



Uso de juegos serios y realidad virtual en la evaluación de la memoria espacial para el diagnóstico temprano de Alzheimer: una revisión sistemática

Ana María Galvis Herrera
David Sánchez Trujillo
Laura Cristina Torres López

Artículo de investigación presentado para optar al título de Psicólogo

Asesor

Juan Pablo Sánchez Escudero, Magíster (MSc) en Psicología

Universidad de Antioquia
Facultad de Ciencias Sociales y Humanas
Psicología
Medellín, Antioquia, Colombia
2022

Cita

(Galvis Herrera et al., 2022)

Referencia

Estilo APA 7 (2020)

Galvis Herrera, A., Sánchez Trujillo, D., & Torres López, L. (2022). *Uso de juegos serios y realidad virtual en la evaluación de la memoria espacial para el diagnóstico temprano de Alzheimer: una revisión sistemática*. [Trabajo de grado profesional]. Universidad de Antioquia, Medellín, Colombia.



CRAI María Teresa Uribe (Facultad de Ciencias Sociales y Humanas)

Repositorio Institucional: <http://bibliotecadigital.udea.edu.co>

Universidad de Antioquia - www.udea.edu.co

Rector: John Jairo Arboleda Céspedes.

Decano/Director: Alba Nelly Gómez García.

Jefe departamento: Alberto Ferrer Botero.

El contenido de esta obra corresponde al derecho de expresión de los autores y no compromete el pensamiento institucional de la Universidad de Antioquia ni desata su responsabilidad frente a terceros. Los autores asumen la responsabilidad por los derechos de autor y conexos.

Resumen

En la última década, evidencia de múltiples investigaciones ha incitado la evaluación de la Memoria Espacial (ME) como marcador preclínico para la detección de la Enfermedad de Alzheimer (EA). La baja correlación de las pruebas neuropsicológicas de lápiz y papel con la ME en la vida cotidiana, y, dificultades en la evaluación de este proceso en entornos diferentes al investigativo, ha impulsado el desarrollo de instrumentos basados en tecnologías como la Realidad Virtual (RV) y los Juegos Serios (JS) para su evaluación. Dado el aumento de la producción científica en el campo y la falta de evidencia respecto a la calidad metodológica de los instrumentos desarrollados, en esta revisión sistemática se presenta la síntesis del análisis de la calidad metodológica de estos estudios. Para su evaluación se utilizó el *Consensus-based Standards for de Selection on health Measurement Instruments* (COSMIN). De acuerdo con los resultados, tanto en el desarrollo como en el análisis de las propiedades psicométricas de estos instrumentos, se presentan sesgos que limitan la posibilidad de recomendarlos como tecnologías válidas, confiables y aplicables a contextos poblacionales. Por lo anterior, se considera que la evaluación de la ME a través de herramientas tecnológicas como la RV y JS se encuentra en evolución y aún no está preparada para reemplazar las pruebas tradicionales en la evaluación de marcadores cognitivos de la EA, por lo que se recomienda, enfáticamente, el desarrollo de investigaciones posteriores que aporten nueva evidencia y suplan las limitaciones encontradas.

Palabras clave: juegos serios, realidad virtual, memoria espacial, enfermedad de alzhéimer, cosmin, deterioro cognitivo leve, revisión sistemática.

Abstract

In the last decade, evidence from multiple investigations has prompted the assessment of Spatial Memory (SM) as a preclinical marker for the detection of Alzheimer's Disease (AD). The low correlation of pencil and paper neuropsychological tests with SM in everyday life, and difficulties in the assessment of this process in environments different from research, has prompted the development of instruments based on technologies such as Virtual Reality (VR) and Serious Games (SG) for its evaluation. Due to the increase in scientific production in the field and the lack of evidence regarding the methodological quality of the instruments developed, this systematic review presents a synthesis of the analysis of the methodological quality of these studies. For this purpose, The Consensus-based Standards for Selection on Health Measurement Instruments (COSMIN) was used for their evaluation. According to the results, both in the development and in the analysis of the psychometric properties of these instruments there are biases that limit the possibility of recommending them as valid, reliable and applicable in population contexts. Therefore, it is considered that the assessment of SM through technological tools such as VR and SG is still evolving and is not ready yet to replace traditional tests in the assessment of cognitive markers of AD, so the development of further researchs is strongly recommended to provide new evidence and overcome the limitations found.

Keywords: serious games, virtual reality, spatial memory, alzheimer 's disease, systematic review, cosmin, mild cognitive impairment.

Introducción

Debido a la transición demográfica se espera que la prevalencia e incidencia de enfermedades asociadas al envejecimiento como la demencia por Enfermedad de Alzheimer (EA) alcancen niveles epidémicos para 2030 (World Health Organization [WHO], 2018). Ante la falta de tratamientos farmacológicos efectivos, los esfuerzos de la comunidad internacional se han centrado en la identificación temprana de la población en riesgo y el inicio de tratamientos preventivos en una etapa en la que la enfermedad no esté tan avanzada (WHO, 2018). Es así, que la búsqueda e identificación de marcadores cognitivos se ha propuesto como una opción viable para hacer frente a esta necesidad.

En la literatura especializada de los últimos 10 años, se ha propuesto que procesos cognitivos como la memoria espacial (ME) son sensibles a los cambios neuropatológicos tempranos característicos de la EA (Coughlan, et al., 2018). La memoria espacial se define como la capacidad para codificar, almacenar y recuperar información de las relaciones espaciales de los objetos en el entorno a partir de diferentes marcos de referencia (Vasilyeva y Lourenco, 2012; Konishi y Bohbot, 2013); esta información es indispensable para la formación de mapas cognitivos, que son representaciones mentales del espacio que permiten la navegación espacial, entendida a su vez como "*el proceso de determinar y mantener una trayectoria entre diferentes puntos del entorno*" (Coughlan, et al., 2018, p.3).

