

Original

Efecto del aprendizaje basado en juegos en las habilidades de pensamiento de orden superior en matemáticas de los estudiantes: un metanálisis



Bambang Sri Anggoro^{a,*}, Andi Harpeni Dewantara^b, Suherman Suherman^{a,c}, Rosida Rakhmawati Muhammad^a, y Sari Saraswati^d

^a Departamento de Educación Matemática, Universitas Islam Negeri Raden Intan Lampung, Bandar Lampung, Indonesia

^b Investigación y Evaluación Educativa, Universitas Negeri Yogyakarta, Yogyakarta, Indonesia

^c Escuela Doctoral de Educación, Universidad de Szeged, Szeged, Hungría

^d Departamento de Educación Matemática, Universitas Hasyim Asy'ari, Jombang, Indonesia

INFORMACIÓN DEL ARTÍCULO

Historia del artículo:

Recibido el 17 de mayo de 2024

Aceptado el 10 de octubre de 2024

On-line el 22 de noviembre de 2024

Palabras clave:

Aprendizaje basado en juegos
Habilidades de pensamiento de orden superior en matemáticas
Metanálisis
Habilidades de pensamiento

R E S U M E N

Las tendencias en la investigación educativa demuestran un creciente interés en comprender el impacto de los juegos en el logro matemático. Sin embargo, solo un número limitado de estudios ha sintetizado investigaciones cuantitativas anteriores para investigar cómo el aprendizaje basado en juegos (ABJ) influye en la mejora de las habilidades de pensamiento de orden superior (HPOS) en matemáticas de los estudiantes. Para abordar esta brecha, este estudio de metanálisis tiene como objetivo investigar la efectividad del ABJ en las HPOS matemáticas de los estudiantes. El estudio ha analizado 40 tamaños del efecto de 13 estudios empíricos publicados entre 2010 y 2024, recopilados de bases de datos electrónicas como Scopus, Eric y Pro-Quest. Se ha empleado el software R para analizar los tamaños del efecto, detectar sesgos de publicación y realizar análisis de subgrupos. Utilizando el algoritmo k-means, se han excluido dos estudios identificados como atípicos. Después de eliminar estos atípicos, los hallazgos han revelado un efecto positivo del ABJ en las HPOS matemáticas de los estudiantes, con un efecto general que cae dentro de la categoría pequeña ($g = .134, p < .001$). Además, las variables moderadoras, incluyendo el nivel educativo, las habilidades de pensamiento medidas, el continente, el tratamiento de control y la duración de la intervención, han influido significativamente en la mejora de las HPOS utilizando ABJ, mientras que el factor tamaño de muestra no ha mostrado un impacto significativo. Todos los hallazgos, limitaciones e implicaciones se discuten en este artículo.

© 2024 Universidad de País Vasco. Publicado por Elsevier España, S.L.U. Se reservan todos los derechos, incluidos los de minería de texto y datos, entrenamiento de IA y tecnologías similares.

Effect of game-based learning on students' mathematics high order thinking skills: A meta-analysis

A B S T R A C T

Educational research trends demonstrate a growing interest in understanding the impact of games on mathematics achievement. However, only a limited number of studies have synthesized previous quantitative studies to investigate how game-based learning (GBL) influences the enhancement of students' high-order thinking skills (HOTS) in mathematics. To address this gap, this meta-analysis study aims to investigate the effectiveness of GBL on students' mathematics HOTS. The study analyzed 40 effect sizes from 13 empirical studies published from 2010 to 2024, gathered from electronic databases such as Scopus, Eric, and Pro-Quest. R software was employed for analyzing effect sizes, detecting publication bias, and conducting subgroup analyses. Using the k-means algorithm, two studies identified as outliers were excluded. After removing these outliers, the findings revealed a positive effect of GBL on students' mathematics HOTS, with an overall effect falling within the small category ($g = .134, p < .001$). Additionally,

Keywords:
Game-Based Learning
Mathematics Higher Order Thinking Skills
Meta-Analysis
Thinking Skills

Véase contenido relacionado en DOI: <https://doi.org/10.1016/j.psicoe.2024.500158>

* Autor para correspondencia.

Correo electrónico: bambangsrianggoro@radenintan.ac.id (B.S. Anggoro).

moderator variables, including educational level, measured thinking skills, continent, control treatment, and intervention duration, significantly influenced the improvement of HOTS using GBL, while the sample size factor showed no significant impact. All findings, limitations, and implications are discussed in this article.

© 2024 Universidad de País Vasco. Published by Elsevier España, S.L.U. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

Introducción

Los conceptos matemáticos son aplicables en un amplio espectro de escenarios cotidianos, que van desde tareas sencillas como determinar descuentos, ganancias y pérdidas (Suherman et al., 2020; Wibawa et al., 2022), hasta aplicaciones más complejas en diversos campos como la ingeniería, la economía, la farmacéutica y la contabilidad (López-Díaz y Peña, 2021; Mumcu, 2018). Además, las matemáticas se han considerado una materia fundamental y la base de la ciencia y la tecnología (Yeh et al., 2019). En relación con esos roles esenciales, el enfoque instruccional en las matemáticas del aula no debe enfatizar únicamente la comprensión procedural o el conocimiento teórico, sino que también debe tener como objetivo desarrollar habilidades de pensamiento de orden superior (HPOS) (Lovianova et al., 2022), particularmente el pensamiento analítico, entre los aprendices de matemáticas. Esta competencia no solo es ventajosa académicamente, sino que también tiene una relevancia sustancial en la capacidad para abordar y resolver desafíos del mundo real de manera efectiva.

Desarrollar HPOS, como la resolución de problemas, el razonamiento, el pensamiento crítico y el pensamiento computacional, es crucial para los estudiantes. Sin embargo, los educadores de matemáticas a menudo enfrentan desafíos significativos en este empeño (Ganapathy et al., 2017). Una dificultad principal que enfrentan los docentes es la búsqueda de métodos atractivos y efectivos para estimular el pensamiento de orden superior de los estudiantes. Estos desafíos surgen de las características abstractas, lógicas y sistemáticas de las matemáticas, junto con la complejidad de sus símbolos y fórmulas (Acharya, 2017). Los desafíos para mejorar a los estudiantes son multifacéticos, abarcando diversas necesidades de los estudiantes que surgen de sus variadas habilidades iniciales (Smith et al., 2022), creencias matemáticas y limitaciones prácticas asociadas con la implementación de métodos de enseñanza específicos (Yerizon et al., 2022).

Una estrategia potencial para abordar estos desafíos es integrar juegos en el proceso de aprendizaje. El aprendizaje basado en juegos (ABJ) ha suscitado un interés significativo tanto en investigadores como en practicantes (Qian y Clark, 2016). La integración de juegos en el aprendizaje es intrigante de explorar, ya que la mayoría de los estudiantes disfrutan naturalmente de jugar (White y McCoy, 2019). El ABJ implica incorporar juegos como herramientas educativas, sumergiendo a los estudiantes en un entorno de aprendizaje dinámico e interactivo. El ABJ ha surgido como un enfoque de aprendizaje innovador que puede mejorar la motivación, el compromiso emocional y la diversión de los estudiantes (Hartt et al., 2020). El ABJ también se presenta como un enfoque innovador con el potencial de ayudar a los educadores a elevar las HPOS de los estudiantes.

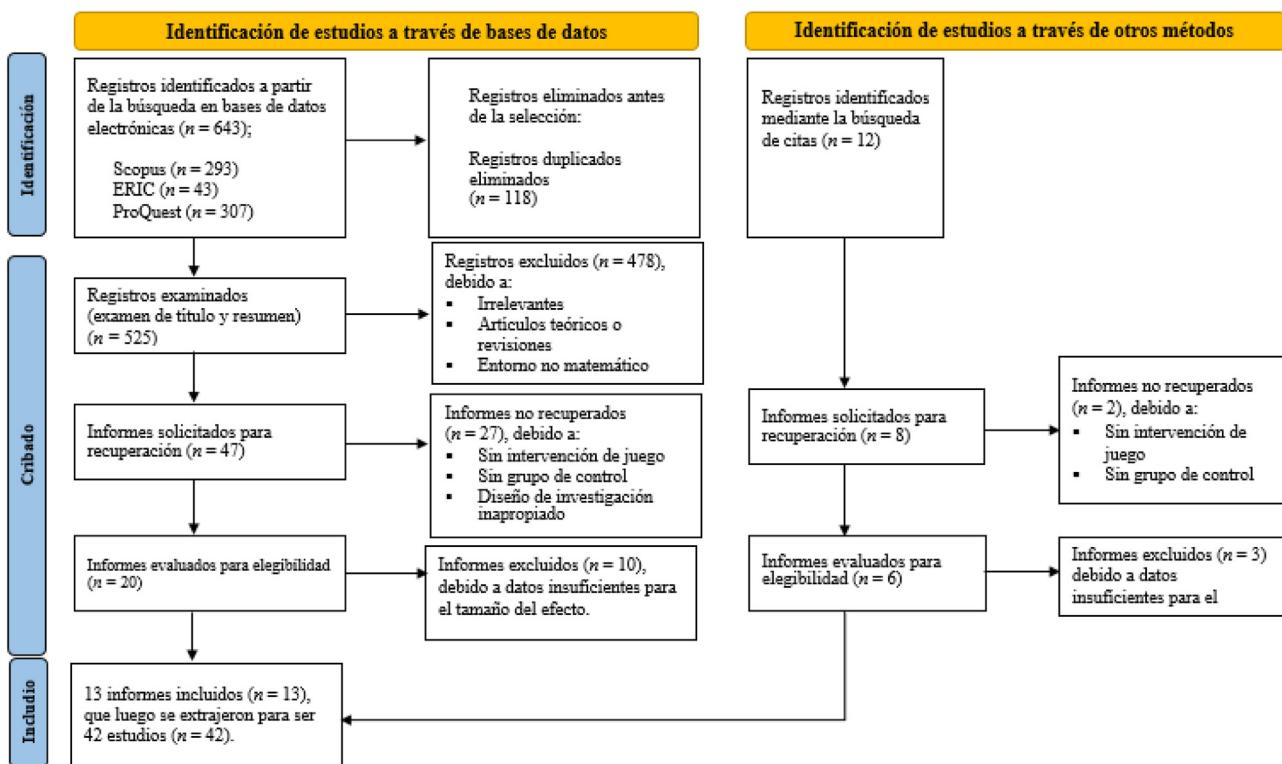
Varios estudios anteriores han proporcionado valiosas perspectivas sobre los impactos del ABJ, destacando su mejora significativa en el compromiso educativo de los estudiantes a través de diversos niveles educativos (Jabbar y Felicia, 2015; Shu y Liu, 2019). El ABJ tiene el potencial de hacer que el aprendizaje sea agradable e interactivo, fomentando el interés y el compromiso de los estudiantes, incluida la implicación cognitiva (Deng et al., 2020). En el ABJ, los estudiantes no solo son desafiados a avanzar a niveles más altos, sino que también obtienen satisfacción y alegría al completar con éxito tareas o misiones del juego (Freitas et al., 2017). En un

contexto más específico, integrar juegos en el proceso de aprendizaje tiene el potencial de mejorar habilidades de pensamiento cognitivo avanzado, como el razonamiento, el pensamiento crítico y la resolución de problemas (Angelelli et al., 2023; Bourke, 2019). Varios estudios anteriores respaldan la idea de que el ABJ tiene el potencial de mejorar las habilidades de pensamiento matemático. Asri y Jamaludin (2022) han destacado la efectividad de los juegos de Scratch en el desarrollo de habilidades de pensamiento. Varios estudios han demostrado la mejora de las habilidades de resolución de problemas a través del ABJ utilizando juegos de mente (Demirel y Yilmaz, 2019), y un juego epistémico con temática de arquitectura (Ke, 2019). Indriani et al. (2019) han encontrado que el aprendizaje basado en problemas asistido por un juego de monopolio ha mejorado la capacidad de pensamiento crítico, mientras que Izzati et al. (2022) han mostrado un mejoramiento del pensamiento divergente con juegos de serpientes y escaleras.

A pesar de que varios estudios han explorado el uso del ABJ en la educación matemática, los hallazgos siguen siendo diversos y, a veces, contradictorios. Algunos estudios abogan por el impacto positivo significativo del ABJ en las HPOS de los estudiantes (Al-Absi, 2017; Izzati et al., 2022; Lu et al., 2023; Ma et al., 2023; Taja-on, 2019), mientras que otros sugieren que no hay diferencias sustanciales en comparación con los métodos de enseñanza tradicionales (Emihovich et al., 2022; Lee et al., 2014). Esta incongruencia plantea una pregunta fundamental: ¿cómo podemos establecer conclusiones consistentes sobre los efectos del ABJ en las HPOS matemáticas de los estudiantes? Para abordar esta pregunta de manera integral, es necesario realizar un metanálisis que sintetice de manera sistemática los resultados de múltiples estudios.

