conjunto, los trabajos meta analizados presentaron bajo riesgo de sesgo. No obstante, algunos mostraron alto riesgo en los dominios tasa de respuesta, tipo de muestreo y administración de las pruebas. Este elemento se debe tener en consideración a la hora de evaluar el efecto promedio encontrado.

A partir de lo señalado se mencionan algunas implicancias. Entre los criterios de admisión a las carreras universitarias las pruebas de PC podrían identificar candidatos con mayor probabilidad de tener éxito académico. Igualmente, ayudar a reconocer aquellos que podrían manifestar dificultades académicas. Es necesario tener en cuenta, que los resultados son referenciales y no necesariamente el único criterio predictor del éxito académico. Asimismo, se precisa reforzar el PC en los universitarios mediante talleres, programas y capacitaciones. Igualmente, sensibilizar sobre la importancia del PC en el rendimiento académico con la implementación de cursos específicos y contenidos transversales (Comer et al., 2019).

Futuras investigaciones pudieran considerar pruebas como la WGCTA, frente a otras con una menor magnitud en la evaluación del PC. Debido a que los instrumentos empleados son de opción múltiple podrían no captar la complejidad

de las dimensiones del PC (Anders et al. 2019; Rear, 2019). Esto sugeriría una operacionalización que considere todo el constructo teniendo en cuenta también factores psicosociales del contexto. Se recomienda examinar sistemáticamente la influencia moderadora de factores como el tiempo dedicado al estudio, las horas de trabajo o las principales estrategias de aprendizaje empleadas. Asimismo, el momento en que son medidos el PC y los indicadores de éxito académico.

Finalmente, en los estudiantes universitarios el PC se correlacionó positivamente con el éxito académico con una magnitud débil. Sin embargo, sus niveles se asociaron con menor fuerza en los estudiantes de Ciencias de la Salud en comparación con aquellos provenientes de otras facultades. Asimismo, la correlación con el éxito académico fue más baja en los estudios que emplearon el CCTDI respecto al WGCTA. Igualmente, para los estudios que midieron dos dimensiones del PC respecto a los que consideraron una de ellas. Por último, los resultados evidenciaron que la magnitud de los efectos fue mayor en los estudios que midieron las habilidades respecto a aquellos que consideraron las disposiciones. De acuerdo con esto, las habilidades jugarían un rol moderador en la relación entre las disposiciones al PC y el logro académico.

5. Referencias bibliográficas

Adamcik, B., Hurley, S., y Erramouspe, J. (1996). Assessment of pharmacy students' critical-thinking and problem-solving abilities. *American Journal of Pharmaceutical Education*, *60*(3), 256-264.

Afshar, H. S., y Movassagh, H. (2017). On the relationship among critical thinking, language learning strategy use and university achievement of Iranian English as a foreign language majors. *The Language Learning Journal*, *45*(3), 382-398. https://doi.org/10.1080/09571736.2014.914238

AkbariLakeh, M., Naderi, A., y Arbabisarjou, A. (2018). Critical thinking and emotional intelligence skills and relationship with students' academic achievement. *Prensa Med Argentina*, *104*(2), e1000280. https://doi.org/10.4172/0032-745X.1000280

Anders, P. L., Stellrecht, E. M., Davis, E. L., y McCall Jr, W. D. (2019). A systematic review of critical thinking instruments for use in dental education. *Journal of dental education*, *83*(4), 381-397. https://doi.org/10.21815/JDE.019.043

Baberi, F. K. S., y Bozorgi, Z. D. (2016). The Relationship between Goal Orientation Dimensions, Critical Thinking and Achievement Motivation with Students' Academic Performance. *International Journal of Humanities and Cultural Studies, 3*, 1768-1778.

Bayat, N. (2014). The Relationship between Prospective Teachers' Levels of Critical Thinking and Their Success in Academic Writing. *Egitim ve Bilim*, *39*(173), 155-169.

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge.

Comer, R. D., Schweiger, T. A., y Shelton, P. (2019). Impact of students' strengths, critical thinking skills and disposition on academic success in the first year of a pharmd program. *American Journal of Pharmaceutical Education*, *83*(1), 93-99. https://doi.org/10.5688/ajpe6499

Cottrell, S. (2017). *Critical thinking skills: Effective analysis, argument and reflection*. Macmillan International Higher Education.