Evidencia científica en la última década, destaca que la navegación espacial puede considerarse como un potencial marcador cognitivo para la detección de la EA desde etapa preclínica (Coughlan et al., 2018). Dado que, a diferencia de otros marcadores cognitivos tradicionales como la memoria episódica, los procesos visuoespaciales no muestran deterioro en adultos mayores como producto de la edad, así como tampoco en otros tipos de demencia (e.g. demencia fronto-temporal) (Fernandez-Baizan et al., 2019).

Es así como, durante los últimos años, se han propuesto múltiples paradigmas para la evaluación objetiva de las habilidades de navegación espacial en población en riesgo de EA. A pesar de su relevancia, el desarrollo de este tipo de paradigmas ha supuesto un desafío para los investigadores en el área, dada la baja capacidad predictiva de las pruebas neuropsicológicas tradicionales, de lápiz y papel, sobre las habilidades de navegación en entornos reales (Costa et al., 2020; Serino et al., 2014). Ante estas limitaciones, la Realidad Virtual (RV) y los Juegos Serios

(JS) se han propuesto como alternativas de evaluación viables de estos potenciales marcadores en contextos de tamizaje de la EA preclínica en población en riesgo (Germine et al., 2019).

La RV consiste en la emulación de un entorno en tiempo real por medio de múltiples canales sensoriales, donde priman los canales visuales y auditivos (Díaz y Florez, 2018). Los JS por su parte, son juegos que se desarrollan con un objetivo específico, distinto al entretenimiento, y en el contexto de la investigación en salud han sido utilizados como herramientas para el diagnóstico, el tratamiento, la estimulación y la rehabilitación de sujetos con enfermedades neurodegenerativas como la EA (Ben-sadoun et al., 2018). Estas técnicas ofrecen beneficios en comparación con la evaluación neuropsicológica tradicional, de lápiz y papel, como la inclusión de un componente lúdico que promueve y mantiene la motivación del evaluado, así como la creación de un ambiente de evaluación regulado que reproduce escenarios de la vida cotidiana y que pueden llegar a ser percibidos como poco intrusivos por parte de los evaluados (Valladares et al., 2016). Además, posibilitan una medición estandarizada y precisa (i.e. en el rango de milisegundos) de los tiempos de respuesta, una mejor aleatorización de la presentación de los estímulos y la prevención de eventos incontrolables (Valladares et al., 2016; Peñate et al., 2014).

Aunque la RV y los JS destacan como una alternativa prometedora para la evaluación neuropsicológica, su implementación como instrumentos de medición en salud suponen superar una serie de limitaciones metodológicas (Valladares et al., 2016). Revisiones sistemáticas recientes (Montana et al., 2018; Cibeira et al., 2020) mencionan la falta de estudios frente a la calidad psicométrica de estas herramientas y la necesidad de encontrar evidencia científica respecto a la validez de estos. Por ejemplo, la mayoría de los análisis que se realizan a la hora de evaluar características como la consistencia interna y la validez de la RV y los JS se encuentran en un estado exploratorio, mientras que la consistencia en el tiempo para medir el impacto o el efecto de la práctica en este tipo de pruebas ha sido poco estudiada (Valladares et al., 2016). El método más usado para medir la validez de estas pruebas es a través de la validez convergente o divergente, mientras que casi ningún estudio realiza validez de constructo (Valladares et al., 2016; Cibeira et al., 2020).

Ahora bien, dado que utilizar instrumentos de medición de resultados de baja o desconocida calidad es una impropiedad metodológica en tanto constituyen un desperdicio de recursos y faltas éticas (Terwee et al, 2018), es relevante realizar un análisis de la calidad metodológica y riesgo de sesgo de las propiedades psicométricas en los instrumentos de medición basados en RV y JS. Hasta

donde es de conocimiento de los autores, dichos aspectos no han sido evaluados por anteriores revisiones sistemáticas, las cuales, se han centrado en la revisión de literatura a través de la guía *Preferred Reporting Items for Systematic Reviews and Meta-Analysis* (PRISMA) (Valladares et al., 2016) o el análisis de la exactitud diagnóstica (i.e. validez discriminante) a través de la guía *STAndards for the Reporting of Diagnostic accuracy studies* (STARD) (Tuena et al., 2020).

Por lo tanto, se propone una revisión sistemática de la literatura de validación de instrumentos de medición basados en software, para la detección de la EA. De esta manera, se plantea evaluar la calidad metodológica y riesgo de sesgo de los estudios publicados en la última década a través de la guía de verificación *Consensus-based Standards for de Selection on health Measurement Instruments* (COSMIN). Esta es una herramienta estandarizada para evaluar estos aspectos a través de calificaciones cuantitativas referentes a la validez y la confiabilidad de los instrumentos de medición en salud (Terwee, et al, 2018). Así pues, el objetivo de este estudio es sintetizar la información disponible respecto a las propiedades psicométricas de los principales paradigmas para la evaluación de estos procesos cognitivos, que hagan uso de herramientas como la RV y JS en contextos de tamizaje y diagnóstico de la EA.

Método

Diseño

Para el desarrollo de esta revisión sistemática se implementaron los pasos sugeridos por la guía COSMIN. De esta manera, en una primera etapa se aplicó el protocolo PRISMA (Moher et al., 2015), para la revisión de la literatura, tamizaje y selección de los artículos; y, en una segunda etapa, se utilizó la lista de chequeo COSMIN para la evaluación de la calidad metodológica general de los estudios y para la evaluación del riesgo de sesgo en el análisis y reporte de las propiedades psicométricas de los instrumentos de medición (Terwee et al., 2018). Los formatos originales de extracción de datos fueron adaptados para incluir criterios de evaluación como: la tecnología usada para el desarrollo del juego, los correlatos neuropsicológicos, el entorno de RV y los biomarcadores.

