


Evaluación de las funciones ejecutivas integradas en plataformas de realidad virtual para población infantil: revisión sistemática

Evaluation of executive functions integrated into virtual reality platforms for child population: systematic review

Eduardo Duarte-González,¹ Bernardo Pino-Belaunzarán,¹ Marisol Castañeda-Franco^{2*} 

https://doi.org/10.36105/psic_anah.2026v1n2.03

¹ Facultad de Psicología, Universidad Anáhuac México

² Instituto Nacional de Psiquiatría Ramón de la Fuente Muñiz, Ciudad de México

*Autor de correspondencia: Marisol Castañeda-Franco, castaneda@inprf.gob.mx, Calzada México-Xochimilco No. 101, Col. San Lorenzo Huipulco, C.P. 14370, Tlalpan, Ciudad de México

Fecha de recepción: 22 de enero de 2025
Fecha de aceptación: 24 de febrero de 2026

CÓMO CITAR: Duarte-González, E., Pino-Belaunzarán, B. & Castañeda-Franco, M. (2026). Evaluación de las funciones ejecutivas integradas en plataformas de realidad virtual para población infantil: revisión sistemática. *Investigación y Avances en Psicología*, 1 (2),53-73. https://doi.org/10.36105/psic_anah.2026v1n2.03




Resumen

Las funciones ejecutivas (FE) son procesos cognitivos superiores esenciales para la conducta dirigida a metas, clasificados en componentes “fríos” (cognitivos) y “cálidos” (emocionales). Pese a su relevancia en la maduración cerebral infantil, las pruebas neuropsicológicas tradicionales presentan limitaciones en su validez ecológica. Ante esto, la realidad virtual (RV) surge como una herramienta coadyuvante capaz de simular escenarios de alta fidelidad. El objetivo fue realizar una revisión sistemática durante el período 2021-2026 para identificar plataformas de realidad virtual (RV) diseñadas para evaluar las FE en población infantil. Siguiendo la declaración PRISMA, se consultaron PubMed y PsycInfo en dos fases (2024 y 2026) usando términos booleanos específicos. Se incluyeron artículos originales de casos y controles, publicados en inglés o español y que emplearan la RV en menores. Se identificaron plataformas de RV que miden componentes ejecutivos en entornos inmersivos. Los hallazgos describen propiedades psicométricas y la capacidad para evaluar procesos ejecutivos fríos y cálidos, aunque todavía persisten limitaciones metodológicas. La RV representa un avance fundamental en la evaluación neuropsicológica infantil al ofrecer mediciones precisas en entornos ecológicos. Sin embargo, su consolidación futura requiere protocolos estandarizados y accesibles que faciliten la detección de disfunciones ejecutivas.

Palabras clave: evaluación neuropsicológica, realidad virtual, funciones ejecutivas, infancia, cognición.

Abstract

Executive functions (EF) are higher-order cognitive processes essential for goal-directed behavior, classified into “cold” (cognitive) and “hot” (emotional) components. Despite their relevance to child brain maturation, traditional neuropsychological tests present limitations in ecological validity. In response, virtual reality (VR) emerges as a supportive tool capable of simulating high-fidelity scenarios. The objective was conducting a systematic review for



the 2021–2026 period to identify virtual reality (VR) platforms designed to assess EF in the pediatric population. Following the PRISMA statement, PubMed and PsycInfo were consulted in two phases (2024 and 2026) using specific boolean terms. Original case-control articles published in English or Spanish that employed VR in minors were included. VR platforms measuring executive components in immersive environments were identified. The findings describe psychometric properties and the capacity to evaluate both cold and hot executive processes, although methodological limitations persist. VR represents a fundamental advancement in pediatric neuropsychological assessment by offering precise measurements in ecological settings. However, its future consolidation requires standardized and accessible protocols that facilitate the detection of executive dysfunctions.

Keywords: neuropsychological assessment, virtual reality, executive functions, childhood, cognition.

Introducción

Las funciones ejecutivas (FE) se definen como procesos cognitivos de orden superior fundamentales para la orquestación de conductas dirigidas a metas u objetivos específicos. De acuerdo con Diamon (2012), estos procesos surgen de la integración sistémica de diversos procesos cognitivos que abarcan desde la recepción sensorial y la ejecución motora hasta la actualización de la memoria y la valoración emocional. Esta interacción multidimensional permite transformar la información procesada en acciones deliberadas y adaptativas. Bajo la perspectiva de Ward (2020), las FE actúan como un sistema de supervisión metacognitiva de dominio general, funcionando como un mecanismo de gestión global que regula diversas funciones cerebrales. Existe un consenso científico sólido que identifica tres componentes esenciales dentro de las FE: el control inhibitorio, la memoria de trabajo y la flexibilidad cognitiva (Miyake *et al.*, 2000).

Estudios de neuroimagen han revelado principios organizativos clave de las FE. Stuss y Alexander (2000) destacan una especialización

hemisférica donde la corteza prefrontal (CPF) lateral izquierda se vincula con la preparación de tareas mientras que la derecha se enfoca en el monitoreo del desempeño. Badre y D'Esposito (2009) proponen una jerarquía anteroposterior de la CPF lateral, distinguiendo una ruta dorsal dedicada a la planificación de la acción y una ruta ventral asociada al procesamiento lingüístico y de objetos. Un principio emergente y ampliamente aceptado clasifica las FE en "frías" (procesos puramente cognitivos) y "cálidas" (procesos vinculados a la emoción y recompensa), lo que permite integrar el afecto en el control cognitivo dentro de un espectro influenciado por el contexto (Ward, 2020). Si bien las FE frías se gestionan mayoritariamente en la CPF dorsolateral y las cálidas en la CPF orbito-medial, este esquema se extiende a una red compleja que involucra áreas corticales, subcorticales y al cíngulo anterior (Salehinejad et al., 2021).