D'Alessio, F. A., Avolio, B. E., y Charles, V. (2019). Studying the impact of critical thinking on the academic performance of executive MBA students. *Thinking Skills and Creativity*, *31*, 275-283. https://doi.org/10.1016/j.tsc.2019.02.002

Deeks, J. J., Higgins, J. P., Altman, D. G., y Cochrane Statistical Methods Group. (2019). Analysing data and undertaking meta□analyses. En J. Higgins et al., (Coord.), *Cochrane handbook for systematic reviews of interventions* (pp. 241-284). Hoboken (NJ): John Wiley & Sons. https://doi.org/10.1002/9781119536604.ch10

Denial, A., y Pitcher, M. (2007). Optometric GPA, NBEO and Clinical Performance Compared To Critical Thinking Skill And Disposition. *Optometric Education*, *32*(3), 79-84.

El Hassan, K., y Madhum, G. (2007). Validating the Watson Glaser critical thinking appraisal. *Higher Education*, *54*(3), 361-383. https://doi.org/10.1007/s10734-006-9002-z

Emir, S. (2009). Education faculty students' critical thinking disposition according to achedemic achievement. *Procedia-Social and Behavioral Sciences*, *1*(1), 2466-2469. https://doi.org/10.1016/j.sbspro.2009.01.433

Facione, P. A., Crossetti, M. D. G. O., y Riegel, F. (2017). Holistic Critical Thinking in the Nursing Diagnostic Process. *Rev Gaúcha Enferm*, *38*(3), e75576. https://doi.org/10.1590/1983-1447.2017.03.75576

Fong, C. J., Kim, Y., Davis, C. W., Hoang, T., y Kim, Y. W. (2017). A meta-analysis on critical thinking and community college student achievement. *Thinking Skills and Creativity*, *26*, 71-83. https://doi.org/10.1016/j.tsc.2017.06.002

Franco, A., y Almeida, L. D. S. (2015). Real-World Outcomes y Pensamiento Crítico: Análisis Diferencial por Área y Género. *Paidéia*, *25*(61), 173-181. https://doi.org/10.1590/1982-43272561201505

Gadzella, B. M., Balogu, M., y Stephens, R. (2002). Prediction of GPA with educational psychology grades and critical thinking scores. *Education*, *122*(3), 618-623.

Ghanizadeh, A. (2017). The interplay between reflective thinking, critical thinking, self-monitoring, and academic achievement in higher education. *Higher Education*, *74*(1), 101-114. https://doi.org/10.1007/s10734-016-0031-y

Huhn, K., y Parrott, S. J. (2017). Exploration of relationships among the health sciences reasoning test, the national physical therapy licensing examination, and cognitive admission variables. *Journal of Physical Therapy Education*, *31*(1), 7-13.

Ip, W. Y., Lee, D. T., Lee, I. F., Chau, J. P., Wootton, Y. S., y Chang, A. M. (2000). Disposition towards critical thinking: a study of Chinese undergraduate nursing students. *Journal of Advanced Nursing*, *32*(1), 84-90. https://doi.org/10.1046/j.1365-2648.2000.01417.x

Jenkins, E. K. (1998). The significant role of critical thinking in predicting auditing students' performance. *Journal of Education for Business*, *73*(5), 274-279. https://doi.org/10.1080/08832329809601644

Kennison, M. M. (2006). The evaluation of students' reflective writing for evidence of critical thinking. *Nursing Education Perspectives*, *27*(5), 269-273.