Estrategia de búsqueda

Se realizó una búsqueda electrónica entre el 26 de junio de 2021 y el 16 de julio del mismo año en las bases de datos PubMed/Medline, Scopus y EBSCO. Reconocidas por el acceso que ofrecen a artículos, revistas, libros y publicaciones de alta calidad. Esta búsqueda fue revisada por pares.

Los términos de la ecuación de búsqueda se eligieron deliberadamente en relación con los constructos (i.e. memoria espacial y navegación espacial), la herramienta de medida (RV y JS), el desenlace de salud (EA y Deterioro Cognitivo Leve [DCL]) y el contexto del uso de los instrumentos (evaluación). Se utilizaron los tesauros de *Descriptor en Ciencias de la Salud* (DeCS) (Panamerican Health Organization [PAHO] et al., 2017) y *Medical Subject Headings* (MeSH) (National Library of Medicine, s.f.), para la redacción de los términos de búsqueda. En la Tabla 1, se presenta la ecuación de búsqueda utilizada.

Tabla 1

Ecuación de búsqueda

Boleano	Categorías	Términos de búsqueda
	Constructo	("Spatial Navigation" OR "Spatial Memory")
AND	Herramienta de medida	("Serious Game*" OR "Virtual Reality" OR "Virtual" OR "Maze Learning" OR "Task")
AND	Desenlace de salud	("Alzheimer*" OR "Alzheimer Disease" OR "Mild Cognitive Impairment" OR "Dementia")
AND	Contexto de uso	("Screening" OR "Diagnos*" OR "Assess*" OR "Evaluat*")
NOT	Términos de exclusión	(Train* OR Rehabilitat* OR Drug* OR "Parkinson Disease" OR "Vascular Dementia" OR "VIH" OR "Diabetes" OR "Mice" OR "Rat" OR "Dog").

Criterios de elegibilidad

Los estudios se incluyeron en la revisión si (1) informaban de datos originales sobre el análisis de las propiedades psicométricas de instrumentos o tareas basados en software para la evaluación de la navegación espacial y la memoria espacial; (2) incluían participantes con riesgo de EA, en especial adultos mayores (>60 años) o participantes más jóvenes (<60 años) con factores de riesgo de EA de aparición temprana; (3) proporcionaban datos suficientes para analizar al menos una de las propiedades de medición incluidas en las directrices COSMIN; (4) utilizaban JS o RV

para adaptar o desarrollar la tarea; (5) el texto completo estaba disponible; (6) se publicó en inglés o español.

Se excluyeron otros tipos de literatura como resúmenes, literatura gris, documentos de trabajo, documentos de investigación argumentativos, documentos de definición, documentos interpretativos, editoriales y conferencias. Los estudios de intervención, así como otras revisiones sistemáticas o de literatura, no se consideraron para esta revisión.

Selección de los estudios

Los estudios encontrados en la consulta de bases de datos se importaron a la aplicación web Rayyan (Ouzzani, et al., 2016), la cual es una herramienta colaborativa para la síntesis de información en revisiones sistemáticas. En la primera etapa del tamizaje se eliminaron los artículos duplicados. En la segunda etapa del tamizaje, cuatro revisores independientes y capacitados en el manejo de la herramienta realizaron una lectura del título y resumen de cada estudio para el cumplimiento de los criterios de inclusión.

Se excluyeron aquellos artículos que cumplieran con uno o más de los criterios de exclusión, descritos en la Tabla 2. Las razones de exclusión se registraron a través de la herramienta de etiquetas proporcionada por Rayyan. Las dudas presentadas en relación con la adecuación de los estudios se resolvieron a través del consenso de los tres revisores y un cuarto experto.