La evaluación de las FE durante la infancia y adolescencia es crítica, pues permite mapear la gestión de la información lógica frente a la reactividad emocional. Kusi-Mensah et al. (2021) subrayan que, en la adolescencia, este análisis es vital para comprender la impulsividad derivada del desfase madurativo entre el sistema límbico y la CPF. No obstante, aunque pruebas tradicionales como el Test de Wisconsin, el Stroop o la Torre de Londres siguen siendo estándares clínicos, su limitada validez ecológica ha generado críticas sobre su capacidad para predecir el comportamiento en la vida diaria. Como respuesta, la neuropsicología contemporánea ha integrado la Realidad Virtual (RV), un sistema computacional que genera entornos tridimensionales inmersivos donde los usuarios interactúan con estímulos artificiales de forma natural (Reyes & López, 2023). La principal ventaja de usar plataformas de RV frente a las pruebas tradicionales de lápiz-papel para evaluar las FE es su validez ecológica ya que al simular escenarios reales permite evaluar cómo el niño o adolescente organiza su conducta, inhibe distracciones y toma decisiones en contextos con una demanda cognitiva similar a la real. A pesar de sus beneficios, la revisión sistemática de Reyes y López (2023) reveló que, entre 2017 y 2021, el uso de RV se concentró en adultos a través de herramientas como la *Virtual Cooking Task* (Giglioli et al., 2021), el sistema *Vrai*

(Robitaille et al., 2016) o el entorno *VIS* (Nir-Hadad et al., 2015). En contraste, la evidencia en población infanto-juvenil resultó limitada, reportándose mayoritariamente el sistema del *Virtual Reality Medical Center* (Fang et al., 2019) enfocado solo en la evaluación de procesos atencionales.

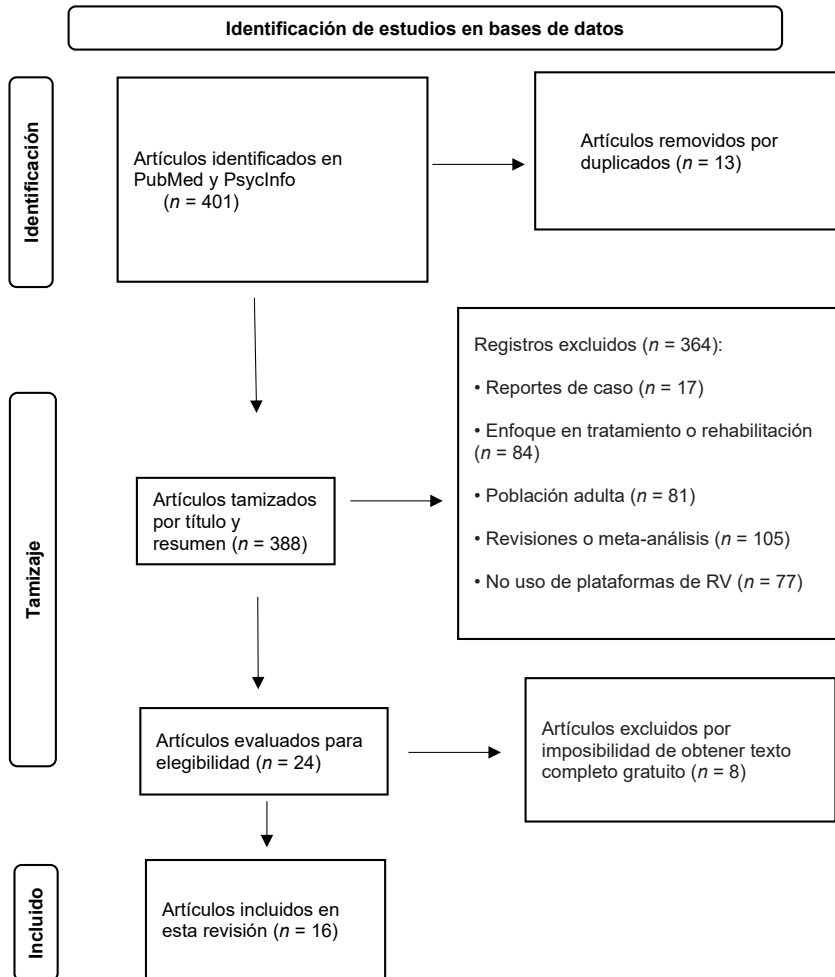
Considerando este vacío de información, el presente artículo tiene como objetivo realizar una revisión sistemática actualizada en el período del 2021 al 2026 para identificar y sintetizar las plataformas de RV diseñadas específicamente para la evaluación de las FE en niños y adolescentes. De manera específica, se busca: 1) identificar las plataformas de RV empleadas desarrolladas en este último quinquenio; 2) determinar los componentes ejecutivos evaluados (control inhibitorio, memoria de trabajo o flexibilidad cognitiva); 3) analizar la inclusión de FE "frías" o "cálidas" en entornos relevantes (aulas, parques, contextos sociales); y 4) sintetizar las limitaciones actuales para proponer futuras líneas de investigación. Esta revisión se justifica teóricamente por la necesidad de cubrir la carencia de estudios en menores, proporcionando una base sólida sobre cómo la RV puede aportar información valiosa en la evaluación neuropsicológica de las etapas más tempranas del desarrollo.

Método

La identificación, cribado y selección de los artículos incluidos en este trabajo se llevó a cabo siguiendo los lineamientos establecidos en la declaración PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). El flujo detallado de este proceso, desde la búsqueda inicial hasta la muestra final, se presenta en forma esquemática en la Figura 1.

Figura 1

Diagrama de flujo PRISMA de los artículos identificados, seleccionados e incluidos en esta revisión sistemática



Fuente: elaboración propia.