Varios estudios de metanálisis anteriores han examinado el ABJ, destacando su promesa como un enfoque pedagógico que mejora efectivamente el desarrollo cognitivo de los estudiantes (Turgut y Temur, 2017; Wang et al., 2022) y los resultados de aprendizaje (Bai et al., 2020; Barz et al., 2024). Algunos metanálisis se han centrado específicamente en las HPOS de los estudiantes, como los efectos del ABJ en el pensamiento computacional (Lu et al., 2023; Ma et al., 2023) y el pensamiento crítico (Mao et al., 2022). Sin embargo, estos estudios a menudo abordan habilidades de pensamiento generales en lugar de centrarse específicamente en las matemáticas. Por otro lado, hay metanálisis que examinan los efectos del ABJ en un contexto matemático, pero su enfoque está en el logro matemático general en lugar de las HPOS. Por ejemplo, se ha informado que el ABJ es un método educativo efectivo para mejorar el rendimiento matemático de los estudiantes en niveles educativos específicos, como K-12 (Byun y Joung, 2018) y niveles de grado Pre K-12 (Tokac et al., 2019). Sin embargo, la investigación que examina específicamente el impacto del ABJ en las HPOS matemáticas de los estudiantes a través de varios niveles educativos sigue siendo limitada.

Para llenar esta brecha, este metanálisis actual tiene como objetivo evaluar la efectividad del ABJ, en comparación con métodos tradicionales no-ABJ, en las HPOS matemáticas de los estudiantes. Al integrar estudios con diversos diseños de investigación y resultados, el estudio tiene como objetivo proporcionar una comprensión matizada de la efectividad del ABJ en la mejora de las HPOS matemáticas de los estudiantes, como la resolución de problemas, el pensamiento crítico, el razonamiento matemático, el pensamiento

**Identificación de estudios a través de otros métodos****Figura 1.** Diagrama de flujo PRISMA.

computacional y otras habilidades de pensamiento cognitivo relevantes (Suherman y Vidákovich, 2022). A través de este examen sistemático, el estudio busca contribuir con valiosas perspectivas a educadores e investigadores, arrojando luz sobre el potencial del ABJ para elevar las HPOS matemáticas de los estudiantes.

Método

Este estudio emplea un enfoque de metanálisis, un método estadístico para sintetizar hallazgos de numerosos estudios cuantitativos dentro de un dominio de investigación particular (Pigott y Polanin, 2020). En particular, este estudio ha aplicado un metanálisis de contraste de grupos, que implica la medición de una o más variables en dos o más grupos de encuestados, seguido de comparaciones subsecuentes (Borenstein et al., 2009).

Marco teórico

Para realizar un metanálisis integral sobre el efecto del ABJ en las HPOS matemáticas de los estudiantes, este estudio ha empleado el nuevo método de Elementos de Informe Preferidos para Revisiones Sistemáticas y Meta-Análisis (PRISMA) 2020 (Page et al., 2021). Las fuentes de datos han consistido en estudios (artículos de revistas revisados por pares y artículos de actas). Se han consultado bases de datos electrónicas, incluyendo Scopus, Eric y Pro-Quest, en busca de estudios relevantes publicados desde enero de 2010 hasta enero de 2024. Este período ha visto avances significativos en la tecnología educativa y el diseño de juegos (Qian y Clark, 2016), un aumento en la adopción del aprendizaje basado en juegos en entornos educativos (Plass et al., 2015) y un cambio hacia la enfatización de las habilidades de pensamiento de orden superior en los planes de estudio de matemáticas (Scherer y Beckmann, 2014). Además, las metodologías de investigación se han vuelto más sofisticadas (Merchant et al., 2014), y las políticas educativas han enfatizado cada vez más la integración de la tecnología en

la educación matemática (Hsu et al., 2020). Este marco temporal proporciona, por tanto, un rico cuerpo de investigación reciente que refleja los contextos y prácticas educativas actuales, asegurando que los hallazgos del metanálisis sean aplicables a la instrucción matemática contemporánea y a los desarrollos futuros en el aprendizaje basado en juegos.

Los términos de búsqueda han incluido variaciones de «aprendizaje basado en juegos», «educación matemática» y «HPOS», en las que las palabras clave relevantes se han combinado con los términos «Y» y «O» como operadores booleanos (Wang et al., 2022). El primer conjunto de palabras clave ha abarcado términos como «aprendizaje basado en juegos», «ABJ», «juego educativo», «aprendizaje gamificado», «gamificación», «juego de computadora» y «juego digital». El segundo conjunto ha incluido palabras clave como «matemáticas», «aprendizaje matemático», «educación matemática». El tercer conjunto ha cubierto términos como «habilidades de pensamiento de orden superior», «HPOS», «habilidades de pensamiento», «pensamiento cognitivo», «resolución de problemas», «pensamiento crítico», «pensamiento creativo», «pensamiento analítico», «pensamiento computacional», «razonamiento».

De la búsqueda de literatura, se han identificado un total de 293 registros en Scopus, 43 en ERIC y 307 en ProQuest (Figura 1). Además, se ha llevado a cabo una búsqueda manual de literatura, que ha abarcado el examen de bibliografías de literatura relevante para el tema del metanálisis. Posteriormente, la literatura adquirida ha sido sometida a un proceso de selección basado en los criterios de inclusión establecidos por el investigador.

Selección de literatura y criterios de inclusión

Se han establecido ocho criterios de inclusión como pautas para la selección de literatura. La literatura que no cumpla con todos los siguientes criterios de inclusión será excluida. Los ocho criterios de inclusión empleados en este estudio son: (1) La literatura debe

haber sido publicada entre 2010 y 2024; (2) La literatura debe estar publicada en revistas revisadas por pares o actas de conferencias, excluyendo trabajos teóricos, informes de estudio, reseñas de libros y capítulos de libros; (3) Los estudios deben haberse realizado en entornos educativos de todos los niveles, desde la educación primaria hasta la educación superior; (4) Los estudios deben tener contenido de texto completo en inglés. Se han excluido los estudios no publicados en inglés; (5) La investigación debe tener un enfoque específico en la implementación de intervenciones de ABJ en aulas de matemáticas; (6) Los estudios deben reportar resultados relacionados con las HPOS, abarcando la resolución de problemas, el pensamiento crítico, el pensamiento computacional, el pensamiento creativo, el razonamiento matemático y otras habilidades cognitivas relevantes; (7) Los estudios deben investigar empíricamente el impacto del ABJ en las HPOS en el aprendizaje de matemáticas (se han excluido los estudios que investigan los efectos del ABJ en las HPOS de los estudiantes en materias distintas de las matemáticas); (8) Los estudios deben haber incluido al menos un grupo utilizando ABJ y un grupo de control con un entorno tradicional o una intervención no relacionada con juegos (que no involucrara ABJ). Se han excluido los estudios sin grupo de control; y (9) se ha requerido que los estudios proporcionen datos suficientes para calcular tamaños de efecto. Esto incluye la presentación de datos cuantitativos como medias, desviaciones estándar y tamaños de muestra para ambos grupos, experimental y de control, o la presentación de t-tests, F-tests o valores p exactos junto con los tamaños de muestra para ambos grupos. Se han excluido los estudios con datos insuficientes para calcular tamaños de efecto.

Extracción de datos, esquema de codificación y variables moderadoras

El proceso de extracción de datos y codificación en este metanálisis se adhiere a un enfoque meticuloso y sistemático, que abarca varios pasos desde la búsqueda de literatura hasta el análisis final de datos. La fase inicial ha involucrado una búsqueda de literatura utilizando palabras clave específicas para identificar estudios relevantes relacionados con el impacto del ABJ en la mejora de las HPOS matemáticas de los estudiantes. Posteriormente, se ha realizado la selección de literatura basada en criterios de inclusión predefinidos.

Una vez que la literatura ha sido recopilada con éxito, se ha iniciado un cuidadoso proceso de extracción de datos a partir de los estudios seleccionados. Este proceso ha involucrado la captura de información sobre el diseño del estudio, el tamaño de la muestra, los resultados reportados y otras variables esenciales. El metanálisis ha analizado trece estudios empíricos, incluyendo artículos de revistas y actas de conferencias, generando 42 tamaños de efecto para un análisis posterior. Un resumen del proceso de búsqueda de literatura se muestra en el nuevo diagrama de flujo PRISMA ([Figura 1](#)).

Los trece trabajos seleccionados, que cumplían con los criterios de inclusión, han pasado por un detallado proceso de extracción de datos, resultando en la derivación de 42 tamaños de efecto. A continuación, se ha llevado a cabo un proceso de codificación para identificar las variables asociadas a los tamaños de efecto, incluyendo la media, la desviación estándar (DE), el tamaño de la muestra (n) y los tamaños de efecto relacionados con las puntuaciones del posttest para ambos grupos, experimental y de control. Además, se han identificado las variables moderadoras presentes en cada estudio durante el proceso de codificación.

El libro de códigos (esquema de codificación) sirve como una guía integral para codificar cada estudio, conteniendo información esencial como: (1) Detalles de la literatura, incluyendo el apellido del autor o autores, año de publicación, título de los estudios y tipo de publicación (artículos revisados por pares o actas); (2) Tamaño de la muestra: pequeño ($n < 30$) (S) y grande ($n > 30$) (L).

Este tamaño de muestra se refiere a la muestra de la clase de experimento en los estudios analizados; (3) Habilidades de pensamiento cognitivo medidas: resolución de problemas (RP), pensamiento computacional (PC), pensamiento crítico (PCr), razonamiento (Rz) y otros (O); (4) Nivel educativo: educación primaria (EP), educación secundaria (ES) y educación universitaria (EU). Estas categorías de nivel educativo utilizadas en este metanálisis se han determinado en función de las clasificaciones proporcionadas en los artículos originales, en lugar de rangos de edad específicos. Sin embargo, los rangos de edad para cada estudio analizado se presentan en la [Tabla 1 \(Apéndice\)](#); (5) Tratamiento de control: tradicional (Tr) y multimedia (MM); (6) Continente: Asia (A), Europa (E), América del Norte (AN) y Otros (O); (7) Duración de la intervención: < 10 horas (ID 1), 10 – 20 horas (ID 2), > 20 horas (ID 3) y no especificado (Ne); e (8) Información estadística para calcular tamaños de efecto, incluyendo el total de la muestra del grupo experimental (NE), puntuación media del grupo experimental (ME) y desviación estándar del grupo experimental (DEE), total de la muestra del grupo de control (NC), puntuación media del grupo de control (MC) y desviación estándar del grupo de control (DEC).

Además, dos evaluadores han participado en la evaluación de la codificación de la literatura realizada por los investigadores en este metanálisis. Para garantizar la fiabilidad, se ha medido la consistencia entre evaluadores utilizando las estadísticas kappa ([Ho et al., 2019](#)).

Cálculo del tamaño del efecto y modelos estadísticos

Esta investigación ha comparado el logro cognitivo en matemáticas entre los grupos de control y experimental, y luego se ha calculado el tamaño del efecto utilizando la diferencia de medias estandarizada (DME) propuesta por [Hedges \(1981\)](#). Las estadísticas descriptivas de la mayoría de los estudios, incluyendo la media (\bar{X}_1 , \bar{X}_2), la desviación estándar (s_1 , s_2) y el tamaño de la muestra (n_1 , n_2) para ambos grupos, nos permiten calcular los valores de DME utilizando la siguiente fórmula ([Lipsey y Wilson, 2001](#)):

$$d = \frac{\bar{X}_2 - \bar{X}_1}{S_{pooled}}$$

Donde:

$$S_{pooled} = \frac{\sqrt{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}}{n_1 - n_2 - 2}$$

[Hedges \(1981\)](#) ha encontrado que el valor d sobreestima ligeramente el parámetro poblacional absoluto, lo que ha llevado a una modificación propuesta conocida como g de Hedges.

[Hedges \(1981\)](#) ha encontrado que el valor d sobreestima ligeramente el parámetro poblacional absoluto, lo que ha llevado a una modificación propuesta conocida como g de Hedges. Para convertir de d a g, se necesita un factor de corrección llamado J, calculado con la siguiente fórmula:

$$j = 1 - \frac{3}{4df - 1}$$

Donde df son los grados de libertad calculados como: $df = n_1 + N_2 - 2$

con el tamaño del efecto corregido, g de Hedges se ha calculado como:

$$g = j$$

Después de calcular el tamaño del efecto para cada estudio, se ha calculado el efecto agregado (efecto medio) utilizando un modelo de efectos aleatorios (EA). La selección de un modelo de efectos aleatorios sobre un modelo fijo se ha basado en la suposición de

los investigadores de que los tamaños del efecto en varios estudios podrían derivarse potencialmente de poblaciones distintas, y estas diferentes poblaciones tienen sus propias distribuciones de muestreo (Borenstein et al., 2009). Esto se refuerza por la opinión de Cooper (2016) de que la diversidad en los entornos experimentales en cada estudio (por ejemplo, género, país, nivel educativo) sería más adecuada para el análisis utilizando un modelo de efectos aleatorios.