Tabla 2
Criterios de exclusión

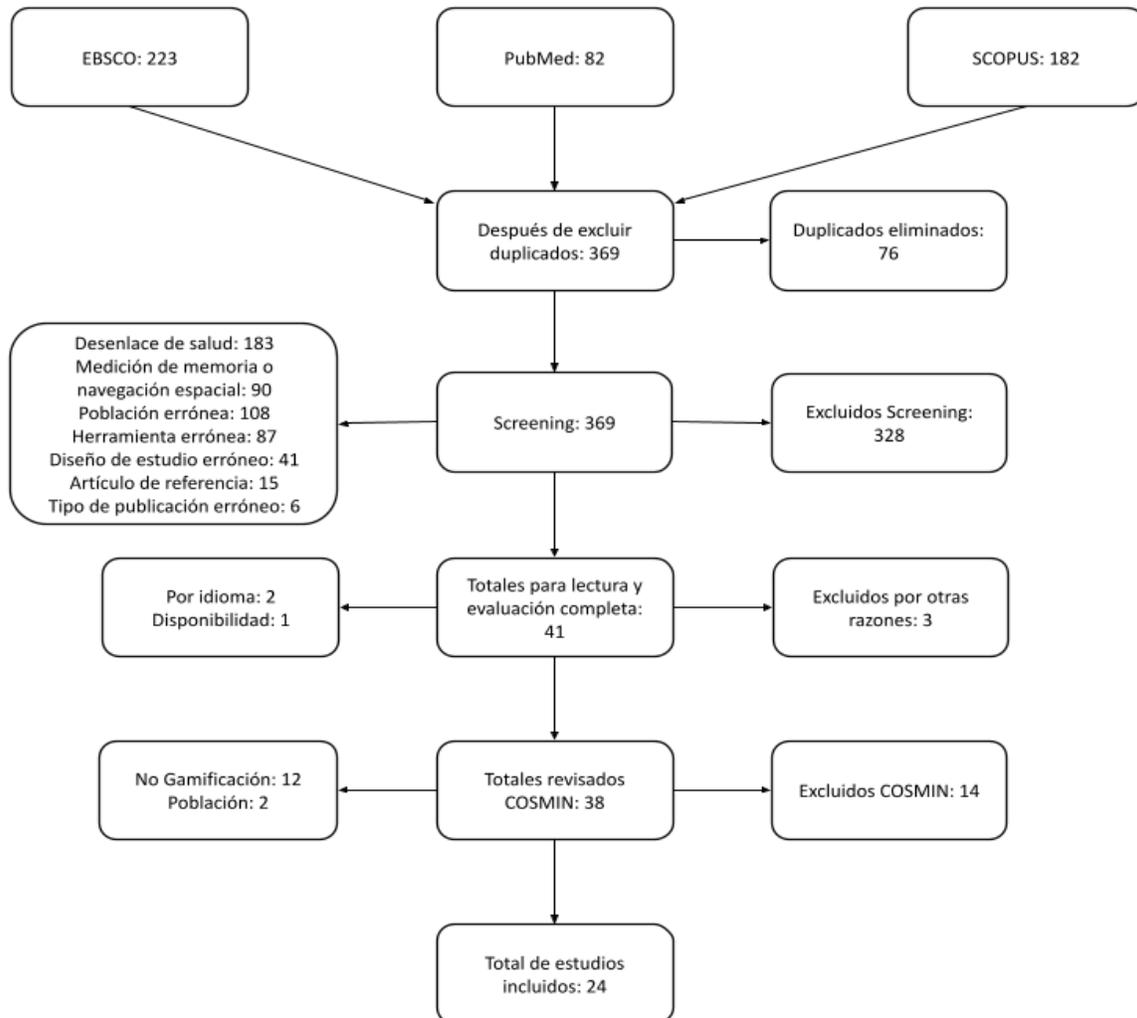
Criterio	Descripción
Artículo de referencia	Documentos que ofrecen datos relevantes sobre la patología y que podrían ser citados como parte del manuscrito. Sin embargo, no informa de pruebas adicionales de análisis de propiedades psicométricas. Esta etiqueta no se utilizaría si otra etiqueta es más adecuada.
Lenguaje erróneo	Artículo publicado en un idioma distinto al español o al inglés.
Desenlace de salud	El resultado primario estudiado no es la EA o el DCL.
Población errónea	Los participantes no son población de riesgo de EA o DCL.
Tipo de población errónea	No se trata de un trabajo de investigación analítica ni de un trabajo de investigación experimental.
Diseño de estudio erróneo	Se excluirán los ensayos clínicos aleatorios, sistemáticos y de revisión de la literatura.

Nota. EA: enfermedad de Alzheimer; DCL: deterioro cognitivo leve.

Tras culminar el proceso de tamizaje, los revisores extrajeron los datos relevantes de los artículos: autores y año de publicación, descripción de los instrumentos, tamaño de la muestra, correlatos neuropsicológicos, principales hallazgos en propiedades psicométricas, población, país, lenguaje y contexto de uso. Estos datos fueron registrados en una hoja de Microsoft Excel siguiendo las directrices de la guía COSMIN. En la Figura 1 se presenta el diagrama de flujo de la selección de los artículos.

Figura 1

Diagrama de flujo de selección de artículos.

**Evaluación de la calidad metodológica**

Se evaluó la calidad metodológica del análisis de las propiedades psicométricas de los instrumentos de medición reportados en los estudios a través de la lista de verificación de riesgo de sesgo de COSMIN. Esta lista está conformada por 10 cajas de revisión, cada una de estas permite calificar las propiedades psicométricas de los instrumentos a través del análisis del riesgo de sesgo presente en los estudios. Los dominios evaluados y su definición se presentan en la Tabla 3.

Tabla 3*Dominios evaluados y definición según COSMIN.*

Dominio	Propiedad de medición	Definición
Validez de contenido	Caja 1 - Calidad del diseño	Parte 1: Medida en que una PROM describe claramente su constructo y población objetivo. Parte 2: Medida en que una PROM realiza prueba piloto o entrevista cognitiva brindando una clara descripción del procedimiento
	Caja 2 - Validez de contenido	Medida en que los elementos de una PROM son representativos de la construcción que están destinados a medir
Estructura interna	Caja 3 - Validez estructural	Medida en que los elementos de una PROM reflejan la dimensionalidad del constructo (es decir, los ítems forman una sola escala [unidimensional] o múltiples subescalas [una escala multidimensional])
	Caja 4 - Consistencia interna	Medida en que hay consistencia de los resultados entre los elementos del PROM (es decir, dentro de una escala o subescala especificada)
	Caja 5 - Validez transcultural / Invarianza de la medida	Medida en que las propiedades de medición de lo traducido o culturalmente PROM adaptado refleja el rendimiento de la versión original del PROM
	Caja 6 - Confiabilidad Test-retest	Medida en que la PROM produce puntajes que son reproducibles (estables) a lo largo momento en que no ha habido cambio en el concepto que se está evaluando
Propiedades restantes	Caja 7 - Error de medida	El error sistemático y aleatorio de la puntuación de una persona en el PROM que no se atribuye a los cambios en el constructo a medir
	Caja 8 - Validez de criterio	La medida en que las puntuaciones de una PROM reflejan las puntuaciones de una prueba o medida considerado como el "estándar de oro"
	Caja 9 - Prueba de hipótesis para la validez de constructo	La medida en que las puntuaciones de una PROM son consistentes con las hipótesis. Por ejemplo, con respecto a las relaciones internas, las relaciones con decenas de otros instrumentos o diferencias entre grupos relevantes. Se basa en la suposición que la PROM es una medida válida del constructo
	Caja 10 - Sensibilidad	La capacidad de una PROM para detectar cambios, como se esperaba, a lo largo del tiempo en el constructo para medirse cuando hay un cambio real en la condición o el tratamiento de una persona

Nota. PROM: patient-reported outcome measures.