Estrategia de búsqueda

El proceso de recolección de información se estructuró en dos fases cronológicas: la primera búsqueda se llevó a cabo entre mayo-agosto

de 2024 por BP y ED, mientras que la actualización y segunda fase de búsqueda fue realizada por MC durante el primer bimestre (enero-febrero) de 2026. Ambas exploraciones se llevaron a cabo en las bases de datos de PubMed y PsycInfo, utilizando la siguiente estrategia de búsqueda booleana: 'virtual reality' AND 'assessment' AND 'executive functions' AND 'children'.

Para la selección de la muestra, se aplicaron los siguientes criterios de inclusión: 1) artículos originales en revistas con revisión por pares, 2) publicaciones en idiomas inglés o español, 3) diseños de investigación de casos y controles, 4) empleo de plataformas de realidad virtual (RV) como herramienta central, 5) aplicación de métodos estadísticos orientados a la comparación entre grupos, 6) estudios publicados dentro del período 2021-2026. Por el contrario, se establecieron como criterios de exclusión los artículos duplicados, reportes de caso, estudios de intervención o rehabilitación, investigaciones en población adulta, revisiones sistemáticas y/o meta-análisis.

La búsqueda bibliográfica arrojó 401 registros. Tras eliminar 13 duplicados, se tamizaron 388 artículos mediante la lectura del título y resumen. De estos, se excluyeron 364 artículos por los siguientes criterios: reportes de caso (17), enfoque en tratamiento o rehabilitación (84), población adulta (81), revisiones o meta-análisis (105) y no uso de plataformas de RV (77). Finalmente, se identificaron 24 artículos de los cuales se excluyeron 8 al no poder obtenerse el texto completo de forma gratuita. Consecuentemente, 16 artículos fueron incluidos en la revisión sistemática.

Evaluación de la calidad de los estudios

La evaluación de la calidad metodológica de los artículos incluidos fue realizada de forma independiente por los investigadores BP, ED y MC, empleando para ello la herramienta Quality Assessment of Case-Control Studies del National Heart, Lung, and Blood Institute (NHLBI). Este instrumento consta de 12 reactivos diseñados para analizar la robustez del diseño y minimizar el riesgo de sesgo a través de dominios críticos tales como: claridad del objetivo de investigación,

representatividad de la población y muestra, justificación estadística del tamaño muestral, criterios de elegibilidad (inclusión/exclusión), precisión en la definición de casos y controles, aleatoriedad en la selección, validez del grupo control, niveles de exposición y empleo de evaluadores cegados. La valoración de cada criterio se registró mediante las opciones de respuesta: 'SÍ', 'NO' o 'No reportado/No determinable'. Finalmente, se calculó el porcentaje de respuestas afirmativas en cada reactivo para cada artículo, utilizándolo como indicador cuantitativo de la calidad metodológica general de la evidencia analizada.

Resultados

La Tabla 1 presenta los 16 artículos incluidos en esta revisión sistemática. Con respecto a las plataformas de RV empleadas para evaluar FE en población infantil se encontró que, en el último quinquenio, la evaluación de las FE con RV en población infantil ha transitado hacia entornos de alta validez ecológica, utilizando plataformas que van desde sistemas de proyección envolvente hasta dispositivos móviles y visores de inmersión total.

Las plataformas de RV con sistemas de inmersión por proyección y visores (HMD) son: a) CAVE (Bulgarelli *et al.*, 2023) es un sistema de visualización de cuatro lados con visión estéreo-activa y seguimiento de movimiento de seis grados de libertad. Mediante un aula virtual, los niños interactúan con avatares usando una "varita mágica". Evalúa la interacción social y el control motor, permitiendo medir la conectividad funcional (fNIRS) durante la elección y juego con avatares preferidos o asignados, y b) VR-CAT (Shen *et al.*, 2022) usa un visor de inmersión total operado por computadora bajo una narrativa de gamificación. Evalúa tres dominios específicos: control inhibitorio (gestión de centinelas), memoria de trabajo (replicación de secuencias criptográficas) y flexibilidad cognitiva (emparejamiento de patrones dinámicos).

Para las plataformas con entornos ecológicos de la vida diaria se identifican los siguientes: a) SmartAction-VR (Romero *et al.*, 2024): basada en el paradigma de múltiples recados, traslada al niño a cinco escenarios (cocina, dormitorio, calle, etcétera) el cual evalúa la planificación

y ejecución mediante el análisis detallado de errores de omisión (acciones faltantes) y comisión (perseveraciones, sustituciones e inclusiones innecesarias), b) EPELI (Seesjärvi *et al.*, 2022): recrea un departamento completo donde el niño debe cumplir 70 sub-tareas distribuidas en 13 escenarios cotidianos; es una herramienta robusta para evaluar la planificación y la memoria prospectiva, ya que incluye tareas basadas en el tiempo y en eventos (señales sonoras) bajo la presión de distractores auditivos y visuales, y c) AttnKare-D (Oh *et al.*, 2024): espacio virtual guiado por un avatar («fantasma») donde se realizan tareas de organización, cálculo y secuenciación lógica (ej. empacar mochilas o planificar horarios) que mide la eficiencia del desempeño a través de la precisión y el tiempo, detectando conductas disruptivas.

Las plataformas de RV con escenarios escolares y paradigmas de rendimiento continuo se identificaron los siguientes: a) AULA (Fernández-Martin *et al.*, 2023): emplea un entorno escolar virtual para aplicar paradigmas Go/No-Go y proporciona medidas clásicas de atención sostenida e inhibición (omisiones, comisiones, tiempo de reacción) e innova al integrar sensores de movimiento en las gafas para registrar la desviación del foco atencional ante distractores realistas (compañeros, ruidos de tráfico, etcétera), y b) VR Classroom (Ide-Okochi *et al.*, 2023): aplicación experimental que sitúa al participante en un aula para evaluar el seguimiento ocular (fijación de la mirada) y la percepción interoceptiva (precisión en el conteo de latidos cardiacos) ante estímulos de perturbación ambiental.