Kidd, R. S., y Latif, D. A. (2003). Traditional and novel predictors of classroom and clerkship success of pharmacy students. *American Journal of Pharmaceutical Education*, *67*(4), 860-865.
Kuhn, D. (2019). Critical thinking as discourse. *Human Development*, *62*(3), 146-164. https://doi.org/10.1159/000500171
Manalo, E., Kusumi, T., Koyasu, M., Michita, Y., y Tanaka, Y. (2013). To what extent do culture-related factors influence university students' critical thinking use? *Thinking Skills and Creativity*, *10*, 121-132. https://doi.org/10.1016/j.tsc.2013.08.003
Miller, D. A., Sadler, J. Z., y Mohl, P. C. (1993). Critical thinking in preclinical course examinations. *Academic Medicine, 68*(3), 303-305. https://doi.org/10.1097/00001888-199304000-00020
Miller, D. R. (2003). Longitudinal assessment of critical thinking in pharmacy students. *American Journal of Pharmaceutical Education*, *67*(4), 890-897.
Miller, M. A. (1992). Outcomes evaluation: Measuring critical thinking. *Journal of Advanced Nursing*, *17*(12), 1401-1407. https://doi.org/10.1111/j.1365-2648.1992.tb02810.x
Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., y Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic reviews*, *4*(1), e7647. https://doi.org/10.1136/bmj.g7647
Nornoo, A. O., Jackson, J., y Axtell, S. (2017). Investigating the correlation between pharmacy student performance on the health science reasoning test and a critical thinking assignment. *American journal of pharmaceutical education*, *81*(2), 1-6. https://doi.org/ 10.5688/ajpe81224
Pardamean, B. (2012). Measuring change in critical thinking skills of dental students educated in a PBL curriculum. *Journal of Dental Education*, *76*(4), 443-453. https://doi.org/10.1002/j.0022-0337.2012.76.4.tb05276.x
Partido, B. B., y Soto, I. (2019). Association between critical thinking and academic performance in dental hygiene students. *Journal of Dental Education*, *83*(1), 32-38. https://doi.org/10.21815/JDE.019.004
Poce, A., Amenduni, F., De Medio, C., y Re, M. R. (2019). Road to Critical Thinking automatic assessment: a pilot study. *Form@re-Open Journal per la formazione in rete*, *19*(3), 60-72. https://doi.org/10.13128/form-7705
Rear, D. (2019). One size fits all? The limitations of standardised assessment in critical thinking. *Assessment & Evaluation in Higher Education*, *44*(5), 664-675. https://doi.org/10.1080/02602938.2018.1526255
Rebele, J. E., y Pierre, E. K. S. (2019). A commentary on learning objectives for accounting education programs: The importance of soft skills and technical knowledge. *Journal of Accounting Education*, *48*, 71-79. https://doi.org/10.1016/j.jaccedu.2019.07.002
Ross, D., Loeffler, K., Schipper, S., Vandermeer, B., y Allan, G. M. (2013). Do scores on three commonly used measures of critical thinking correlate with academic success of health professions trainees? A systematic review and meta-analysis. *Academic Medicine*, *88*(5), 724-734. https://doi.org/10.1097/ACM.0b013e31828b0823
Schwanz, K., y McIlreavy, M. (2015). Academic performance of introductory psychology students: the importance of critical thinking. *Res. Psychol. Behav. Sci*, *3*(2), 25-31. https://doi.org/10.12691/rpbs-3-2-2
Scott, J. N., y Markert, R. J. (1994). Relationship between critical thinking skills and success in preclinical courses. *Academic medicine, 69*(11), 920-924. https://doi.org/10.1097/00001888-199411000-00015
Shirazi, F., y Heidari, S. (2019). The relationship between critical thinking skills and learning styles and academic achievement of nursing students. *The Journal of Nursing Research*, *27*(4), e38. https://doi.org/10.1097/jnr.0000000000000307
Stevenson, J., y Clegg, S. (2013). 'My past is a double edge sword': temporality and reflexivity in mature learners. *Studies in Continuing Education*, *35*(1), 17-29. https://doi.org/10.1080/0158037X.2012.684794
Stroup, D. F., Berlin, J. A., Morton, S. C., Olkin, I., Williamson, G. D., y Rennie, D. (2009). Meta-analysis of observational studies in epidemiology: a proposal for reporting. *Jama*, *283*(15), 2008-2012. https://doi.org/10.1001/jama.283.15.2008
Taghinezhad, A., Riasati, M. J., y Behjat, F. (2019). The effect of teaching critical thinking strategies on students' academic writing, critical thinking ability, and critical thinking dispositions. *International Journal of Foreign Language Teaching and Research*, *7*(28), 37-55.
Tiessen, J. B. (1987). Critical thinking and selected correlates among baccalaureate nursing students. *Journal of Professional Nursing*, *3*(2), 118-123.
Tsai, T. H. (2014). Validating use of a critical thinking test for the dental admission test. *Journal of Dental Education*, *78*(4), 552-557. https://doi.org/10.1002/j.0022-0337.2014.78.4.tb05705.x
Turner, H., y Williams, R. L. (2007). Vocabulary development and performance on multiple-choice exams in large entry-level courses. *Journal of College Reading and Learning*, *37*(2), 64-81. https://doi.org/10.1080/10790195.2007.10850198
Vendrely, A. M. (2007). An investigation of the relationships among academic performance, clinical performance, critical thinking, and success on the physical therapy licensure examination. *Journal of allied health*, *36*(2), 108-123.
Wang, S. (2017). An exploration into research on critical thinking and its cultivation: An overview. *Theory and Practice in Language Studies*, *7*(12), 1266-1280. https://doi.org/10.17507/tpls.0712.14
Wells, G. A., Shea, B., O'connell, D., Peterson, J., Welch, V., Losos, M., y Tugwell, P. (20 de julio de 2020). *The Newcastle-Ottawa Scale (NOS) for assessing the quality if nonrandomized studies in meta-analyses*. The Ottawa Hospital. Our Research. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp
Whitney, E. M., Aleksejuniene, J., y Walton, J. N. (2016). Critical thinking disposition and skills in dental students: development and relationship to academic outcomes. *Journal of Dental Education*, *80*(8), 948-958. https://doi.org/10.1002/j.0022-0337.2016.80.8.tb06175.x
Williams, K. B., Glasnapp, D. R., Tilliss, T. S., Osborn, J., Wilkins, K., Mitchell, S. Kershbaum, W. y Schmidt, C. (2003). Predictive validity of critical thinking skills for initial clinical dental hygiene performance. *Journal of dental education*, *67*(11), 1180-1192. https://doi.org/10.1002/j.0022-0337.2003.67.11.tb03709.x
Williams, K. B., Schmidt, C., Tilliss, T. S., Wilkins, K., y Glasnapp, D. R. (2006). Predictive validity of critical thinking skills and disposition for the national board dental hygiene examination: a preliminary investigation. *Journal of Dental Education*, *70*(5), 536-544. https://doi.org/10.1002/j.0022-0337.2006.70.5.tb04109.x