Resultados

En total, 24 estudios fueron seleccionados como elegibles y evaluados a través de la lista de verificación de riesgo de sesgo COSMIN. Las propiedades psicométricas de los instrumentos de medición de cada artículo y su respectiva valoración se presentan en la Tabla 4. Los dominios se calificaron con base en una escala de cinco descriptores: “muy bueno”, “adecuado”, “dudoso”, “inadecuado” y “no aplica”, según el protocolo. Un dominio fue evaluado solo si el estudio presentaba los datos suficientes para su valoración. La calidad metodológica de cada dominio se determinó tomando la calificación más baja de cualquiera de los elementos de la caja, siguiendo el “*the worst-score counts principle*”, es decir, la calificación global del dominio correspondió al puntaje más bajo de la caja, aunque otros elementos presentaran calificaciones altas.

Tabla 4
Evaluación de Propiedades Psicométricas.

Referencia	B1. Calidad del diseño	B3. Validez de estructura	B4. Consistencia interna	B6. Confiabilidad	B8. Validez de criterio	B9a. Validez convergente	B9b. Validez discriminativa
Allison et al, 2016	I		D		MB		MB
Allison et al, 2019	I	I	MB	I	MB		A
Bellassen et al, 2011	I	I			I		A
Caffo et al, 2012	I				MB		A
Caffo et al, 2017	I		D				A
Coughlan et al, 2020	I			D			MB
Da Costa et al, 2021	I				MB	D	A
Davis y Sikorskii, 2020	I						D
Kerbler et al, 2015	I					MB	
Konishi et al, 2018	I						D
Laczo et al, 2012	I					A	A
Laczo et al, 2014	I						A
Lee et al, 2014	I					A	
Lesk et al, 2014	I					D	D
Migo et al, 2015	I						MB
Mohammadi et al, 2018	I						A
Morganti et al, 2013	I					MB	A
Parizkova et al, 2017	I					MB	MB
Pengas et al, 2010	I				MB	MB	MB
Serino et al, 2015	I					MB	MB
Serino et al, 2018	I					D	D
Tarnanas et al, 2012	I				MB	I	MB
Weniger et al, 2011	I					D	MB
Widmann et al, 2012	I				D		MB

Nota: I = Inadecuado, D=Dudoso, A= Adecuado, MB= Muy bueno.

Todos los artículos evaluados presentaron un nivel *inadecuado* en la calidad del diseño de los estudios (Box 1). Un total de cinco (20.8%) de los estudios no proveyeron una descripción clara del constructo a medir, además siete (29.12%) estudios no informaron un marco de referencia conceptual bien definido. La representatividad de la muestra fue clasificada como dudosa en 22 (91.5%) de los estudios evaluados. Ninguno de los estudios evaluados realizó una entrevista

cognitiva o prueba piloto (Box 2) para profundizar en la comprensibilidad y exhaustividad de la herramienta de medida usada.

Solo dos (8%) de los artículos seleccionados realizaron análisis de validez de estructura (Box 3). Aunque uno de ellos llevó a cabo un análisis factorial confirmatorio, alcanzando una clasificación como *muy bueno*, y, el otro ejecutó un análisis factorial exploratorio, alcanzando una clasificación de *adecuado*, ambos obtuvieron una calificación de *inadecuado* en este dominio debido al tamaño de la muestra. Para evaluar el tamaño de muestra se siguió el criterio de regla de oro (*thumb rule*), en el cual se consideró que muestras mayores a 200 participantes tienen una calificación de *muy bueno*, las muestras entre 150 – 200 de *adecuado*, de 100 – 150 de *dudoso* y menores a 100 de *inadecuado*. Según esto, ambos estudios tuvieron un tamaño de muestra *inadecuado*, puesto que el primero presentó una muestra total de 91 participantes y el segundo de 94 participantes.

Tres (12.5%) de los estudios evaluaron consistencia interna (Box 4). De estos, dos (66.6%) no clarificaron si contaban con subescalas, por lo que obtuvieron una calificación de *dudoso*, mientras que el artículo restante calculó consistencia interna para cada una de ellas, por lo que se calificó como *muy bueno*. Los tres artículos utilizaron los estadísticos de alfa de Cronbach u Omega de McDonald para calcular la consistencia en las puntuaciones continuas y uno de ellos utilizó KR-20 para puntuaciones dicotómicas, por lo que obtuvieron una calificación de *muy bueno* en este criterio del dominio.

En el caso de la confiabilidad (Box 6), solo dos (8.3%) investigaciones realizaron análisis de test-retest. Ambos estudios presentaron un intervalo de tiempo apropiado entre la aplicación de las pruebas, por lo que se les calificó como *muy bueno* en este criterio. Igualmente, proporcionaron información suficiente para asumir que las condiciones de ambas aplicaciones fueron similares, de esta manera, obtuvieron un puntaje de *adecuado* en este criterio. El estadístico usado en ambos estudios fue el de correlación intraclass sin una descripción detallada, alcanzando una calificación de *adecuado*, sin embargo, uno de los estudios no calculó Kappa para las puntuaciones ordinales, por lo que se le asignó una puntuación de *inadecuado* en el dominio. El otro artículo obtuvo una calificación de *dudoso* en el dominio debido a que no informó claramente la estabilidad del constructo en los pacientes entre ambas aplicaciones.

Referente a la validez de criterio (Box 8), ocho (33.3%) de los artículos la estimó. De estos, siete (87.5%) evaluaron el área bajo la curva para puntuaciones continuas, obteniendo una

calificación de *muy bueno*, y tres (37.5%) sensibilidad y especificidad para las puntuaciones dicotómicas. Algunos sesgos importantes que surgieron fueron la no justificación del *Gold Standard* y la no descripción del punto de corte óptimo para las curvas ROC.