Finalmente, las plataformas de RV con simuladores específicos y tareas de navegación incluyen: a) EcoSupermarketX (Mouga *et al.*, 2021): tarea no inmersiva de un supermercado virtual que evalúa la memoria de trabajo y la carga ejecutiva mediante listas de compras de dificultad creciente, integrando además variables de cognición social mediante señalizaciones o pistas proporcionadas por avatares y b) MVR (Realidad Virtual Móvil) (Tabibi *et al.*, 2023): replica un entorno peatonal cercano a una escuela y evalúa la toma de decisiones y el control inhibitorio en tiempo real, donde el niño debe decidir el momento seguro para cruzar una calle con tráfico bidireccional desde una perspectiva de primera persona.

Los componentes ejecutivos evaluados en las plataformas de RV analizadas en esta revisión se centran en diversos dominios de FE. Instrumentos como AttnKare-D (Oh *et al.*, 2024), VR-CAT (Shen *et al.*, 2022) y el aula simulada en 3D (Stokes *et al.*, 2022) ofrecen una evaluación integral que abarca el control inhibitorio, la memoria de trabajo y la flexibilidad cognitiva. De manera similar, el entorno CAVE (Bulgarelli *et al.*, 2023), el VR Classroom (Ide-Okochi *et al.*, 2023) y el VHLM (Castilla *et al.*, 2021) integran la flexibilidad y el control inhibitorio, sumando variables como la coordinación visomotora y la planificación. Por su parte, el sistema VADDAI (Bernardelli *et al.*, 2025) y el aula virtual de Verhoef *et al.* (2021) refuerzan la evaluación de la memoria de trabajo y la flexibilidad junto a procesos de autorregulación y monitoreo. Finalmente, otras herramientas se especializan en componentes específicos, como el SmartAction-VR (Romero *et al.*, 2024), el EcoSupermarketX (Mouga *et al.*, 2021) y el EPELI (Seesjärvi *et al.*, 2022), los cuales priorizan la planificación y el monitoreo, mientras que el AULA (Fernández-Martin *et al.*, 2023) y el MVR (Tabibi *et al.*, 2023) se enfocan en la autorregulación, la atención y el control inhibitorio aplicados a contextos funcionales.

La inclusión de FE “frías” o “cálidas” en entornos relevantes (aulas, parques, contextos sociales) muestran un predominio en la evaluación de FE “frías”. El entorno escolar es uno de los escenarios más utilizados debido a su relevancia en la vida diaria infantil; en este contexto, herramientas como AULA (Fernández-Martin *et al.*, 2023), VR Classroom (Ide-Okochi *et al.*, 2023), CAVE (Bulgarelli *et al.*, 2023) y el aula simulada en 3D (Stokes *et al.*, 2022) evalúan procesos atencionales y de control ejecutivo. Otros estudios trasladan estas mediciones de FE frías al ámbito doméstico o urbano, como es el caso de EPELI (Seesjärvi *et al.*, 2022), VADDAI (Bernardelli *et al.*, 2025) y SmartAction-VR (Romero *et al.*, 2024) en el hogar, o el sistema MVR (Tabibi *et al.*, 2023) en entornos ciudadanos. Finalmente, propuestas como el VR-CAT (Shen *et al.*, 2022) y el VHLM (Castilla *et al.*, 2021) utilizan escenarios lúdicos o seminaturalistas, como castillos y laberintos, para analizar componentes fríos en entornos motivadores para el usuario.

Por otro lado, la inclusión de FE "cálidas" que involucran regulación emocional y toma de decisiones con carga motivacional es menos frecuente pero significativa en entornos sociales; destaca el artículo de Verhoef *et al.* (2021) en un aula virtual diseñada para la interacción social, y EcoSupermarketX (Mouga *et al.*, 2021), que logra evaluar tanto FE frías como cálidas dentro de un entorno de supermercado.

A pesar de los avances en la evaluación de las FE mediante RV, se identifican limitaciones transversales que condicionan la generalización de los resultados y definen las metas para futuras investigaciones. En primer lugar, es necesario comentar sobre la validez metodológica y representatividad de la muestra que es una limitación persistente en casi la totalidad de los estudios analizados incluyendo a Bulgarelli *et al.* (2023), Oh *et al.* (2024), Shen *et al.* (2022), Bernardelli *et al.* (2025) y Stokes *et al.* (2022) debido al reducido tamaño de las muestras y son artículos de carácter piloto o exploratorio. Esto compromete la potencia estadística, impide el análisis de subgrupos (como los subtipos de TDAH o la influencia de la medicación) y genera sesgos de representatividad. Autores como Fernández-Martin *et al.* (2023) y Seesjärvi *et al.* (2022) señalan además el uso de muestreos no probabilísticos y sesgos subjetivos en las escalas de comparación (como el BRIEF), lo que dificulta una validación externa robusta.

En segundo lugar, la complejidad de las tareas y aislamiento de procesos que se observa en plataformas como el EcoSupermarketX (Mouga *et al.*, 2021) y VADDAI (Bernardelli *et al.*, 2025) subrayan la dificultad de aislar qué FE específica está alterada debido a la naturaleza multidimensional de las tareas de RV. Asimismo, se reconoce que el desempeño en entornos simulados puede diferir de la conducta espontánea en la vida real, ya que es complejo replicar la incertidumbre total de la cotidianidad o la influencia de la motivación por el carácter lúdico de la tecnología.