En nuestro análisis, empleamos un marco frecuentista, que se basa en métodos estadísticos clásicos para estimar parámetros e inferir características de la población basadas en datos observados. Este marco se utilizó para evaluar el tamaño del efecto agrupado y la influencia de las variables moderadoras. Además, utilizamos ponderación de varianza inversa para la integración meta-analítica, como la forma más común de calcular un tamaño de efecto agrupado (Harrer et al., 2021). Este método otorga más peso a los estudios con estimaciones de tamaño de efecto más precisas, que son típicamente aquellos con tamaños de muestra más grandes o varianzas más pequeñas. El tamaño del efecto general se ha calculado como un promedio ponderado de los tamaños de efecto de estudios individuales, donde los pesos corresponden a la precisión de la estimación del tamaño de efecto de cada estudio. Este enfoque nos permite incorporar la variabilidad tanto dentro como entre los estudios, lo que conduce a una estimación más precisa del tamaño del efecto general.

La heterogeneidad entre los tamaños de efecto se ha evaluado utilizando la estadística Q y se ha cuantificado con el índice I^2 . La varianza entre estudios se ha estimado utilizando el método de DerSimonian y Laird. También se han incluido los intervalos de confianza correspondientes para las estimaciones de heterogeneidad. Este enfoque nos permite tener en cuenta la variabilidad tanto dentro como entre los estudios, proporcionando una estimación más precisa del tamaño del efecto general.

Para calcular el tamaño del efecto de cada estudio y el efecto agregado, se ha utilizado el software R versión 2023.9.0.463 (Posit Team, 2023) con el paquete «meta» (Balduzzi et al., 2019). El modelo de efectos aleatorios implementado en este paquete permite incorporar la variabilidad tanto dentro como entre los estudios, y proporciona una estimación robusta del tamaño del efecto general. El enfoque frecuentista utilizado en este análisis implica calcular estimaciones puntuales e intervalos de confianza para hacer inferencias sobre los efectos poblacionales.

Detección de valores atípicos

En el metanálisis, se utilizan típicamente efectos aleatorios para tener en cuenta la variación no explicada más allá de la variabilidad dentro del estudio, pero este método puede no ser suficiente al tratar con estudios atípicos (Beath, 2014). Es crucial considerar la identificación de valores atípicos (Riahi y Mokhayeri, 2017) porque estos valores atípicos poseen un tamaño de efecto tan extremo que diverge notablemente del efecto general. Se ha utilizado el software R versión 2023.9.0.463 (Posit Team, 2023) para detectar el(s) valor(es) atípico(s), empleando los paquetes «metafor» y «dmetar». El paquete «dmetar» incorpora la función «find.outliers», implementando el algoritmo de eliminación de valores atípicos (Harrer et al., 2021). Además, se ha utilizado la función «gosh.diagnostic», que utiliza diferentes algoritmos de agrupamiento para identificar patrones en los datos (Harrer et al., 2021). En este estudio, se ha seleccionado específicamente el algoritmo de k-medias (Hartigan y Wong, 1979) de entre estos algoritmos.

Análisis de variables moderadoras

Analizar las variables moderadoras es crucial en el metanálisis, ayudando a los investigadores a entender los factores que influ-

yen en las variaciones en los tamaños de efecto entre estudios e informando la evaluación de intervenciones actuales y la posible formulación de otras más impactantes (Li et al., 2020). En este estudio, las variables moderadoras a analizar incluyen el tamaño de la muestra, el nivel educativo, el continente donde se ha realizado el estudio, así como el tratamiento de control y la duración de la intervención. La exploración de la diversidad en los resultados de la investigación entre varios estudios ha implicado la utilización de una prueba de heterogeneidad (prueba Q). Un hallazgo notable en la estadística Q indica la posible consideración de cada estudio como originado de una población común. Esencialmente, las variaciones sustanciales en el impacto colectivo de cada elemento dentro de la variable moderadora, destacadas por una estadística Q significativa, subrayan la posible importancia del análisis de variables moderadoras. El análisis de todas las variables moderadoras en este estudio se ha realizado utilizando modelos similares a ANOVA. Dentro de estos modelos, se han presentado los efectos medios (g) dentro de los grupos, intervalos de confianza del 95% (IC) y heterogeneidad entre grupos (Q_b). Una estadística Q_b notable señala distinciones significativas en los efectos agregados entre los componentes de la variable moderadora. El análisis de las variables moderadoras ha sido facilitado por el software R versión 2023.9.0.463 (Posit Team, 2023) con el paquete «meta» (Balduzzi et al., 2019).

Evaluación del sesgo de publicación

La evaluación del sesgo de publicación tiene como objetivo identificar sesgos potenciales en la literatura y su impacto en las conclusiones generales del metanálisis. Para evaluar el sesgo de publicación en este estudio de metanálisis, se ha adoptado un enfoque de tres frentes. El examen ha involucrado la implementación de gráficos de embudo y la prueba de regresión de Egger. La inspección visual del gráfico de embudo se empleó como un primer paso para discernir cualquier asimetría. Una distribución simétrica dentro del gráfico de embudo se consideró indicativa de la ausencia de sesgo de publicación (Card, 2011). Esta representación gráfica ha facilitado la identificación de posibles valores atípicos y ha proporcionado una impresión visual general del panorama del sesgo de publicación. Posteriormente, se ha aplicado la prueba de regresión de Egger, reconocida por su evaluación cuantitativa de la asimetría del gráfico de embudo, para evaluar rigurosamente la presencia de sesgo de publicación. Este enfoque estadístico ha cuantificado el grado de asimetría dentro del gráfico de embudo, aportando una dimensión cuantitativa a la evaluación del sesgo. La evaluación ha tenido como objetivo discernir si existía una relación sistemática entre los tamaños de efecto y su precisión, lo que podría indicar un sesgo de publicación.

Además de los métodos mencionados, también hemos aplicado el método Trim and Fill para evaluar aún más el sesgo de publicación. Este es uno de los métodos más comunes para ajustar la asimetría del gráfico de embudo (Duval y Tweedie, 2000). El método Trim and Fill implica identificar posibles estudios faltantes del gráfico de embudo y ajustar el metanálisis para tener en cuenta estos estudios faltantes, proporcionando así una estimación más precisa del tamaño del efecto. Este enfoque integral nos ha permitido abordar sesgos potenciales de manera más robusta y mejorar la credibilidad de nuestros hallazgos.

Para ejecutar el análisis del sesgo de publicación, se ha utilizado el software R versión 2023.9.0.463 (Posit Team, 2023) junto con el paquete «meta» (Balduzzi et al., 2019) y el paquete «metafor» (Viechtbauer, 2010). La integración de estas herramientas estadísticas ha permitido una exploración integral del sesgo de publicación, reforzando la solidez y credibilidad de los hallazgos del metanálisis al esclarecer el impacto del aprendizaje basado en juegos en las habilidades de pensamiento analítico en matemáticas de los estudiantes.

Resultados

Descripción general del estudio

Durante el proceso de revisión de la literatura y selección, se han identificado un total de trece fuentes relevantes, que comprenden estudios experimentales que investigan el impacto del aprendizaje basado en juegos (ABJ) en las habilidades analíticas matemáticas de los estudiantes. La síntesis de estos 13 artículos ha dado lugar a un conjunto de datos que abarca 42 estudios individuales, cada uno de los cuales contribuye con un tamaño de efecto distinto al metanálisis. Los estudios incluidos en este metanálisis cumplen con los criterios de inclusión pre establecidos: son artículos de revistas revisadas por pares o actas de conferencias publicadas en inglés entre 2010 y 2024. Estos estudios investigan empíricamente el impacto del ABJ en las habilidades de pensamiento de orden superior (HPOS) en la educación matemática en todos los niveles (desde primaria hasta pregrado). El diseño de investigación debe ser experimental, involucrando al menos un grupo que use ABJ y un grupo de control con intervenciones tradicionales o no basadas en juegos (no-ABJ). Además, los estudios deben proporcionar suficientes datos para calcular los tamaños de efecto.

Las características de los estudios incluidos en este metanálisis, junto con la distribución detallada de estos estudios, se presentan en el [Apéndice \(Tabla 1\)](#). Esta tabla proporciona una visión general de las diversas fuentes de literatura, tipos de publicación, habilidades de pensamiento medidas, duraciones de intervención, tamaños de muestra, niveles educativos, continentes donde se han realizado los estudios y tratamientos de control. Un total de 42 estudios están involucrados, que comprenden 40 artículos de revistas y dos actas de conferencias, realizados en cuatro continentes diferentes: Asia ($n = 12$), América del Norte ($n = 3$), América del Sur ($n = 1$) y Europa ($n = 22$). Estos estudios examinan el impacto del ABJ en HPOS en estudiantes de matemáticas en los siguientes niveles educativos: escuela primaria ($n = 29$), escuela secundaria ($n = 6$) y pregrado ($n = 5$). La investigación se centra en varios aspectos de

HPOS, incluyendo la resolución de problemas ($n = 3$), razonamiento matemático ($n = 18$) y otras habilidades como el pensamiento computacional, pensamiento crítico y pensamiento divergente ($n = 6$). En los grupos experimentales, los tamaños de muestra se categorizan en dos grupos: muestras pequeñas ($n < 30$) en 11 estudios y muestras grandes ($n < 3$) en 29 estudios. La duración de la intervención ABJ varía y se categoriza en cuatro grupos: 12 estudios con duraciones de intervención de menos de diez horas, tres estudios con duraciones tanto de menos de diez horas como de más de 20 horas, y los estudios restantes carecen de información detallada sobre la duración. En los grupos de control, los métodos de enseñanza incluyen instrucción basada en multimedia en 20 estudios, mientras que el resto utiliza métodos de enseñanza tradicionales.

Para proporcionar una perspectiva integral sobre los estudios incluidos, se ha llevado a cabo un examen de las características clave. Estas características se organizaron en cuatro categorías principales: literatura, diseño del estudio, muestra y características de la intervención (juego). Dentro de cada categoría, se han identificado subcaracterísticas, que han servido como variables moderadoras que proporcionan una perspectiva detallada sobre los estudios. En este estudio, dos evaluadores han valorado la codificación de la literatura realizada por los investigadores. La consistencia entre evaluadores se ha medido utilizando la estadística kappa, que ha arrojado un valor kappa de .78, indicando un «acuerdo sustancial» entre los evaluadores. Esto ha confirmado que las decisiones de los dos evaluadores eran altamente consistentes, lo que respalda la fiabilidad de la codificación de la literatura realizada por los investigadores ([Ho et al., 2019](#)).

Resultados del análisis principal antes de detectar valores atípicos

El análisis utilizando el modelo de efectos aleatorios ha revelado que el tamaño del impacto promedio en los 42 estudios ha sido de .266 ($p < .001$), y el intervalo de confianza del 95% ha oscilado entre .086 y .446 ([Figura 2](#)). Estos hallazgos indican

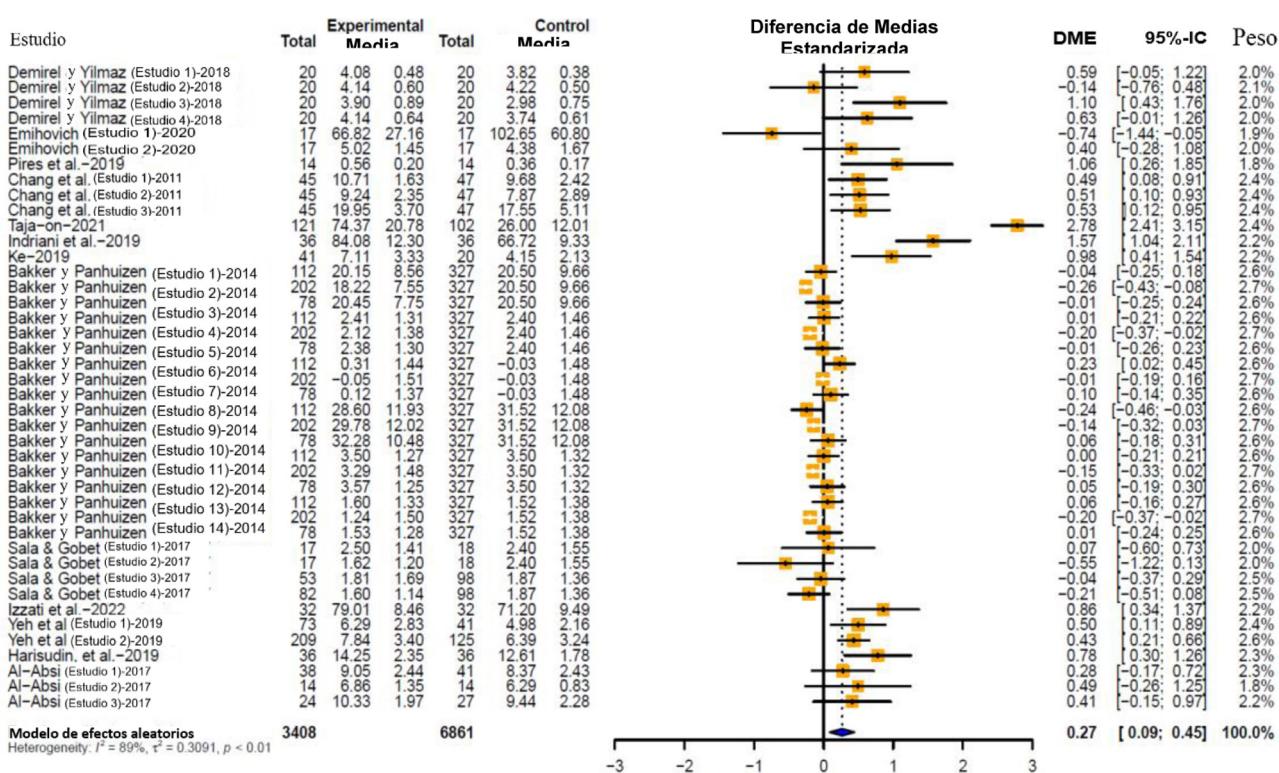


Figura 2. Gráfico de bosque para 42 estudios (detección previa de valores atípicos).