Todos los estudios incluidos en la investigación presentaron información psicométrica relacionada con la prueba de hipótesis para la validez de constructo (Box 9). 12 (50%) de los artículos revisados compararon sus instrumentos con otras pruebas (box 9a), cinco (41.5%) de ellos proveyeron una clara descripción de las pruebas de comparación, y las propiedades de medida del instrumento con el que se compararon son apropiadas, obteniendo un puntaje de *muy bueno*. Dos (16.6%) de los estudios tuvieron carencias menores en la descripción de las propiedades de medida del instrumento de comparación, por lo que no es claro si aplican a la población de estudio, por consiguiente, se les asignó una puntuación de *adecuado*. Por otro lado, cuatro (33.2%) de los estudios presentaron información sobre las propiedades de medida de los instrumentos de comparación, pero no especificaron y no hubo claridad sobre a qué población del estudio se aplicó, por ende, se calificó *dudoso*. Por último, uno (8.3%) de los artículos incluidos no presentó constructos claros en sus instrumentos de comparación, siendo *inadecuado* según la escala de criterios.

De la totalidad de los estudios, 22 (92%) realizaron comparación entre grupos para probar hipótesis (Box 9b) asociadas a la variación de las puntuaciones del instrumento. De ellos, nueve (40.8%) realizaron una adecuada descripción de los subgrupos y calcularon la diferencia entre grupos con métodos estadísticos apropiados, por ende, se calificó como *muy bueno*. Nueve (40.8%) de los artículos revisados obtuvieron *adecuado*, debido a que tienen carencias menores en la descripción de los subgrupos o en la del método estadístico utilizado para la comparación. Por último, cuatro (18%) de las investigaciones no proveen una descripción adecuada de los subgrupos o su método estadístico para realizar la comparación no fue óptimo, en este sentido se les asignó un puntaje de *dudoso*.

Uno de los hallazgos a destacar de la revisión es el uso de periféricos para la evaluación de los participantes. Se encontró que 13 (54.1%) de los estudios revisados utilizaban *joysticks* para moverse dentro de las aplicaciones; también se usaron otras herramientas: tres (12.5%) usaron *gamepads*, dos (8.3%) teclados, tres (12.5%) pantallas táctiles, uno (4.1%) *trackerball*, uno (4.1%) ratón y uno (4.1%) sensores de movimiento. Las características principales de muestra,

instrumento, constructo, tecnología, técnicas, diseño, correlatos neuropsicológicos y biomarcador de cada estudio se presentan en el Anexo 1.

Discusión

En los estudios analizados no se aprecia un marco teórico común para brindar una descripción homogénea del constructo medido, de esta manera, se referencia indiscriminadamente entre navegación espacial (Allison et al., 2016; Allison et al., 2019; Kerbler et al., 2015; Konishi et al., 2018; Laczó et al., 2012; Laczó et al., 2014; Migo et al., 2016; Mohammadi et al., 2018; Morganti et al., 2013; Pariskova et al., 2018; Serino et al., 2015; Serino et al., 2018; Weniger et al., 2011), memoria espacial (Coughlan et al 2020; Widman et al 2012), memoria visoespacial (Lee et al., 2014), memoria espaciotemporal (Bellasen et al., 2012), reorientación espacial (Caffó et al., 2012; Caffó et al., 2018), orientación espacial (Da costa et al., 2021), memoria topográfica (Lesk et al., 2014; Pengas et al., 2010) y cognición espacial (Tarnanas et al., 2012). Esta polisemia en el constructo incide directamente en la calidad de los análisis psicométricos de validez convergente y validez de contenido realizados por los estudios, debido a que el uso de diferentes conceptos para aludir a un mismo fenómeno puede ocasionar variaciones considerables en su operacionalización que generen inconsistencias en el análisis de sus propiedades psicométricas. Por ejemplo, el análisis de validez convergente requiere un punto de comparación para establecer la correspondencia entre dos o más instrumentos, por ende, estos deben medir el mismo constructo. Por otro lado, el análisis de validez de contenido requiere un constructo sólido, claramente definido, para determinar en qué medida los ítems de un instrumento son representativos de este.

En ninguno de los estudios evaluados se implementó un muestreo aleatorio para la selección del grupo clínico. En todos los casos, los participantes fueron seleccionados por medio de un muestreo no probabilístico por conveniencia en centros de investigación universitarios (Allison et al 2016; Allison et al 2019; Lesk et al., 2014), centros clínicos u hospitalarios (Bellasen et al., 2012; Caffó et al., 2012; Da costa et al., 2021; Kerbler et al., 2015; Laczó et al., 2014; Migo et al., 2016; Mohammadi et al., 2018; Morganti et al., 2013; Parizkova et al., 2018; Pengas et al., 2010; Serino et al., 2015; Weniger et al., 2011; Widman et al., 2012), geriátricos (Caffó et al., 2018; Serino et al., 2018) o de otras investigaciones (Davis y Sikorskii, 2020) . Algunos artículos no reportaron el método de selección de la muestra (Coughlan et al 2020; Konishi et al., 2018; Laczó et al., 2012;

Lee et al., 2014; Tarnanas et al., 2012). La implementación de estrategias de muestreo no probabilísticos, como el muestreo por conveniencia afecta la representatividad de la muestra respecto a la población, así como la generalización de los resultados.

El 29% de los estudios analizados se observa una disparidad en el tamaño de los grupos comparados en los cuales se duplica la cantidad de sujetos; por ejemplo, Caffó et al. (2018), compara un grupo de 203 controles sanos con otro grupo de 14 DCL-a Ud, 22 DCL-na Ud, 16 DCL-a Md, 10 DCL-na Md, y, 21 Prob. Dem¹. EL 79% de los estudios realizaron una comparación de grupos a través del análisis de varianza (ANOVA) y análisis de covarianza (ANCOVA) para determinar si existen diferencias entre los mismos, sin embargo, solo ocho realizaron un análisis Post-Hoc para determinar las medias que difieren. Para realizar comparación de grupos usando ANOVA se debe cumplir con el supuesto de la homocedasticidad, la cual se define como la homogeneidad de las varianzas. Una disparidad en los grupos a comparar puede afectar este supuesto, lo que implica que debe ser vigilado a través de estadísticos como la prueba de Levene. Ninguno de los artículos con disparidad en los grupos presenta evidencia de homocedasticidad, en consecuencia, es posible que sus resultados sean imprecisos y no fiables (O'Neill, M. E., & Mathews, K. L., 2002).