En tercer lugar, existen limitaciones de diseño y estímulos ya que se reporta una falta de diversidad en los escenarios y estímulos. Verhoef *et al.* (2021) y Bulgarelli *et al.* (2023) mencionan que el uso de pocos escenarios o avatares con opciones de personalización limitadas restringe la validez de los contextos sociales y de provocación.

Por su parte, Tabibi *et al.* (2023) destacan la ausencia de estímulos multimodales (como el audio del tráfico), señalando que las futuras plataformas deben integrar la audición para una evaluación ecológica completa.

Por último, es importante considerar los desafíos tecnológicos y adaptabilidad ya que existe una brecha en la consideración de la destreza previa de los participantes con dispositivos de RV y videojuegos (Oh *et al.*, 2024; Seesjärvi *et al.*, 2022). Además, investigaciones como la de Castilla *et al.* (2021) sugieren la necesidad de incorporar hardware más preciso, como dispositivos de *eye-tracking* y seguimiento de la dirección de la cabeza, para discriminar mejor entre problemas de planificación y déficits atencionales.

Con respecto a la calidad metodológica de los 16 artículos incluidos, ésta se evaluó mediante la herramienta *Quality Assessment of Case-Control Studies* del National Heart, Lung, and Blood Institute (NHLBI). En términos generales, los estudios demostraron una alta consistencia en la claridad de la pregunta de investigación y en la definición de los criterios de elegibilidad, cumpliendo con estos parámetros en la totalidad de la muestra. No obstante, se identificaron debilidades críticas en dimensiones específicas: la mayoría de las investigaciones no presentaron una justificación del tamaño muestral ni aplicaron mecanismos de aleatorización en la selección de participantes. A pesar de estas carencias en el reporte detallado de métricas de precisión y validez, la estructura general de los diseños de casos y controles permitió categorizar la evidencia disponible como aceptable para los objetivos de esta revisión sistemática.

Tabla 1

Evidencia empírica de plataformas de realidad virtual para la evaluación de funciones ejecutivas en población infantil

| Autores (Año) | Muestra y objetivos | Intrumentos y plataforma de RV | Resultados |
|----------------------------------|--|---|--|
| Castilla <i>et al.</i> (2021) | 10 niños (6-16 años). Probar el VHLM para estudiar el control inhibitorio y flexibilidad mental. | VHLM. | Identificación de estrategias de replanificación y latencia como indicadores de impulsividad y mapas mentales ($p < .05$). |
| Stokes <i>et al.</i> (2022) | 20 niños con TDAH (8-12 años). Evaluar la dinámica temporal de la distracción mediante seguimiento ocular. | Aula simulada en 3D, AX-XPT, Stroop, Math Task. | La presencia de distractores redujo significativamente el rendimiento ($p < 0.005$) e interrumpió la fijación visual. |
| Bernardelli <i>et al.</i> (2025) | 9 niños con TDAH y 10 control (7-11 años). 1. Evaluar usabilidad y aceptación de entorno de RV diseñado a medida. | VADDAI | Alta satisfacción y usabilidad en ambos grupos. El grupo TDAH mostró mayor apreciación por la utilidad de la herramienta. |
| Verhoef <i>et al.</i> (2021) | 32 niños (8-13 años). Evaluar psicometría de una plataforma de interacción social. | Aula virtual interactiva. | Buena validez convergente ($r = .73$) y discriminante ($d = 0.29$). Los niños mostraron mayor entusiasmo con RV que con papel. |
| Salmi <i>et al.</i> (2024) | 72 niños con TDAH y 71 control (9-13 años). Examinar variabilidad de latencias en entornos de interacción libre. | EPELI. | El grupo TDAH mostró latencias de respuesta más cortas ($p < .001$) y menor variabilidad que los controles en tareas cotidianas. |

| Autores (Año) | Muestra y objetivos | Instrumentos y plataforma de RV | Resultados |
|---------------------------------------|--|---|---|
| Ide-Okochi <i>et al.</i> (2023) | 7 niños con discapacidad de desarrollo y 7 control (10-18 años). Explorar mirada e interocepción. | VR Classroom; MAIA (Cuestionario interoceptivo). | Niños con discapacidad mostraron mayor duración de mirada al profesor ($p < 0.05$) y menor percepción de frecuencia cardíaca. |
| Fernández-Martín <i>et al.</i> (2023) | 57 niños con TDAH y 53 controles (9 años). Identificar perfiles conductuales ecológicos. | AULA; SDQ (Strengths and Difficulties Questionnaire). | No se replicaron subtipos tradicionales de TDAH; se hallaron perfiles basados en distracción y actividad motora cefálica. |
| Bulgarelli <i>et al.</i> (2023) | 37 niños (3-5 años). Evaluar desarrollo social mediante RV y fNIRS. | CAVE System; Espectroscopía funcional. | Preferencia por jugar con avatares de igual género/edad (80% en niñas). Patrones específicos de conectividad fronto-temporal. |
| Oh <i>et al.</i> (2024) | 15 niños con TDAH y 5 control (6-12 años). Explorar utilidad clínica y seguridad. | AttnKare-D; K-WISC; Conners CPT. | Excelente rendimiento diagnóstico ($AUC = 0.893$). Alta correlación con escalas de padres K-ARS ($r = 0.62$). |
| Mouga <i>et al.</i> (2021) | 17 TEA y 16 controles (15-16 años). Evaluar vínculo entre FE y cognición social. | EcoSupermarketX; Eye-tracking; Escalas Wechsler. | TEA mostró dependencia de claves contextuales ($p = 0.009$) y mayor fijación visual en claves sociales salientes. |
| Romero <i>et al.</i> (2024) | 40 TDAH y 36 controles (9-17 años). Evaluar funcionamiento ejecutivo en vida diaria. | SmartAction-VR; NEPSY-II; Stroop; Digit Span. | TDAH con menor precisión y más errores de comisión y olvidos ($p < 0.01$). Olvidos correlacionaron con baja independencia real. |
| Seesjärvi <i>et al.</i> (2025) | 71 TDAH y 81 controles (9-12 años). Evaluar memoria prospectiva basada en el tiempo. | EPELI. | TDAH mostró menor monitoreo estratégico del tiempo (mediación completa del efecto del trastorno en el desempeño). |