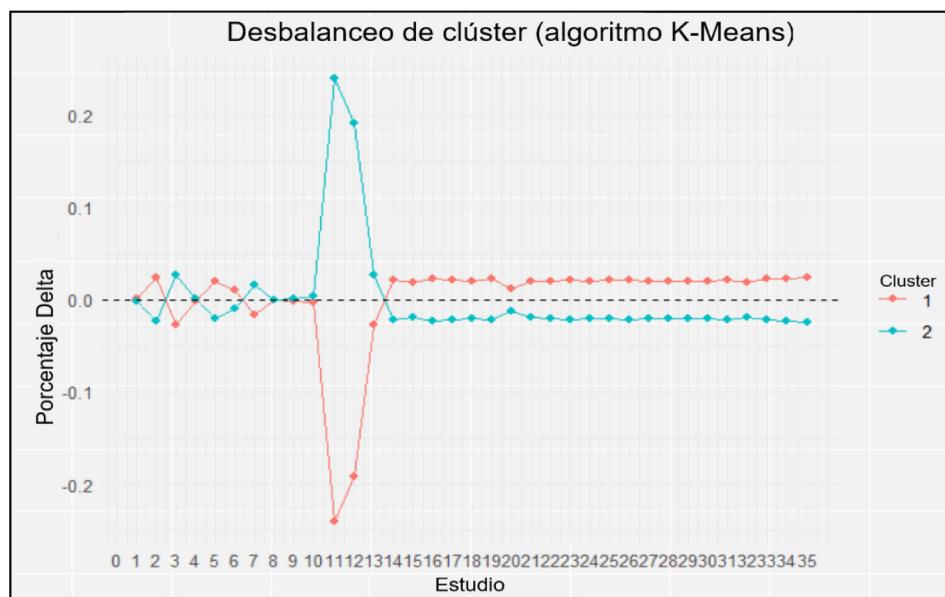


Figura 3. Gráfico de K-Means para detectar valores atípicos.

un impacto estadísticamente significativo del ABJ en las habilidades de pensamiento de orden superior (HPOS) en matemáticas de los estudiantes en comparación con los métodos de aprendizaje tradicionales (no-ABJ), aunque el tamaño del efecto agregado cae en la categoría de «moderado», según lo definido por Cohen (2013), donde los tamaños de efecto entre 0.20 y 0.50 se consideran moderados para estudios experimentales. Así, el impacto positivo de la implementación del ABJ en las HPOS matemáticas, en comparación con los métodos tradicionales, se encuentra dentro de la categoría de tamaño de efecto medio/moderado. Además, el análisis también ha revelado diversos tamaños de efecto entre los 42 estudios ($Q=385.71$, $gl=41$, $p<.001$), con un nivel de heterogeneidad (I^2) alcanzando el 89.4%. Estos hallazgos indican una variación considerable en los tamaños de efecto empleados en este estudio.

Detección de valores atípicos

Identificar valores atípicos, puntos de datos que exhiben una desviación considerable del estándar, puede tener una influencia notable en los resultados generales. Por lo tanto, se recomienda eliminar los valores atípicos y posteriormente reevaluar el impacto acumulativo (Riahi y Mokhayeri, 2017). Los gráficos del análisis de K-means (Figura 3) ilustran estudios con puntos de datos posicionados significativamente distantes de la línea horizontal (estudios 11 y 12), reconocidos como valores atípicos. Los resultados del análisis descubren la existencia de dos estudios identificados como valores atípicos, que abarcan la investigación realizada por Taja-on (2019) e Indriani et al. (2019).

Análisis del efecto resumen después de eliminar valores atípicos

Después de la exclusión de dos estudios identificados como valores atípicos, se ha llevado a cabo una reevaluación del impacto colectivo de los 40 estudios restantes. En estos estudios, los tamaños de efecto (TE) varían de -.74 a 1.10 (Figura 4). Entre ellos, 15 estudios (37.50%) muestran un tamaño de efecto negativo, lo que indica que el ABJ no fue más efectivo en comparación con el aprendizaje tradicional (no-ABJ) en el grupo de control. Por el contrario, 25 estudios (62.50%) muestran un tamaño de efecto positivo, sugiriendo que los estudiantes en el grupo experimental, que implementaron ABJ en matemáticas, lograron habilidades

de pensamiento de orden superior (HPOS) en matemáticas más altas que aquellos en el grupo de control con enfoques no basados en juegos. En resumen, los resultados revelan un tamaño de efecto general positivo y significativo con $g=.134$ ($p<.001$, 95% CI = [.028; .239]) considerablemente más bajo que el tamaño de efecto original antes de eliminar los estudios atípicos ($g=.27$), destacando la importancia de la detección de valores atípicos en los metanálisis (Viechtbauer y Cheung, 2010). En contraste con la condición antes de excluir los valores atípicos, el tamaño de efecto general ahora cae dentro de la categoría de «tamaño de efecto pequeño» (Cohen, 2013). La Figura 4 proporciona el gráfico de bosque, ilustrando los tamaños de efecto de los estudios y los intervalos de confianza. Además, la prueba de heterogeneidad ha arrojado un resultado significativo ($Q=143.62$, $gl=39$, $p<.001$), indicando variación en los tamaños de efecto ($I^2 = 72.8\%$). Según Higgins et al. (2003), un valor I^2 superior al 75% sugiere una heterogeneidad considerable, indicando que una proporción sustancial de la varianza observada refleja diferencias reales en los tamaños de efecto en lugar de un error de muestreo. Esta heterogeneidad subraya la necesidad de explorar posibles moderadores que podrían explicar la variabilidad en la efectividad del ABJ (Borenstein et al., 2009).

Evaluación del sesgo de publicación

El gráfico de embudo en la Figura 5 representa visualmente los resultados de una evaluación del sesgo de publicación utilizando el método de recorte y llenado (Duval y Tweedie, 2000). Este análisis ha revelado un sesgo de publicación sustancial en el metanálisis. La asimetría observada en el gráfico, con menos estudios en el lado izquierdo, ha sugerido inicialmente que faltaban estudios con hallazgos negativos o nulos. El método de recorte y llenado ha confirmado esto, identificando 11 estudios potencialmente faltantes. Crucialmente, después de ajustar por estos estudios faltantes, los resultados han cambiado drásticamente. El tamaño del efecto reportado originalmente se ha reducido significativamente y se ha vuelto estadísticamente no significativo (DME = -.019, 95% CI [-.154; .116], $p=.781$). Este cambio sustancial sugiere que el sesgo de publicación puede haber inflado los hallazgos originales, un problema común en la investigación educativa (Pigott et al., 2013). Esto se alinea con los desafíos conocidos en la investigación edu-

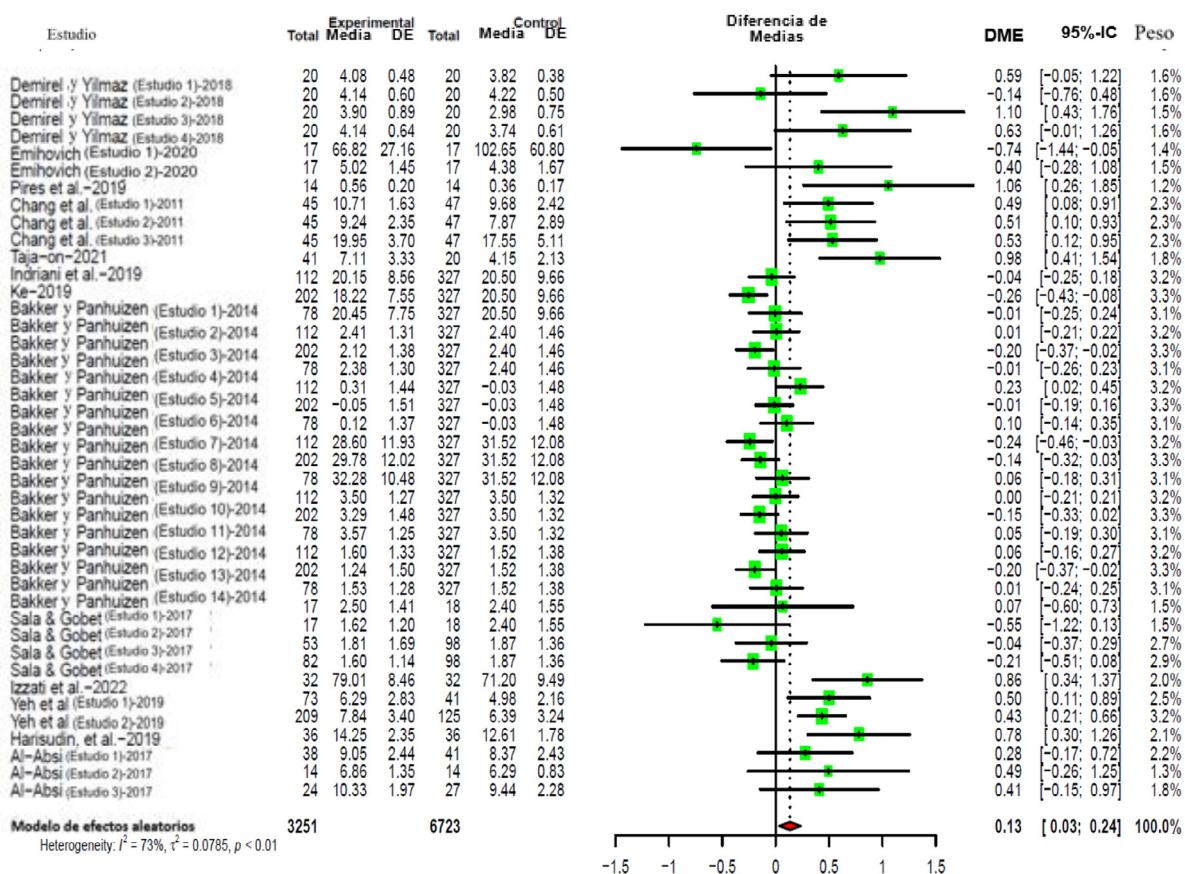


Figura 4. Gráfico de bosque para 40 estudios (detección posterior de valores atípicos).

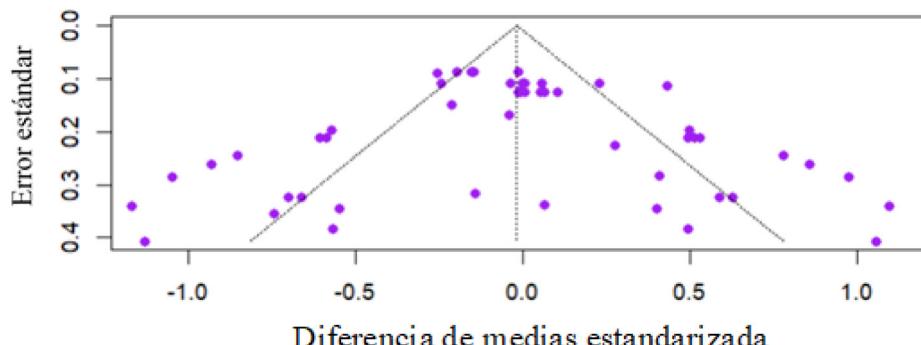


Figura 5. Gráfico de embudo utilizando el método de recorte y llenado.

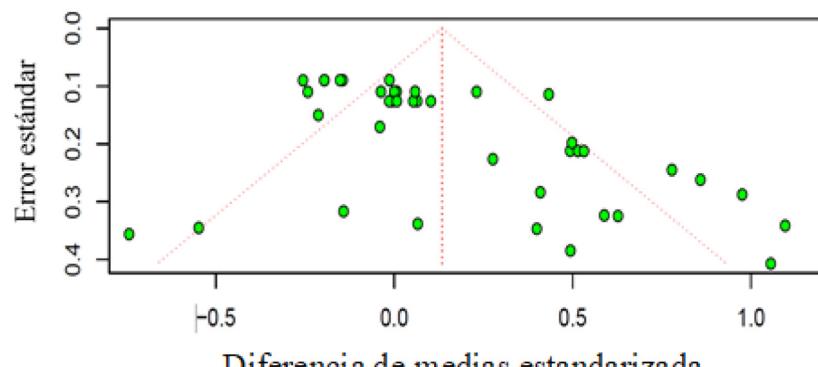


Figura 6. Gráfico de embudo utilizando el método de Egger.