Además de los hallazgos descritos anteriormente se observó que ninguno de los estudios reporta el poder estadístico de la muestra para la implementación de las diferentes mediciones psicométricas, lo que unido a la baja representatividad de la muestra y muestreo no probabilístico limita considerablemente la posibilidad de generalizar los resultados al existir una alta probabilidad de errores tipo I (falso negativo) I en la prueba de hipótesis.

Dentro de los 22 artículos (92%) que evalúan la validez de instrumento a través de la comparación de grupos, es necesario destacar que los subgrupos seleccionados diferían entre las investigaciones; el 62.6% compararon DCL de diversos subtipos (Allison, et al., 2016; Bellasen, et al., 2012; Caffó, et al., 2012; Caffó, et al., 2018; Da costa, et al., 2021; Laczó et al., 2012; Laczó, et al., 2014; Lesk, et al., 2014; Migo, et al., 2016; Mohammadi, et al., 2018; Pariskova, et al., 2018; Pengas, et al., 2010; Serino, 2015; Tarnanas, et al., 2012; Weniger, 2011), el 8.34% otro tipo de demencias (Bellasen, et al., 2012; Pengas, et al., 2010) y el 29.2% comparó personas sanas con y

¹ **DCL-a Ud:** deterioro cognitivo leve amnésico de único dominio; **DCL-a Md:** deterioro cognitivo leve amnésico de múltiple dominio; **DCL-na Md:** deterioro cognitivo leve no amnésico de múltiple dominio; **Prob. Dem.:** probable demencia.

sin biomarcadores (Allison, et al., 2019; Coughlan et al., 2020; Davis y Sikorskii, 2020; Konishi, et al., 2018; Morganti, et al., 2013; Serino, 2018; Widman, et al., 2012). El uso de este método se puede deber a que los estudios se encuentran en una etapa inicial en el desarrollo de sus instrumentos, en la que se busca determinar si estos discriminan el proceso cognitivo a medir en los diferentes subgrupos (Newman y Kohn, 2020). La preferencia por este tipo de validez se puede deber también a limitaciones en el acceso a la población y sesgos de selección que llevan a sesgos de espectro en la medición.

Teniendo en cuenta que la población objetivo de dichas investigaciones está conformada por adultos mayores, es relevante considerar que el uso de algunos periféricos puede afectar la usabilidad del instrumento. En el caso del uso del *gamepad* (Konishi et al., 2018; Serino et al., 2015; Serino et al., 2018), la *trackerball* (Migo et al., 2016) y el uso de sensores de la posición y movimiento del cuerpo (Da Costa et al., 2021), se considera que podría haber sesgos asociados a la usabilidad del instrumento, pues dichos periféricos pueden llegar a requerir habilidades motrices que podrían estar reducidas o jamás haber sido desarrolladas por la población adulta mayor (Jin et al., 2020). Por otro lado, la mayoría de los estudios revisados se visualizaron en pantallas de computador portátil o de escritorio (71%), sin embargo, algunas de las investigaciones utilizan herramientas de RV montados en la cabeza como el *Oculus Rift* (Da Costa et al., 2021), lo que podría generar problemas en la medición debido a que en adultos mayores resulta más difícil usar dispositivos inmersivos dada la baja exposición de este grupo poblacional a la tecnología 3D (Morganti et al., 2013) y pueden generar mareos por el movimiento (Jin et al., 2020).

Recomendaciones

A lo largo de la revisión se recolectó información útil para el crecimiento del campo del conocimiento respecto a la calidad metodológica, validez y confiabilidad de las herramientas tecnológicas usadas en el tamizaje de la EA. Uno de los hallazgos más relevantes en los artículos revisados es que ninguno de ellos reportó haber realizado una etapa de pilotaje de sus instrumentos ni las modificaciones que realizaron a partir de esta. Teniendo en cuenta que la población objetivo de dichas investigaciones está conformada por adultos mayores, se considera importante recolectar información preliminar del ajuste de los participantes a las herramientas tecnológicas, teniendo en

cuenta sus necesidades y limitaciones propias del envejecimiento, que podrían dificultar el uso de la RV y JS (Jin, et al., 2020).

Otro hallazgo importante se relaciona con la validez y la confiabilidad de los instrumentos de tamizaje. Como se mencionó anteriormente, el método más frecuente para determinar la validez en los artículos fue la prueba de hipótesis por comparación de grupos o validez convergente con otras pruebas psicológicas. Sin embargo, se considera importante que se continúe evaluando la validez de estructura, constructo, contenido y criterio, así como la confiabilidad de los instrumentos de tamizaje, para que estos puedan abrirse paso como herramientas de ayuda diagnóstica en el contexto sanitario para la EA (Valladares et al., 2016).

El uso de las aplicaciones tecnológicas para la evaluación de marcadores cognitivos podría traer dificultades debido a las condiciones específicas de la población objetivo de los artículos revisados. Dichas dificultades podrían influir en la medición de los marcadores cognitivos que se pretenden evaluar, por lo que se hace necesario construir herramientas tecnológicas que se ajusten a las necesidades de los adultos mayores y permitan discriminar la navegación espacial sin que existan sesgos relacionados con el ajuste del participante con el instrumento (Morganti, et al., 2013; Jin, et al., 2020).