Anexo 1 – Versión de la Escala Newcastle-Ottawa

Criterios de evaluación de la calidad metodológica de las fuentes primarias. Los posibles niveles de sesgo son alto riesgo de sesgo, bajo riesgo de sesgo y poco claro. Además, a cada elemento se le asignó 1 por bajo riesgo de sesgo y 0 por alto riesgo de sesgo o poco claro. La puntuación de calidad total para cada estudio puede variar de 0 a 5, valores más altos indican un menor riesgo de sesgo. Por último, se consideró que los estudios tenían bajo riesgo de sesgo (≥3 puntos) o alto riesgo de sesgo (<3 puntos).

Dominios	Nivel de sesgo	Indicadores de evaluación	Puntuación
Tasa de respuesta	Alto riesgo	Tasa de respuesta inferior al 60% o no hubo descripción de la tasa de respuesta	0
	Bajo riesgo	Tasa de respuesta superior al 60%	1
	Poco claro	Información insuficiente	0
Tipo de muestreo	Alto riesgo	Métodos de muestreo por conveniencia	0
	Bajo riesgo	Muestreo de toda la cohorte o una muestra aleatoria de la cohorte	1
	Poco claro	Información insuficiente	0
Modo de administración de las pruebas	Alto riesgo	Si las pruebas se enviaron por correo electrónico a los participantes o si los resultados de la prueba se recuperaron retrospectivamente	0
	Bajo riesgo	Cuando la administración de la prueba probablemente se observó, como en sesiones presenciales	1
	Poco claro	Información insuficiente	0
Uso de instrumentos validados	Alto riesgo	Si los instrumentos empleados no reportan validez	0
	Bajo riesgo	Si los instrumentos utilizados reportan validez	1
	Poco claro	Información insuficiente	0
Análisis estadístico	Alto riesgo	No se informaron estadísticas descriptivas, estaban incompletas o no incluían medidas adecuadas de pensamiento crítico	0
	Bajo riesgo	Estadísticas descriptivas informadas para describir la población (por ejemplo, edad, sexo) con medidas adecuadas de pensamiento crítico (por ejemplo, desviación estándar, error estándar, rango, porcentaje	1
	Poco claro	Información insuficiente	0