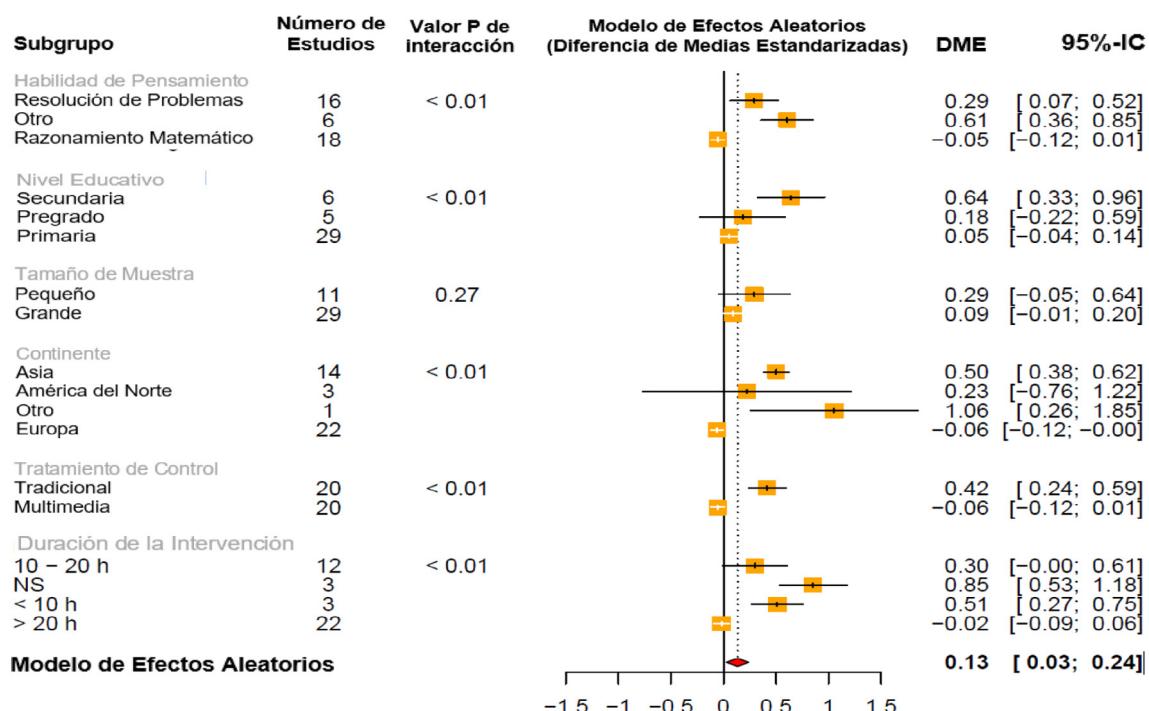


Figura 7. Gráfico de bosque de subgrupos (variables moderadoras).

cativa, donde los estudios con hallazgos positivos o significativos son más propensos a ser publicados, lo que puede sesgar los resultados meta-analíticos. El gráfico de embudo y el análisis asociado sirven como un recordatorio crítico de la importancia de abordar el sesgo de publicación en la síntesis de investigaciones y la necesidad de tener precaución al interpretar hallazgos meta-analíticos no ajustados en campos propensos a dicho sesgo.

Curiosamente, la prueba de Egger (Egger et al., 1997) ha proporcionado un resultado contrastante ($t=4.55, p=.082$), no indicando un sesgo de publicación significativo. El gráfico de embudo asociado (Figura 6) muestra una distribución relativamente simétrica de los tamaños de efecto. Esta discrepancia entre los resultados del método de recorte y llenado y la prueba de Egger no es poco común en los metanálisis y resalta la importancia de utilizar múltiples métodos para evaluar el sesgo de publicación (Sutton et al., 2000). También sugiere que diferentes tipos de sesgo de publicación podrían estar en juego, como el sesgo de reporte de resultados o el sesgo de retraso temporal (Dwan et al., 2008; Ioannidis y Trikalinos, 2005). El gráfico de embudo (Figura 6) revela asimetría, particularmente una falta de estudios en el cuadrante inferior izquierdo, sugiriendo un potencial sesgo de publicación en este metanálisis. Esta asimetría indica que los estudios más pequeños con resultados negativos o nulos pueden estar subrepresentados, un fenómeno conocido como el «problema del cajón de archivos» (Rothstein, 2005). El gráfico muestra una tendencia donde los estudios más pequeños (con mayores errores estándar) tienden a reportar efectos positivos más grandes, mientras que hay una notable ausencia de estudios más pequeños que reporten efectos negativos. Este patrón se alinea con la descripción de Sterne et al. (2011) sobre el sesgo de publicación en gráficos de embudo y apoya los hallazgos anteriores del método de recorte y llenado. Si bien la prueba de Egger puede no haber indicado un sesgo significativo, la inspección visual puede revelar patrones que las pruebas estadísticas podrían pasar por alto (Egger et al., 1997). Esta asimetría subraya la necesidad de una interpretación cautelosa de los resultados meta-analíticos y resalta el impacto potencial del sesgo de publicación en la investigación educativa, donde los estudios pequeños son comunes (Ioannidis y Trikalinos, 2007).

El cambio sustancial en el tamaño del efecto después de aplicar el método de recorte y llenado subraya la importancia crítica de abordar el sesgo de publicación en los metanálisis. Como argumenta Ioannidis (2005), el sesgo de publicación puede llevar a una sobreestimación de los tamaños de efecto y potencialmente a hallazgos falsos positivos en la literatura. En este caso, el efecto positivo pequeño inicial del ABJ en las HPOS matemáticas parece ser en gran medida atribuible al sesgo de publicación, sugiriendo que el efecto real puede ser insignificante o inexistente al considerar el panorama completo de los estudios realizados, incluidos aquellos potencialmente no publicados.

Análisis de variables moderadoras

Los resultados del análisis de variables moderadoras, que abarcan las habilidades de pensamiento medidas, el nivel educativo, el tamaño de la muestra (Li et al., 2020), el continente donde se ha realizado el estudio, el tratamiento de control y la duración de la intervención (Prieto-Rodríguez et al., 2020), se representan en la Figura 7.

Habilidades de pensamiento medidas

La variable moderadora inicial «habilidades de pensamiento medidas» ha consistido en cuatro grupos: resolución de problemas, pensamiento crítico, razonamiento matemático y otras habilidades. Sin embargo, después de la eliminación de valores atípicos, solo han quedado tres grupos de habilidades de pensamiento, ya que ambos estudios eliminados han pertenecido a la categoría de pensamiento crítico. Además, el resultado del análisis muestra que hay una diferencia significativa en la efectividad de ABJ según el tipo de habilidad de pensamiento medida ($Q_b = 32.63, p < .001$). El análisis revela que ABJ tiene un efecto positivo pero pequeño en el aumento de las habilidades de resolución de problemas de los estudiantes, con un tamaño de efecto de $g = .29$. Entre los tres grupos de habilidades de pensamiento medidas, ABJ ha demostrado una efectividad positiva notable para «otras habilidades», como pensamiento crítico, pensamiento computacional y pensamiento divergente, con

un tamaño de efecto más grande de $g = .61$. En contraste, para las habilidades de razonamiento matemático (con $g = -.05$ y un intervalo de confianza del 95% de [-.12; .01]) no es estadísticamente significativo, ya que el intervalo de confianza incluye cero. Esto indica que ABJ no muestra una diferencia significativa en la mejora de las habilidades de razonamiento matemático de los estudiantes en comparación con las intervenciones no-ABJ.

Nivel educativo

En términos de nivel educativo, este estudio ha investigado tres grupos: nivel elemental, secundario y universitario. Los resultados del análisis revelan diferencias significativas en el efecto de ABJ sobre las HPOS matemáticas de los estudiantes según el nivel educativo de los estudiantes ($Q_b = 15.09, p = .0017$). Indica que el nivel educativo influye significativamente en la efectividad de ABJ sobre las HPOS matemáticas de los estudiantes en comparación con el aprendizaje tradicional (no-ABJ). Entre los tres grupos de materias medidas, la implementación de ABJ se muestra más efectiva en el nivel de secundaria ($g = .78$ y 95% CI [.33; .96]). Mientras que, los otros dos niveles educativos (elemental y universitario) han mostrado tamaños de efecto no significativos en los cuales su intervalo de confianza incluye cero (Figura 7).

Tamaño de la muestra

Entre las seis variables moderadoras identificadas, solo se ha encontrado que el tamaño de la muestra no tenía un impacto significativo en la mejora de las HPOS de los estudiantes en ABJ. La variable moderadora del tamaño de la muestra se clasifica en dos grupos: muestras grandes y pequeñas. Los hallazgos sugieren que no hay diferencias significativas en el impacto de ABJ sobre las HPOS matemáticas de los estudiantes al considerar el tamaño de la muestra del estudio ($Q_b = 1.22, p = .27$) (Figura 7).

Continente

La variable moderadora para el continente donde se ha realizado el estudio se clasifica en cuatro grupos: Asia, América del Norte, Europa y otros. Los hallazgos revelan variaciones notables en los tamaños de efecto promedio entre estos grupos continentales ($Q_b = 76.73, p < .001$). El p general $<.01$ (Figura 7) indica una diferencia significativa en la efectividad de ABJ a través de los continentes analizados. Asia muestra un tamaño de efecto positivo y significativo con $g = .50$. La categoría «Otros», que incluye América del Sur (ver Tabla 1 Apéndice), muestra un tamaño de efecto aún mayor con $g = 1.06$ (95% CI [.26; 1.85]). Aunque el tamaño de efecto es sustancial y el intervalo de confianza no incluye cero, esta categoría incluye solo un estudio, por lo que los resultados deben interpretarse con precaución. La falta de estudios comparativos en esta región dificulta sacar conclusiones sólidas sobre la efectividad de ABJ en América del Sur. En contraste, América del Norte presenta un tamaño de efecto más pequeño, estadísticamente no significativo con $g = .23$, indicando un impacto menos pronunciado de ABJ en esta región. Europa también muestra un tamaño de efecto no significativo con $g = -.06$ (95% CI [-.12; .00]), sugiriendo que ABJ no presenta una diferencia clara en comparación con los métodos de aprendizaje convencionales en Europa.

Duración de la intervención

La duración del tratamiento en el grupo experimental se categoriza en cuatro grupos: menos de diez horas, 10-20 horas, más de 20 horas y no especificado. Este estudio ha revelado diferencias significativas en el impacto de ABJ en el aumento de las HPOS matemáticas de los estudiantes en comparación con no-ABJ, según la duración de la intervención ($Q_b = 43.44, p < .001$). Además, se sugiere que dos grupos dentro de esta variable de duración de la

intervención (menos de 10 horas y no especificado) ejercen una influencia. Sin embargo, en la categoría de duración de intervención de 10 a 20 horas, los resultados no son estadísticamente significativos porque el intervalo de confianza incluye 0 (95% CI [-.00; .61]). Además, en los estudios con la duración de tratamiento más larga (> 20 horas), no se ha encontrado que ABJ fuera más efectivo que no-ABJ en la mejora de las HPOS matemáticas de los estudiantes (Figura 7).

Tratamiento de control

La variable de tratamiento de control se divide en dos grupos, incluyendo tradicional y multimedia. Los resultados del análisis indican diferencias significativas en el impacto de ABJ en las HPOS matemáticas de los estudiantes según el tipo de tratamiento de control aplicado ($Q_b = 24.01, p < .001$). Esto sugiere que el tratamiento de control elegido influye significativamente en la efectividad del aprendizaje basado en juegos para mejorar las HPOS matemáticas de los estudiantes en comparación con el aprendizaje no basado en juegos. Entre los dos grupos de tratamiento de control medidos, la implementación de ABJ es más efectiva en comparación con el aprendizaje tradicional no tecnológico o el aprendizaje asistido por multimedia ($g = .42, p < .001$) (Figura 7).

Discusión

A pesar de la creciente tendencia de incorporar juegos en el proceso de aprendizaje, que ha llevado a numerosas investigaciones empíricas sobre la efectividad del ABJ, los estudios individuales muestran hallazgos variados respecto a su influencia en las habilidades de pensamiento académico matemático de los estudiantes. Además, falta una síntesis completa de los resultados de estos estudios. Esta investigación busca cerrar esta brecha a través de un enfoque de metanálisis. Los resultados de esta investigación sugieren que, en general, la implementación del aprendizaje basado en juegos afecta positivamente las habilidades de pensamiento de orden superior en matemáticas de los estudiantes.

El alcance actual de este estudio de metanálisis se centra en la efectividad del ABJ en las HPOS matemáticas de los estudiantes (Chang et al., 2012). Debido a las limitaciones que surgen de la disponibilidad de datos limitados, existen desafíos para comparar directamente los resultados de esta investigación con otros estudios idénticos. Sin embargo, es posible realizar comparaciones entre los resultados de este estudio y metanálisis relevantes previos. Los hallazgos de esta investigación corresponden con los resultados de numerosos estudios meta-analíticos anteriores que han indagado sobre la efectividad del ABJ en el contexto de los logros generales en matemáticas.