| Autores (Año) | Muestra y objetivos | Instrumentos y plataforma de RV | Resultados |
|--------------------------------|--|---|--|
| Seesjärvi <i>et al.</i> (2021) | 47 TDAH y 68 controles (9-12 años). Examinar validez predictiva, discriminante y concurrente. | EPELI; BRIEF; ADHD-RS. | Alta validez discriminante ($AUC=.88$). Correlaciones significativas con BRIEF y ADHD-RS ($r = .31$ a $.57$). |
| Shen <i>et al.</i> (2022) | 24 TCE y 30 con lesión ortopédica (7-17 años). Evaluar utilidad y validez de VR-CAT para TCE. | VR-CAT; SSQ; NIH Toolbox; Escala de Borg. | Confiabilidad test-retest moderada ($ICC = 0.48$). Validez concurrente con NIH Toolbox ($r=0.53$). Alta validez ecológica y niveles de diversión. |
| Seesjärvi <i>et al.</i> (2022) | 100 niños con desarrollo típico (9-12 años). Evaluar consistencia interna y efectos de edad/género en EPELI. | EPELI; BRIEF; Presence Questionnaire; SSQ. | Consistencia interna aceptable ($\alpha=0.70-0.88$). Correlación negativa con BRIEF ($r = -.33$). La eficacia de la tarea explicó el 15.6% de la varianza. |
| Tabibi <i>et al.</i> (2023). | 50 niños (9 años). 56% con TDAH (sin medicación durante la prueba). Evaluar diferencias en habilidades peatonales y FE. | MVR (Simulación móvil); IVA+Plus; BDEFS-CA. | Niños con TDAH realizaron más cruces inseguros ($p < .01$). El riesgo peatonal se vinculó directamente a déficits ejecutivos ($r = .54$). |

Nota. TDAH=Trastorno por Déficit de Atención e Hiperactividad; TEA=Trastorno del Espectro Autista; TCE=Traumatismo Craneoencefálico; FE=Funciones Ejecutivas; RV=Realidad Virtual; SSQ=Simulator Sickness Questionnaire; AUC=Área bajo la curva; ICC=Coefficiente de correlación intraclase.

Fuente: elaboración propia.

Discusión y conclusiones

El objetivo de la presente revisión sistemática fue identificar y sintetizar las plataformas de RV desarrolladas entre 2021 y 2026 para la evaluación de las FE en población infanto-juvenil. Los hallazgos revelan un cambio de paradigma significativo ya que mientras que el periodo

2017-2021 mostraba un vacío de herramientas específicas para menores (Reyes & López, 2023), el último quinquenio ha consolidado el uso de entornos inmersivos que priorizan la validez ecológica a través de sistemas como CAVE y VR-CAT.

El análisis de resultados de este artículo confirma que las evaluaciones neuropsicológicas contemporáneas están logrando superar la brecha entre el laboratorio y la vida diaria. Plataformas como VR-CAT (Shen *et al.*, 2022) demuestran la viabilidad de evaluar la tríada de FE (inhibición, memoria de trabajo y flexibilidad) dentro de una narrativa gamificada. Este enfoque no solo incrementa la motivación del menor, sino que permite observar la organización de comportamientos ejecutivos de una forma que las pruebas de lápiz y papel no logran capturar. El uso de aulas virtuales y sistemas de proyección como CAVE (Bulgarelli *et al.*, 2023) responde directamente a la crítica sobre la limitada capacidad predictiva de las pruebas tradicionales en el comportamiento cotidiano.

Es relevante mencionar que la introducción de avatares y la medición de la conectividad funcional (fNIRS) en entornos sociales permite mapear la interacción entre la corteza prefrontal y el sistema límbico. Esto es particularmente crítico en la adolescencia, donde el desfase madurativo incrementa la impulsividad (Kusi-Mensah *et al.*, 2021). Al integrar la valoración emocional y la interacción social en la tarea ejecutiva, las nuevas plataformas de RV están logrando evaluar no solo la capacidad lógica o FE frías, sino también la regulación afectiva en contextos de alta demanda social.

La incorporación de neuroimagen funcional (fNIRS) junto con la RV representa un hito en la validación de los principios de Stuss y Alexander (2000). La posibilidad de monitorear la actividad de la CPF lateral mientras el niño navega en un entorno 3D permitiría corroborar la especialización hemisférica en tareas de preparación y monitoreo en tiempo real. Asimismo, la diversificación de dispositivos (desde visores HMD hasta sistemas móviles) sugiere una democratización de la tecnología, permitiendo que la evaluación neuropsicológica se traslade de centros de investigación especializados a entornos clínicos y educativos más diversos.

Sin embargo, a pesar del crecimiento en el número de artículos publicados (pasando de una evidencia "limitada" antes de 2021 a

16 estudios rigurosos identificados en este periodo), persisten desafíos como la accesibilidad de datos por la exclusión de artículos por falta de acceso al texto completo sugiere que parte de la evidencia científica sigue restringida por barreras económicas de publicación. Además, aunque las plataformas muestran alta validez ecológica, se requiere mayor consenso en los baremos de puntuación para que los resultados de la RV sean comparables con los estándares clínicos tradicionales como los test de lápiz-papel. Finalmente, aún se requiere profundizar en el análisis de cómo estas herramientas se adaptan a perfiles de neurodesarrollo específicos (como TDAH o TEA), donde las FE juegan un rol determinante.