Tras la exclusión de valores atípicos, el análisis del efecto resumido revela un tamaño de efecto positivo general ($g = .13$), lo que significa una efectividad estadísticamente significativa del ABJ, en comparación con el aprendizaje tradicional no-ABJ, para mejorar las HPOS matemáticas de los estudiantes. Sin embargo, es importante señalar que el tamaño del efecto se encuentra dentro de la categoría de «tamaño de efecto pequeño», tal como lo define Cohen (2013). Este hallazgo difiere notablemente de varios estudios de metanálisis anteriores que han informado de efectos grandes o efectos moderados/medios del ABJ en el rendimiento cognitivo matemático de los estudiantes, como los hallazgos de Turgut y Temur (2017) y Byun y Joung (2018), que indican un efecto «moderado/medio» ($g = .792$ y $g = .370$, respectivamente). Sin embargo, los resultados de este estudio no muestran una diferencia sustancial con los hallazgos observados en el estudio de Tokac et al. (2019), revelando el tamaño de efecto general «pequeño pero marginalmente significativo» ($g = .13, p = .02$). Los resultados sugieren que, a pesar de que múltiples metanálisis indican consistentemente la influen-

cia positiva del ABJ en el rendimiento cognitivo de los estudiantes en matemáticas, el tamaño del efecto agregado exhibe variabilidad.

Varios estudios anteriores también han explorado el impacto del ABJ en las HPOS de los estudiantes (Harisudin et al., 2019), pero con un enfoque de investigación más amplio, no limitado solo al campo de las matemáticas. En comparación, los resultados de este estudio muestran consistencia con esos estudios anteriores, indicando resultados positivos. Sin embargo, la categoría del tamaño del efecto agregado encontrado difiere ligeramente de los resultados de estudios de metanálisis anteriores, como la investigación que examina el efecto del ABJ en las habilidades de pensamiento computacional, como se indica en los estudios de Lu et al. (2023) y Ma et al. (2023), que arrojan valores de $g = .677$ y $g = .600$, ambos en la categoría de «moderado» según los criterios de Cohen (Cohen, 2013). Además, otro estudio que examina las HPOS en el contexto del pensamiento crítico ha reportado tamaños de efecto categorizados como «grandes» con $g = .863$ (Mao et al., 2022). Estos hallazgos refuerzan que el ABJ tiene un efecto positivo en la mejora de las HPOS de los estudiantes, aunque el tamaño del efecto general sigue siendo dinámico.

Además de la información sobre el tamaño del efecto agregado, este estudio de metanálisis también proporciona información sobre los resultados del análisis de variables moderadoras, incluidas las habilidades de pensamiento medidas, el nivel educativo, el tamaño de la muestra, el continente, la duración de la intervención y el tratamiento de control. Estos hallazgos ofrecen una comprensión integral de cómo diversos factores contribuyen al impacto del ABJ en las HPOS matemáticas de los estudiantes. La exploración de la variable «habilidades de pensamiento medidas» indica una diferencia significativa en el tamaño del impacto promedio entre estos grupos, sugiriendo que la efectividad del ABJ en las HPOS de los estudiantes se ve influenciada significativamente por las habilidades de pensamiento específicas medidas. Estos hallazgos refuerzan los metanálisis anteriores que han reportado un efecto positivo del ABJ en las HPOS matemáticas de los estudiantes, específicamente para ciertas habilidades de pensamiento (Aizpurua et al., 2018; Suherman y Vidákovich, 2024), como el pensamiento crítico (Mao et al., 2022) y la habilidad de pensamiento computacional (Lu et al., 2023; Ma et al., 2023).

Este estudio ha examinado cuatro niveles educativos y ha encontrado diferencias significativas en el impacto del ABJ en las HPOS matemáticas de los estudiantes según el nivel académico. Estos hallazgos coinciden con el estudio de Turgut y Temur (2017), indicando una diferencia estadísticamente significativa entre los grupos en relación con los niveles educativos, sugiriendo que el nivel académico es un factor que contribuye a proporcionar un efecto positivo de la educación matemática asistida por juegos en el logro académico (Bakker et al., 2015). En contraste, estos hallazgos contradicen los resultados de un metanálisis previo, que ha concluido que el impacto del ABJ en el rendimiento cognitivo de los estudiantes no se ve afectado por sus niveles académicos (Wang et al., 2022).

Además, se ha encontrado que el tamaño de la muestra fue la única variable moderadora que no impactó significativamente la mejora de las HPOS matemáticas de los estudiantes en el ABJ. Este hallazgo es consistente con los resultados de un estudio de metanálisis anterior (Ma et al., 2023), que indica que el factor del tamaño de la muestra no ha influido significativamente en el efecto del ABJ en el pensamiento crítico de los estudiantes. Aunque no es impactante de manera significativa, un aspecto interesante de los hallazgos de este estudio es que se ha considerado que el ABJ fue más efectivo para mejorar las HPOS matemáticas de los estudiantes en tamaños de muestra pequeños en comparación con tamaños de muestra grandes. Las variaciones en el tamaño del efecto pueden verse influenciadas por el tamaño de la muestra, siendo las muestras más pequeñas típicamente las que producen tamaños de

efecto más grandes en comparación con las muestras más grandes (Slavin y Smith, 2009; Sun et al., 2023). La influencia del tamaño de la muestra en las variaciones del tamaño del efecto sugiere que el ABJ tiende a mostrar resultados de aprendizaje más efectivos con tamaños de muestra más pequeños, una tendencia respaldada por varios estudios anteriores (Petri y Gresse von Wangenheim, 2017).

Otro factor que ha influido significativamente en la efectividad del ABJ para mejorar las HPOS matemáticas de los estudiantes es el continente donde se ha llevado a cabo el estudio (Pires et al., 2019; Sala y Gobet, 2017). Se han observado variaciones notables entre continentes, siendo el ABJ más efectivo en el grupo de «Otros» países, particularmente en América del Sur, seguido de Asia en segundo lugar. Este hallazgo tiene poca similitud con un estudio de metanálisis anterior que indica que el uso de juegos en el aprendizaje, como enfoque de aprendizaje, parece arrojar resultados más favorables en contextos asiáticos en comparación con otros continentes (Bai et al., 2020). Sin embargo, este resultado sigue siendo cuestionable ya que hay un apoyo empírico limitado respecto a la influencia de un país o continente (Tokac et al., 2019) en el impacto del ABJ en las HPOS matemáticas. Estos resultados necesitan ser examinados más a fondo en futuras investigaciones.

El tipo de tratamiento de control también ha influido significativamente en la efectividad del ABJ en las HPOS matemáticas de los estudiantes. El ABJ ha superado al aprendizaje tradicional no tecnológico o asistido por multimedia, enfatizando la importancia de seleccionar cuidadosamente los tratamientos de control en diseños experimentales que involucren ABJ. Este hallazgo contradice los resultados de un estudio previo (Wang et al., 2022), que no ha encontrado diferencias estadísticamente significativas entre el grupo de control tratado con enfoques tradicionales y multimedia en relación con el impacto del aprendizaje basado en juegos digitales en el logro cognitivo de los estudiantes. Se ha realizado poca investigación al respecto, por lo que se necesita una investigación adicional para explorar esta área.

A su vez, se ha encontrado que la variable moderadora de la duración de la intervención/tratamiento influye significativamente en la efectividad del ABJ en las HPOS matemáticas de los estudiantes. La duración del tratamiento en el grupo experimental ha desempeñado un papel crucial, y las intervenciones con duraciones prolongadas suelen exhibir efectos más fuertes (Korkman et al., 1999). Contradicatoriamente, los hallazgos de este estudio indican brevemente que el ABJ resulta efectivo en diversas duraciones, pero su efectividad disminuye en estudios con duraciones de tratamiento que superan las 20 horas (el grupo con la duración más larga). Este resultado podría parecer paradójico, ya que se podría anticipar que las intervenciones más largas serían más efectivas que las más cortas (Tokac et al., 2019). Al desestimar la «duración no especificada» en este subgrupo de variable moderadora, la duración más corta (menos de diez horas) ha proporcionado el impacto más fuerte. Esto es bastante consistente con los resultados de un estudio de metanálisis anterior (Lu et al., 2023), que ha reportado que las intervenciones de ABJ entre cuatro horas y una semana mostraron el efecto más fuerte en el pensamiento crítico de los estudiantes. Esto subraya la importancia de optimizar la duración de la intervención para lograr un impacto óptimo dentro de una duración intermedia, que no sea demasiado corta ni demasiado larga, para evitar causar aburrimiento en los estudiantes.

Los hallazgos de este metanálisis sugieren que, si bien el efecto general es relativamente pequeño, el ABJ impacta positivamente en las HPOS matemáticas de los estudiantes. Esto indica que el ABJ puede ser un enfoque suplementario valioso para mejorar las HPOS matemáticas en diferentes niveles educativos. Sin embargo, los educadores y los responsables de políticas deben ser conscientes de variables específicas que influyen en la efectividad del ABJ, como la alineación entre el juego utilizado y las habilidades de pensamiento objetivo, el nivel educativo de los estudiantes y la

duración de la intervención. Dadas las limitaciones del estudio, incluido el pequeño número de estudios empíricos y las restricciones de idioma, la investigación futura debe buscar incluir un rango más amplio de estudios y explorar variables moderadoras adicionales. Esto ayudará a proporcionar una comprensión más completa de cómo se puede optimizar el ABJ para apoyar el pensamiento matemático de los estudiantes y mejorar las prácticas educativas.

Por último, nuestro metanálisis ha revelado inicialmente un efecto positivo pequeño pero estadísticamente significativo. Sin embargo, al aplicar el método Trim and Fill para ajustar el posible sesgo de publicación, hemos observado un cambio sustancial tanto en la magnitud como en la dirección del tamaño del efecto. El efecto ajustado se ha vuelto ligeramente negativo y estadísticamente no significativo. Este cambio marcado se alinea con los hallazgos de Carter et al. (2019), quienes han demostrado que el sesgo de publicación puede llevar a una sobreestimación de los tamaños de efecto en la investigación psicológica. El cambio sustancial que hemos observado sugiere que el sesgo de publicación probablemente ha ejercido una influencia considerable en nuestros hallazgos originales, en consonancia con el meta-metanálisis de Fanelli et al. (2017), que ha encontrado que el sesgo de publicación es prevalente en todas las disciplinas científicas. Si bien la prueba de Egger no ha indicado sesgo significativo, el cambio notable en el tamaño del efecto tras el ajuste no puede pasarse por alto. Esta discrepancia entre diferentes métodos de evaluación del sesgo de publicación refleja el trabajo de Renkewitz y Keiner (2019), quienes han enfatizado la importancia de utilizar múltiples enfoques para detectar y corregir el sesgo de publicación. El intervalo de confianza del efecto ajustado ahora incluye cero, lo que apoya aún más la posibilidad de que el efecto verdadero pueda ser despreciable o inexistente. Este cambio dramático de un efecto positivo significativo a un resultado no significativo plantea importantes preguntas sobre la solidez y el equilibrio de la literatura disponible en este campo, evocando el «efecto de declive» descrito por Schoeler (2011) en la ciencia psicológica. Nuestros hallazgos sugieren que los estudios publicados pueden no representar completamente el panorama completo de la investigación realizada, con un posible sesgo hacia la presentación de resultados positivos o significativos, un fenómeno bien documentado por Ioannidis (2005). También sirven como un recordatorio para interpretar los resultados meta-analíticos con cautela, especialmente en campos donde el sesgo de publicación puede ser prevalente, como lo han destacado Ferguson y Heene (2012) en su crítica de los metanálisis psicológicos.

Limitaciones e implicaciones

A pesar de ofrecer valiosas ideas sobre el ABJ y su impacto en las HPOS matemáticas de los estudiantes, este estudio tiene varias limitaciones. La revisión está limitada por un pequeño número de estudios empíricos—solo se han incluido trece—lo que puede afectar la generalizabilidad de los hallazgos a poblaciones más amplias y diversos contextos educativos. El alcance limitado se debe en parte a las dificultades para acceder a la base de datos Clarivate Web of Science (WoS) y a tesis o disertaciones relevantes. Además, el estudio se ha restringido a artículos en inglés debido a limitaciones de idioma. Asimismo, aunque se han analizado seis variables moderadoras, otros factores influyentes pueden no haber sido considerados en su totalidad, lo que sugiere la necesidad de análisis más completos en futuras investigaciones. Además, este análisis aborda cada moderador de forma individual y no tiene en cuenta las posibles interacciones entre moderadores, que pueden influir colectivamente en la efectividad del ABJ.

Reconocer estas limitaciones es crucial para una comprensión más profunda y para guiar mejoras en futuras investigaciones sobre ABJ y educación matemática. Esto subraya la necesidad de investigaciones más amplias que involucren un mayor número de estudios

empíricos en diversas poblaciones y entornos educativos para proporcionar una comprensión más completa del impacto del ABJ en las HPOS matemáticas. Además, futuros estudios podrían explorar otras posibles variables moderadoras no cubiertas en este metanálisis, como las características específicas de las intervenciones de ABJ o los niveles de conocimiento previo de los estudiantes. Asimismo, la investigación futura podría explorar los efectos combinados de múltiples moderadores a través de enfoques analíticos más complejos, como los efectos de interacción en modelos de meta-regresión. Adoptar este enfoque más amplio contribuirá a una comprensión más integral de la dinámica del ABJ en la educación matemática.

Una limitación notable de este estudio es también el potencial sesgo de publicación, como lo evidencian la reducción significativa en el tamaño del efecto después del ajuste Trim and Fill. Si bien la prueba de Egger no ha indicado sesgo sustancial, los hallazgos del método Trim and Fill sugieren que los estudios faltantes pueden haber sesgado los resultados originales, posiblemente inflando el tamaño del efecto. Esto resalta la necesidad de precaución al interpretar los hallazgos, ya que el conjunto de datos puede no ser tan robusto como se asumió inicialmente. Además, el análisis está limitado por los datos disponibles, con una posible variabilidad en el diseño del estudio y las características de la muestra entre los estudios incluidos que contribuyen al sesgo observado. Confiar únicamente en correcciones estadísticas como Trim and Fill puede no tener en cuenta todas las dimensiones del sesgo de publicación, dejando margen para una exploración adicional. La investigación futura debería priorizar la replicación de estos hallazgos con conjuntos de datos más completos y emplear una gama más amplia de métodos para detectar y corregir el sesgo de publicación. Ampliar el alcance de las bases de datos buscadas, fomentar la preinscripción de estudios y considerar herramientas adicionales como el análisis de p-curve o modelos de selección podría proporcionar una visión más profunda sobre la influencia del sesgo de publicación, mejorando en última instancia la confiabilidad y validez de los resultados meta-analíticos.

Las implicaciones de este metanálisis para la educación matemática son significativas. A pesar del tamaño de efecto relativamente pequeño, el ABJ ha mostrado potencial como una herramienta suplementaria para mejorar las HPOS matemáticas de los estudiantes. Los educadores deberían considerar integrar estrategias de ABJ en sus planes de estudio para estimular el pensamiento matemático de los estudiantes. Sin embargo, es esencial seleccionar y diseñar cuidadosamente las intervenciones de ABJ para alinearse con objetivos educativos específicos, los tipos de habilidades de pensamiento objetivo y el nivel educativo. La investigación futura también debería investigar cómo las características del juego y la interacción entre el conocimiento previo de los estudiantes y el ABJ pueden proporcionar información para personalizar intervenciones que se ajusten mejor a las necesidades de aprendizaje individuales, particularmente en la mejora de sus HPOS matemáticas. Al abordar estos aspectos, los estudios futuros pueden ofrecer una comprensión más completa de cómo se puede utilizar eficazmente el ABJ para apoyar y mejorar la calidad del aprendizaje en el aula de matemáticas.

Conclusiones

Considerando las diversas conclusiones derivadas de estudios individuales que evalúan la efectividad del ABJ en las HPOS matemáticas de los estudiantes, este metanálisis tiene como objetivo ofrecer una síntesis actualizada del impacto del ABJ en las HPOS matemáticas de los estudiantes. Concluye que el ABJ afecta positivamente las HPOS matemáticas de los estudiantes, con un tamaño de efecto global que se encuentra dentro de la categoría de «pequeño». Aunque el tamaño de efecto agregado encontrado es relativamente pequeño, es suficiente para proporcionar información de que el ABJ

puede considerarse como un enfoque alternativo para mejorar las HPOS matemáticas de los estudiantes en diversos niveles educativos. Este estudio también revela que variables como el tipo de habilidades de pensamiento medidas, el nivel educativo, el continente donde se ha realizado el estudio, el tipo de tratamiento de control y la duración de la intervención tienden a influir significativamente en el impacto del ABJ en las HPOS matemáticas de los estudiantes. Sin embargo, el factor de tamaño de muestra no ha mostrado un impacto significativo en la efectividad del ABJ. En general, estos hallazgos indican que, si bien el ABJ puede apoyar el desarrollo de las HPOS matemáticas de los estudiantes, se debe prestar atención a las variables específicas que afectan su efectividad.

Agradecimientos

Los autores agradecen profundamente a la Universitas Islam Negeri Raden Intan Lampung, Indonesia, por su apoyo en la investigación. También extendemos nuestro agradecimiento a la Escuela de Doctorado en Educación de la Universidad de Szeged, Hungría, la Universidad Hasyim Asy'ari, Indonesia, y al Fondo de Dotación de Indonesia para la Educación (Lembaga Pengelola Dana Pendidikan - LPDP) del Ministerio de Indonesia, por su valioso apoyo, sugerencias y discusiones que contribuyeron a esta investigación.

Declaración de contribución de autoría CRediT

Bambang Sri Anggoro: supervisión, obtención de fondos, redacción – revisión y edición.

Andi Harpeni Dewantara: metodología, redacción - borrador original, análisis formal.

Suherman Suherman: conceptualización, redacción - borrador original, análisis formal, metodología, edición y visualización.

Rosida Rakhmawati Muhammad: redacción - borrador original, análisis formal.

Sari Saraswati: redacción – revisión y edición.

Anexo. Material adicional

Se puede consultar material adicional a este artículo en su versión electrónica disponible en <https://doi.org/10.1016/j.psicod.2024.500158>.

Referencias

- Las referencias marcadas con un * fueron incluidas en el metaanálisis.
- Acharya, B. R. (2017). Factors affecting difficulties in learning mathematics by mathematics learners. *International Journal of Elementary Education*, 6(2), 8-8. <https://doi.org/10.11648/j.ijeedu.20170602.11>
- Aizpurua, A., Lizaso, I., y Iturbe, I. (2018). Estrategias de aprendizaje y habilidades de razonamiento de estudiantes universitarios. *Revista de Psicodidáctica*, 23(2), 110–116. <https://doi.org/10.1016/j.psicod.2018.01.001>
- *Al-Absi, M. (2017). The effect of using puzzles and games on students' mathematical thinking at the faculty of educational sciences and arts (UNRWA). *An-Najah University Journal for Research - B (Humanities)*, 31(10), 1867–1888. <https://doi.org/10.35552/0247-031-010-007>
- Angelelli, C. V., Ribeiro, G. M., de, C., Severino, M. R., Johnstone, E., Borzenkova, G., y da Silva, D. C. O. (2023). Developing critical thinking skills through gamification. *Thinking Skills and Creativity*, 49, 101354-101354. <https://doi.org/10.1016/j.tsc.2023.101354>
- Asri, A. S. M., y Jamaludin, K. A. (2022). Potential scratch games in developing students' thinking skills. *Malaysian Journal of Social Sciences and Humanities (MJSSSH)*, 7(12), e002004-e002004. <https://doi.org/10.47405/mjssh.v7i12.2004>
- Bai, S., Hew, K. F., y Huang, B. (2020). Does gamification improve student learning outcome? Evidence from a meta-analysis and synthesis of qualitative data in educational contexts. *Educational Research Review*, 30, 100322-100322. <https://doi.org/10.1016/j.edurev.2020.100322>
- *Bakker, M., van den Heuvel-Panhuizen, M., y Robitzsch, A. (2015). Effects of playing mathematics computer games on primary school students' multiplicative reasoning ability. *Contemporary Educational Psychology*, 40, 55–71. <https://doi.org/10.1016/j.cedpsych.2014.09.001>
- Balduzzi, S., Rücker, G., y Schwarzer, G. (2019). How to perform a meta-analysis with R: A practical tutorial. *Evidence Based Mental Health*, 22(4), 153–160. <https://doi.org/10.1136/ebmental-2019-300117>
- Barz, N., Benick, M., Dörrenbächer-Ulrich, L., y Perels, F. (2024). The effect of digital game-based learning interventions on cognitive, metacognitive, and affective-

- motivational learning outcomes in school: A meta-analysis. *Review of Educational Research*, 94(2), 193–227. <https://doi.org/10.3102/00346543231167795>
- Beath, K. J. (2014). A finite mixture method for outlier detection and robustness in meta-analysis. *Research Synthesis Methods*, 5(4), 285–293. <https://doi.org/10.1002/jrsm.1114>
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., y Rothstein, H. R. (2009). *Introduction to meta-analysis*. John Wiley & Sons.
- Bourke, B. (2019). Using gamification to engage higher-order thinking skills. En M. Khosrow-Pour (Ed.), *Handbook of research on promoting higher-order skills and global competencies in life and work* (pp. 632–652). IGI Global. <https://doi.org/10.4018/978-1-7998-3022-1.ch033>
- Byun, J., y Joung, E. (2018). Digital game-based learning for K–12 mathematics education: A meta-analysis. *School Science and Mathematics*, 118(3–4), 113–126. <https://doi.org/10.1111/ssm.12271>
- Card, N. A. (2011). *Applied meta-analysis for social science research*. Guilford Publications.
- Carter, E. C., Schönbrodt, F. D., Gervais, W. M., y Hilgard, J. (2019). Correcting for bias in psychology: A comparison of meta-analytic methods. *Advances in Methods and Practices in Psychological Science*, 2(2), 115–144. <https://doi.org/10.1177/2515245918847196>
- *Chang, K.-E., Wu, L.-J., Weng, S.-E., y Sung, Y.-T. (2012). Embedding game-based problem-solving phase into problem-posing system for mathematics learning. *Computers & Education*, 58(2), 775–786. <https://doi.org/10.1016/j.compedu.2011.10.002>
- Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Academic press.
- Cooper, H. (2016). *Research synthesis and meta-analysis: A step-by-step approach*. Sage Publication.
- *Demirel, T., y Yilmaz, T. K. (2019). The effects of mind games in math and grammar courses on the achievements and perceived problem-solving skills of secondary school students. *British Journal of Educational Technology*, 50(3), 1482–1494. <https://doi.org/10.1111/bjet.12624>
- Deng, R., Benckendorff, P., y Gannaway, D. (2020). Learner engagement in MOOCs: Scale development and validation. *British Journal of Educational Technology*, 51(1), 245–262. <https://doi.org/10.1111/bjet.12810>
- Duval, S., y Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56(2), 455–463. <https://doi.org/10.1111/j.0006-341X.2000.00455.x>
- Dwan, K., Altman, D. G., Arnaiz, J. A., Bloom, J., Chan, A.-W., Cronin, E., Decullier, E., Easterbrook, P. J., Von Elm, E., y Gamble, C. (2008). Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PloS One*, 3(8), e3081. <https://doi.org/10.1371/journal.pone.0003081>
- Egger, M., Smith, G. D., Schneider, M., y Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *Bmj*, 315(7109), 629–634. <https://doi.org/10.1136/bmj.315.7109.629>
- *Emihovich, B., Roque, N., y Mason, J. (2022). Can video gameplay improve undergraduates' problem-solving skills? *International Journal of Game-Based Learning*, 10(2), 1–18. <https://doi.org/10.4018/IJGBL.2020040102>
- Fanelli, D., Costas, R., y Ioannidis, J. P. A. (2017). Meta-assessment of bias in science. *Proceedings of the National Academy of Sciences*, 114(14), 3714–3719. <https://doi.org/10.1073/pnas.1618569114>
- Ferguson, C. J., y Heene, M. (2012). A vast graveyard of undead theories: Publication bias and psychological science's aversion to the null. *Perspectives on Psychological Science*, 7(6), 555–561. <https://doi.org/10.1177/17456916124590>
- Freitas, S. A. A., Lacerda, A. R. T., Calado, P. M. R. O., Lima, T. S., y Canedo, E. D. (2017). *Gamification in education: A methodology to identify student's profile*, 1–8.
- Ganapathy, M., Mehar Singh, M. K., Kaur, S., y Kit, L. W. (2017). Promoting higher order thinking skills via teaching practices. *3L The Southeast Asian Journal of English Language Studies*, 23(1), 75–85. <https://doi.org/10.1757/3L-2017-2301-06>
- *Harisudin, L., Susanto, S., y Hobri, H. (2019). The development of mathematics learning tools through the bridge games based on lesson study for learning community and its relationship with the higherorder thinking skills in probability theory. *Journal of Physics: Conference Series*, 1211, 012075–012075. <https://doi.org/10.1088/1742-6596/1211/1/012075>
- Harrer, M., Cuijpers, P., Furukawa, T. A., y Ebert, D. D. (2021). *Doing meta-analysis with R: A hands-on guide*. CRC press.
- Hartigan, J. A., y Wong, M. A. (1979). Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series c (Applied Statistics)*, 28(1), 100–108.
- Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, 6(2), 107–128.
- Higgins, J. P., Thompson, S. G., Deeks, J. J., y Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *Bmj*, 327(7414), 557–560. <https://doi.org/10.1136/bmj.327.7414.557>
- Ho, G.-Y., Leonhard, M., Volk, G. F., Foerster, G., Pototschnig, C., Klinge, K., Granitzka, T., Zienau, A.-K., y Schneider-Stickler, B. (2019). Inter-rater reliability of seven neurolaryngologists in laryngeal EMG signal interpretation. *European Archives of Oto-Rhino-Laryngology*, 276(10), 2849–2856. <https://doi.org/10.1007/s00405-019-05553-y>
- Hsu, C.-Y., Liang, J.-C., y Tsai, M.-J. (2020). Probing the structural relationships between teachers' beliefs about game-based teaching and their perceptions of technological pedagogical and content knowledge of games. *Technology, Pedagogy and Education*, 29(3), 297–309. <https://doi.org/10.1080/1475939X.2020.1752296>
- *Indriani, M. N., Isnarto, I., y Mariani, S. (2019). The implementation of PBL (problem based learning) model assisted by monopoly game media in improving critical