A pesar de los hallazgos significativos, esta revisión sistemática presenta ciertas limitaciones que deben considerarse al interpretar los resultados tales como la restricción de bases de datos ya que la búsqueda se limitó a PubMed y PsycInfo, es posible que se hayan omitido desarrollos tecnológicos documentados en bases de datos de ingeniería o ciencias de la computación (como IEEE Xplore o Scopus), donde suelen publicarse los prototipos iniciales de plataformas de RV. Otra limitación que observamos es el sesgo de selección por la imposibilidad de acceder al texto completo de forma gratuita representa un sesgo de disponibilidad. Esto podría haber dejado fuera datos relevantes sobre plataformas emergentes que aún no cuentan con una distribución abierta en la comunidad científica. Una limitación más que vale la pena comentar es la ventana temporal proyectiva, dado que la revisión incluye una fase de búsqueda situada en el primer bimestre de 2026, la representatividad de dicho año es parcial. Esto implica que las conclusiones sobre las tendencias del último año del quinquenio son preliminares y podrían modificarse con las publicaciones subsiguientes de ese mismo ciclo. Finalmente, al limitar la inclusión de artículos en inglés y español, se excluyó potencialmente evidencia relevante proveniente de mercados tecnológicos robustos en Asia (China, Japón o Corea del Sur), donde el desarrollo de RV aplicada a la educación y salud infantil ha mostrado un crecimiento acelerado.

En conclusión, podemos señalar que la RV ha dejado de ser una herramienta experimental enfocada en población adulta para convertirse en un pilar de la evaluación neuropsicológica infanto-juvenil. La

transición hacia evaluaciones más inmersivas y ecológicas permite una comprensión más profunda del funcionamiento ejecutivo en el mundo real. Futuras investigaciones deberían centrarse en la creación de protocolos clínicos estandarizados y en el desarrollo de plataformas de bajo costo que permitan una detección temprana de disfunciones ejecutivas en el ámbito escolar.

Financiamiento: este artículo no recibió financiamiento para su elaboración.


Conflicto de intereses: los autores declaran no tener algún conflicto de intereses.

Referencias

- Badre, D., & D'Esposito, M. (2009). Is the rostro-caudal axis of the frontal lobe hierarchical? *Nature Reviews Neuroscience*, 10(9), 659-669. <https://doi.org/10.1038/nrn2667>
- Bernardelli, G., Arlati, S., Scaglione, A., Saligari, E., Frigerio, P., Sotgiu, A., Flori, V., Lucini, D., Vignoli, A., & Zangiacomi, A. (2025). Acceptance of virtual reality to promote attention orientation in children: a qualitative study among children with ADHD and neurotypical development. *Annals Of Medicine*, 57(1), 2548387. <https://doi.org/10.1080/07853890.2025.2548387>
- Bulgarelli, C., Pinti, P., Aburumman, N., & Jones, E. J. H. (2023). Combining wearable fNIRS and immersive virtual reality to study preschoolers' social development: a proof-of-principle study on preschoolers' social preference. *Oxford Open Neuroscience*, 2, kvad012. <https://doi.org/10.1093/oons/kvad012>
- Castilla, A., Borst, G., Cohen, D., Fradin, J., Lefrançois, C., Houdé, O., Zaoui, M., & Berthoz, A. (2021). A new paradigm for the study of cognitive flexibility in children and adolescents: The "Virtual House Locomotor Maze" (VHLM). *Frontiers In Psychiatry*, 12, 708378. <https://doi.org/10.3389/fpsyt.2021.708378>
- Diamond, A. (2012). Executive functions. *Annual Review of Psychology*, 64(1), 135-168. <https://doi.org/10.1146/annurev-psych-113011-143750>
- Fang, Y., Han, D., & Luo, H. (2019). A virtual reality application for assessment for attention deficit hyperactivity disorder in school-aged children. *Neuropsychiatric Disease And Treatment*, 15, 1517-1523. <https://doi.org/10.2147/ndt.s206742>

- Fernández-Martín, P., Rodríguez-Herrera, R., Cánovas, R., Díaz-Orueta, U., De Salazar, A. M., & Flores, P. (2023). Data-driven profiles of attention-deficit/hyperactivity disorder using objective and ecological measures of attention, distractibility, and hyperactivity. *European Child & Adolescent Psychiatry*, 33(5), 1451-1463. <https://doi.org/10.1007/s00787-023-02250-4>
- Giglioli, I. A. C., Gálvez, B. P., Granados, A. G., & Raya, M. A. (2021). The virtual cooking task: a preliminary comparison between neuropsychological and ecological virtual reality tests to assess executive functions alterations in patients affected by alcohol use disorder. *Cyberpsychology Behavior and Social Networking*, 24(10), 673-682. <https://doi.org/10.1089/cyber.2020.0560>
- Ide-Okochi, A., Matsunaga, N., & Sato, H. (2022). A preliminary study of assessing gaze, interoception and school performance among children with neurodevelopmental disorders: The Feasibility of VR Classroom. *Children*, 9(2), 250. <https://doi.org/10.3390/children9020250>
- Kusi-Mensah, K., Nuamah, N. D., Wemakor, S., Agorinya, J., Seidu, R., Martyn-Dickens, C., & Bateman, A. (2021). Assessment tools for executive function and adaptive function following brain pathology among children in developing country contexts: a scoping review of current tools. *Neuropsychology Review*, 32(3), 459-482. <https://doi.org/10.1007/s11065-021-09529-w>
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: A latent variable analysis. *Cognitive Psychology*, 41(1), 49-100. <https://doi.org/10.1006/cogp.1999.0734>
- Mouga, S., Duarte, I. C., Café, C., Sousa, D., Duque, F., Oliveira, G., & Castelo-Branco, M. (2021). Attentional cueing and executive deficits revealed by a virtual supermarket task coupled with eye-tracking in autism spectrum disorder. *Frontiers In Psychology*, 12, 671507. <https://doi.org/10.3389/fpsyg.2021.641507>
- Nir-Hadad, S. Y., Weiss, P. L., Waizman, A., Schwartz, N., & Kizony, R. (2015). A virtual shopping task for the assessment of executive functions: Validity for people with stroke. *Neuropsychological Rehabilitation*, 27(5), 808-833. <https://doi.org/10.1080/09602011.2015.1109523>
- Oh, S., Joung, Y., Chung, T., Lee, J., Seok, B. J., Kim, N., & Son, H. M. (2024). Diagnosis of ADHD using virtual reality and artificial intelligence: an exploratory study of clinical applications. *Frontiers In Psychiatry*, 15, 1383547. <https://doi.org/10.3389/fpsyg.2024.1383547>