- thinking ability and self confidence. *Journal of Primary Education*, 8(2), 200–208. <https://doi.org/10.15294/jpe.v8i2.25991>
- Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Medicine*, 2(8), e124. <https://doi.org/10.1371/journal.pmed.1004085>
- Ioannidis, J. P., y Trikalinos, T. A. (2005). Early extreme contradictory estimates may appear in published research: The Proteus phenomenon in molecular genetics research and randomized trials. *Journal of Clinical Epidemiology*, 58(6), 543–549. <https://doi.org/10.1016/j.jclinepi.2004.10.019>
- Ioannidis, J. P., y Trikalinos, T. A. (2007). The appropriateness of asymmetry tests for publication bias in meta-analyses: A large survey. *Cmaj*, 176(8), 1091–1096. <https://doi.org/10.1503/cmaj.060410>
- Izzati, G. N., Dwijanto, D., y Cahyono, A. N. (2022). The effectivity of problem-based learning assisted by snakes and ladders games on students' divergent thinking ability. *Jurnal Pijar Mipa*, 17(6), 732–736. <https://doi.org/10.29303/jpm.v17i6.4177>
- Jabbar, A. I. A., y Felicia, P. (2015). Gameplay engagement and learning in game-based learning. *Review of Educational Research*, 85(4), 740–779. <https://doi.org/10.3102/0034654315577210>
- *Ke, F. (2019). Mathematical problem solving and learning in an architecture-themed epistemic game. *Educational Technology Research and Development*, 67(5), 1085–1104. <https://doi.org/10.1007/s11423-018-09643-2>
- Korkman, M., Barron-Linnankoski, S., y Lahti-Nuutila, P. (1999). Effects of age and duration of reading instruction on the development of phonological awareness, rapid naming, and verbal memory span. *Developmental Neuropsychology*, 16(3), 415–431. https://doi.org/10.1207/S15326942DN1603_24
- Lee, T. Y., Mauriello, M. L., Ahn, J., y Bederson, B. B. (2014). CTArcade: Computational thinking with games in school age children. *International Journal of Child-Computer Interaction*, 2(1), 26–33. <https://doi.org/10.1016/j.ijcci.2014.06.003>
- Li, X., Dusseldorf, E., Su, X., y Meulman, J. J. (2020). Multiple moderator meta-analysis using the R-package Meta-CART. *Behavior Research Methods*, 52(6), 2657–2673. <https://doi.org/10.3758/s13428-020-01360-0>
- Lipsey, M. W., y Wilson, D. B. (2001). *Practical meta-analysis*. SAGE publications, Inc.
- López-Díaz, M. T., y Peña, M. (2021). Mathematics training in engineering degrees: An intervention from teaching staff to students. *Mathematics*, 9(13), 1475–1475. <https://doi.org/10.3390/math9131475>
- Lovianova, I. V., Yu Kaluhin, R., Kovalenko, D. A., Rovenska, O. G., y Krasnoshchok, A. V. (2022). Development of logical thinking of high school students through a problem-based approach to teaching mathematics. *Journal of Physics: Conference Series*, 2288(1), 012021–012021. <https://doi.org/10.1088/1742-6596/2288/1/012021>
- Lu, Z., Chiu, M. M., Cui, Y., Mao, W., y Lei, H. (2023). Effects of game-based learning on students' computational thinking: A meta-analysis. *Journal of Educational Computing Research*, 61(1), 235–256. <https://doi.org/10.1177/07356331221100740>
- Ma, J., Zhang, Y., Zhu, Z., Zhao, S., y Wang, Q. (2023). Game-based learning for students' computational thinking: A meta-analysis. *Journal of Educational Computing Research*, 61(7), 1430–1463. <https://doi.org/10.1177/07356331231178948>
- Mao, W., Cui, Y., Chiu, M. M., y Lei, H. (2022). Effects of game-based learning on students' critical thinking: A meta-analysis. *Journal of Educational Computing Research*, 59(8), 1682–1708. <https://doi.org/10.1177/07356331211007098>
- Merchant, Z., Goetz, E. T., Cifuentes, L., Keehn-Kennicutt, W., y Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Computers & Education*, 70, 29–40. <https://doi.org/10.1016/j.compedu.2013.07.033>
- Mumcu, H. Y. (2018). Examining mathematics department students' views on the use of mathematics in daily life. *International Online Journal of Education and Teaching (IOJET)*, 5(1), 61–80.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, n71–n71. <https://doi.org/10.1136/bmj.n71>
- Petri, G., y Gresse von Wangenheim, C. (2017). How games for computing education are evaluated? A systematic literature review. *Computers & Education*, 107, 68–90. <https://doi.org/10.1016/j.compedu.2017.01.004>
- Pigott, T. D., y Polanin, J. R. (2020). Methodological guidance paper: High-quality meta-analysis in a systematic review. *Review of Educational Research*, 90(1), 24–46. <https://doi.org/10.3102/0034654319877153>
- Pigott, T. D., Valentine, J. C., Polanin, J. R., Williams, R. T., y Canada, D. D. (2013). Outcome-Reporting Bias in Education Research. *Educational Researcher*, 42(8), 424–432. <https://doi.org/10.3102/0013189X13507104>
- *Pires, F., Maquine Lima, F. M., Melo, R., Serique Bernardo, J. R., y de Freitas, R. (2019). *Gamification and engagement: Development of computational thinking and the implications in mathematical learning*, 362–366. <https://doi.org/10.1109/ICALT.2019.00112>
- Plass, J. L., Homer, B. D., y Kinzer, C. K. (2015). Foundations of game-based learning. *Educational Psychologist*, 50(4), 258–283. <https://doi.org/10.1080/00461520.2015.1122533>
- Posit Team. (2023). *RStudio: Integrated development environment for R*.(2023.3.1.446).
- Prieto-Rodríguez, E., Sincock, K., y Blackmore, K. (2020). STEM initiatives matter: Results from a systematic review of secondary school interventions for girls. *International Journal of Science Education*, 42(7), 1144–1161. <https://doi.org/10.1080/09500693.2020.1749909>
- Qian, M., y Clark, K. R. (2016). Game-based learning and 21st century skills: A review of recent research. *Computers in Human Behavior*, 63, 50–58. <https://doi.org/10.1016/j.chb.2016.05.023>
- Renkewitz, F., y Keiner, M. (2019). How to detect publication bias in psychological research: A comparative evaluation of six statistical methods. *Zeitschrift Für Psychologie*, 227(4), 261–279. <https://doi.org/10.1027/2151-2604/a000386>
- Riahi, S. M., y Mokhayeri, Y. (2017). Methodological issues in a meta-analysis. *Current Medical Research and Opinion*, 33(10), 1813–1813. <https://doi.org/10.1080/03007995.2017.1359152>
- Rothstein, H. R. (2005). *Publication bias in meta-analysis*. Publication Bias in Meta-Analysis-Prevention, Assessment and Adjustment/John Wiley y Sons. <https://doi.org/10.1002/0470870168>
- *Sala, G., y Gobet, F. (2017). Does chess instruction improve mathematical problem-solving ability? Two experimental studies with an active control group. *Learning and Behavior*, 45(4), 414–421. <https://doi.org/10.3758/s13420-017-0280-3>
- Scherer, R., y Beckmann, J. F. (2014). The acquisition of problem solving competence: Evidence from 41 countries that math and science education matters. *Large-Scale Assessments in Education*, 2(1), 10. <https://doi.org/10.1186/s40536-014-0010-7>
- Schooler, J. (2011). Unpublished results hide the decline effect. *Nature*, 470(7335), 437–437.
- Shu, L., y Liu, M. (2019). Student engagement in game-based learning: A literature review from 2008 to 2018. *Journal of Educational Multimedia and Hypermedia*, 28(2), 193–215.
- Slavin, R., y Smith, D. (2009). The relationship between sample sizes and effect sizes in systematic reviews in education. *Educational Evaluation and Policy Analysis*, 31(4), 500–506. <https://doi.org/10.3102/1062373709352369>
- Smith, K., Maynard, N., Berry, A., Stephenson, T., Spiteri, T., Corrigan, D., Mansfield, J., Ellerton, P., y Smith, T. (2022). Principles of problem-based learning (PBL) in STEM education: Using expert wisdom and research to frame educational practice. *Education Sciences*, 12(10), 728. <https://doi.org/10.3390/eduscii12100728>
- Sterne, J. A., Sutton, A. J., Ioannidis, J. P., Terrin, N., Jones, D. R., Lau, J., Carpenter, J., Rücker, G., Harbord, R. M., y Schmid, C. H. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *Bmj*, 343. <https://doi.org/10.1136/bmj.d4002>
- Suherman, S., Prananda, M. R., Probongrum, D. I., Pratama, E. R., Laksono, P., y Amiruddin, A. (2020). Improving higher order thinking skills (hots) with project based learning (pjbl) model assisted by geogebra. *Journal of Physics: Conference Series*, 1467(1), 012027. <https://doi.org/10.1088/1742-6596/1467/1/012027>
- Suherman, S., y Vidákovich, T. (2022). Assessment of mathematical creative thinking: A systematic review. *Thinking Skills and Creativity*, 44, 101019. <https://doi.org/10.1016/j.tsc.2022.101019>
- Suherman, S., y Vidákovich, T. (2024). Relationship between ethnic identity, attitude, and mathematical creative thinking among secondary school students. *Thinking Skills and Creativity*, 51, 101448. <https://doi.org/10.1016/j.tsc.2023.101448>
- Sun, L., Guo, Z., y Hu, L. (2023). Educational games promote the development of students' computational thinking: A meta-analytic review. *Interactive Learning Environments*, 31(6), 3476–3490. <https://doi.org/10.1080/10494820.2021.1931891>
- Sutton, A. J., Duval, S. J., Tweedie, R. L., Abrams, K. R., y Jones, D. R. (2000). Empirical assessment of effect of publication bias on meta-analyses. *Bmj*, 320(7249), 1574–1577. <https://doi.org/10.1136/bmj.320.7249.1574>
- *Taja-on, E. P. (2019). Game-aided instruction: Enhancing critical thinking through logical-mathematical games. *Online Submission*, 1(1), 1–15. <https://doi.org/10.13140/RG.2.2.16436.17288>
- Tokac, U., Novak, E., y Thompson, C. G. (2019). Effects of game-based learning on students' mathematics achievement: A meta-analysis. *Journal of Computer Assisted Learning*, 35(3), 407–420. <https://doi.org/10.1111/jcal.12347>
- Turgut, S., y Temur, Ö. D. (2017). The effect of game-assisted mathematics education on academic achievement in Turkey: A meta-analysis study. *International Electronic Journal of Elementary Education*, 10(2), 195–206. <https://doi.org/10.26822/iejee.2017236115>
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3). <https://doi.org/10.18637/jss.v036.i03>
- Viechtbauer, W., y Cheung, M. W.-L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, 1(2), 112–125. <https://doi.org/10.1002/rsmi.11>
- Wang, L.-H., Chen, B., Hwang, G.-J., Guan, J.-Q., y Wang, Y.-Q. (2022). Effects of digital game-based STEM education on students' learning achievement: A meta-analysis. *International Journal of STEM Education*, 9(1), 1–13. <https://doi.org/10.1186/s40594-022-00344-0>
- White, K., y McCoy, L. P. (2019). Effects of game-based learning on attitude and achievement in elementary mathematics. *Networks: An Online Journal for Teacher Research*, 21(1), 1–17. <https://doi.org/10.4148/2470-6353.1259>
- Wibawa, K. A., Payadnya, P. A. A., Yasa, G. U., y Prahmana, R. C. I. (2022). The learning trajectory of entrepreneurship arithmetic content using a traditional market. *Mathematics Teaching Research Journal*, 14(3), 144–169.
- *Yeh, C. Y. C., Cheng, H. N. H., Chen, Z.-H., Liao, C. C. Y., y Chan, T.-W. (2019). Enhancing achievement and interest in mathematics learning through Math-Island. *Research and Practice in Technology Enhanced Learning*, 14(1), 5–5. <https://doi.org/10.1186/s41039-019-0100-9>
- Yerizon, Y., Triwani, T., y Musdi, E. (2022). Effectiveness of mathematics learning devices based on flipped classroom to improve mathematical critical thinking ability students. *International Journal of Education and Management Engineering*, 12(3), 41–46. <https://doi.org/10.5815/ijeme.2022.03.05>