- Reyes, A. C. D., & López, J. V. S. (2023). Evaluación neuropsicológica y realidad virtual: una revisión sistemática. *Informes Psicológicos*, 23(1), 108-124. <https://doi.org/10.18566/infpsic.v23n1a07>
- Robitaille, N., Jackson, P. L., Hébert, L. J., Mercier, C., Bouyer, L. J., Fecteau, S., Richards, C. L., & McFadyen, B. J. (2016). A Virtual Reality avatar interaction (VRai) platform to assess residual executive dysfunction in active military personnel with previous mild traumatic brain injury: proof of concept. *Disability And Rehabilitation Assistive Technology*, 12(7), 758-764. <https://doi.org/10.1080/17483107.2016.1229048>
- Romero-Ayuso, D., Del Pino-González, A., Torres-Jiménez, A., Juan-González, J., Celdrán, F. J., Franchella, M. C., Ortega-López, N., Triviño-Juárez, J. M., Garach-Gómez, A., Arrabal-Fernández, L., Medina-Martínez, I., & González, P. (2024). Enhancing ecological validity: virtual reality assessment of executive functioning in children and adolescents with ADHD. *Children*, 11(8), 986. <https://doi.org/10.3390/children11080986>
- Salehinejad, M. A., Ghanavati, E., Rashid, M. H. A., & Nitsche, M. A. (2021). Hot and cold executive functions in the brain: A prefrontal-cingular network. *Brain And Neuroscience Advances*, 5, 23982128211007769. <https://doi.org/10.1177/23982128211007769>
- Salmi, J., Merzon, L., Eräste, T., Seesjärvi, E., Huhdanpää, H., Aronen, E. T., Mannerkoski, M., MacInnes, W. J., & Laine, M. (2023). Fluctuations of attention during self-paced naturalistic goal-directed behavior in attention-deficit/hyperactivity disorder. *JAACAP Open*, 2(3), 188-198. <https://doi.org/10.1016/j.jaacop.2023.12.002>
- Seesjärvi, E., Puhakka, J., Aronen, E. T., Hering, A., Zuber, S., Merzon, L., Kliegel, M., Laine, M., & Salmi, J. (2022). EPELI: a novel virtual reality task for the assessment of goal-directed behavior in real-life contexts. *Psychological Research*, 87(6), 1899-1916. <https://doi.org/10.1007/s00426-022-01770-z>
- Seesjärvi, E., Puhakka, J., Aronen, E. T., Lipsanen, J., Mannerkoski, M., Hering, A., Zuber, S., Kliegel, M., Laine, M., & Salmi, J. (2021). Quantifying ADHD symptoms in open-ended everyday life contexts with a new virtual reality task. *Journal Of Attention Disorders*, 26(11), 1394-1411. <https://doi.org/10.1177/10870547211044214>
- Seesjärvi, E., Zuber, S., Joly-Burra, E., Kliegel, M., & Salmi, J. (2025). A naturalistic virtual reality task reveals difficulties in time-based prospective memory and strategic time-monitoring in children with ADHD. *Scientific Reports*, 15(1), 24722. <https://doi.org/10.1038/s41598-025-08944-w>

- 
- Shen, J., Koterba, C., Samora, J., Leonard, J., Li, R., Shi, J., Yeates, K. O., Xiang, H., & Taylor, H. G. (2022). Usability and validity of a virtual reality cognitive assessment tool for pediatric traumatic brain injury. *Rehabilitation Psychology*, 67(4), 587-596. <https://doi.org/10.1037/rep0000464>
- Stokes, J. D., Rizzo, A., Geng, J. J., & Schweitzer, J. B. (2022). Measuring attentional distraction in children with ADHD using virtual reality technology with eye-tracking. *Frontiers In Virtual Reality*, 3. <https://doi.org/10.3389/frvir.2022.855895>
- Stuss, D. T., & Alexander, M. P. (2000). Executive functions and the frontal lobes: a conceptual view. *Psychological Research*, 63(3-4), 289-298. <https://doi.org/10.1007/s004269900007>
- Tabibi, Z., Schwebel, D. C., & Juzdani, M. H. (2023). How does attention deficit hyperactivity disorder affect children's road-crossing? A case-control study. *Traffic Injury Prevention*, 24(4), 315-320. <https://doi.org/10.1080/15389588.2023.2181664>
- Verhoef, R. E. J., Van Dijk, A., Verhulp, E. E., & De Castro, B. O. (2021). Interactive virtual reality assessment of aggressive social information processing in boys with behaviour problems: A pilot study. *Clinical Psychology & Psychotherapy*, 28(3), 489-499. <https://doi.org/10.1002/ccp.2620>
- Ward, J. (2020). *The student's guide to cognitive neuroscience*. London: Taylor & Francis.