Teniendo en cuenta las limitaciones expuestas hasta el momento puede concluirse, de manera similar a como anteriores revisiones sistemáticas lo han hecho (Montana et al, 2018; Cibeira et al., 2020; Valladares et al., 2016; Tuena et al., 2020), que la evaluación de la NE a través de herramientas tecnológicas como la RV y JS se encuentra en evolución y aún no está preparada para reemplazar a las pruebas tradicionales en la evaluación de marcadores cognitivos de la EA, por lo que se recomienda enfáticamente el desarrollo de investigaciones posteriores que aporten nueva evidencia y suplan las limitaciones encontradas. Para este fin, la síntesis realizada en esta revisión sistemática es de utilidad para identificar sesgos comunes y vacíos en las investigaciones de este campo, que permitan que los autores que construyan instrumentos de tamizaje para la EA en el futuro tengan más información sobre la calidad metodológica y las propiedades psicométricas de las herramientas usadas hasta el momento, para que de esta forma, puedan encaminar su investigación hacia una mayor validez y confiabilidad en sus pruebas de medición.

Referencias

- Allison, S. L., Fagan, A. M., Morris, J. C., & Head, D. (2016). Spatial Navigation in Preclinical Alzheimer's Disease. *Journal of Alzheimer's Disease*, *52*(1), 77–90. <https://doi.org/f8jq4k>
- Allison, S. L., Rodebaugh, T. L., Johnston, C., Fagan, A. M., Morris, J. C., & Head, D. (2019). Developing a Spatial Navigation Screening Tool Sensitive to the Preclinical Alzheimer Disease Continuum. *Archives of Clinical Neuropsychology*, *34*(7), 1138–1155. <https://doi.org/h2zw>
- Bellassen, V., Igloi, K., de Souza, L. C., Dubois, B., & Rondi-Reig, L. (2012). Temporal Order Memory Assessed during Spatiotemporal Navigation As a Behavioral Cognitive Marker for Differential Alzheimer's Disease Diagnosis. *Journal of Neuroscience*, *32*(6), 1942–1952. <https://doi.org/f3wkpx>
- Ben-sadoun, G., Manera, V., Alvarez, J., Sacco, G., Robert, P., Kuljiš, R. O., & García-Betances, R. I. (2018). Recommendations for the Design of Serious Games in Neurodegenerative Diseases. *Frontiers in Aging Neuroscience*, *10* (13). <https://doi.org/gf4h3h>
- Caffò, A. O., de Caro, M. F., Picucci, L., Notarnicola, A., Settanni, A., Livrea, P., Lancioni, G. E., & Bosco, A. (2012). Reorientation Deficits Are Associated With Amnestic Mild Cognitive Impairment. *American Journal of Alzheimer's Disease & Other Dementias*, *27*(5), 321–330. <https://doi.org/f34k8b>
- Caffò, A. O., Lopez, A., Spano, G., Serino, S., Cipresso, P., Stasolla, F., Savino, M., Lancioni, G. E., Riva, G., & Bosco, A. (2017). Spatial reorientation decline in aging: the combination of geometry and landmarks. *Aging & Mental Health*, *22*(10), 1372–1383. <https://doi.org/10.1080/13607863.2017.1354973>
- Cibeira, N., Lorenzo López, L., Maseda, A., López López, R., Moreno Peral, P., & Millán Calenti, J. C. (2020). Realidad virtual como herramienta de prevención, diagnóstico y tratamiento del deterioro cognitivo en personas mayores: revisión sistemática. *Revista de Neurología*, *71*(06), 205-212. <https://doi.org/h2zx>

- Costa, R. Q. M. D., Pompeu, J. E., Viveiro, L. A. P. D., & Brucki, S. M. D. (2020). Spatial orientation tasks show moderate to high accuracy for the diagnosis of mild cognitive impairment: a systematic literature review. *Arquivos de Neuro-Psiquiatria*, 78(11), 713–723. <https://doi.org/h2z2>
- Coughlan, G., Laczó, J., Hort, J., Minihane, A. M., & Hornberger, M. (2018). Spatial navigation deficits — overlooked cognitive marker for preclinical Alzheimer disease? *Nature Reviews Neurology*, 14(8), 496–506. <https://doi.org/gd2pmv>
- Coughlan, G., Puthusseryppady, V., Lowry, E., Gillings, R., Spiers, H., Minihane, A. M., & Hornberger, M. (2020). Test-retest reliability of spatial navigation in adults at-risk of Alzheimer’s disease. *PLOS ONE*, 15(9), 1-13. <https://doi.org/ghhpx6>
- Davis, R., & Sikorskii, A. (2020). Eye Tracking Analysis of Visual Cues during Wayfinding in Early Stage Alzheimer’s Disease. *Dementia and Geriatric Cognitive Disorders*, 49(1), 91–97. <https://doi.org/gg64k2>
- Díaz Pérez, E., & Flórez Lozano, J. A. (2018). Realidad virtual y demencia. *Revista de Neurología*, 66(10), 344-352. <https://doi.org/h2z3>
- Fernandez-Baizan, C., Diaz-Caceres, E., Arias, J., & Mendez, M. (2019). Egocentric and allocentric spatial memory in healthy aging: performance on real-world tasks. *Brazilian Journal of Medical and Biological Research*, 52(4), 1-7. <https://doi.org/h2z4>
- Germine, L., Reinecke, K., & Chaytor, N. S. (2019). Digital neuropsychology: Challenges and opportunities at the intersection of science and software. *The Clinical Neuropsychologist*, 33(2), 271–286. <https://doi.org/gmnpkg>
- Jin, R., Pillozzi, A., & Huang, X. (2020). Current Cognition Tests, Potential Virtual Reality Applications, and Serious Games in Cognitive Assessment and Non-Pharmacological Therapy for Neurocognitive Disorders. *Journal of Clinical Medicine*, 9(10), 1-20. <https://doi.org/gmcpzm>.

- Kerbler, G. M., Nedelska, Z., Fripp, J., Laczó, J., Vyhnalek, M., Lisý, J., Hamlin, A. S., Rose, S., Hort, J., & Coulson, E. J. (2015). Basal Forebrain Atrophy Contributes to Allocentric Navigation Impairment in Alzheimer's Disease Patients. *Frontiers in Aging Neuroscience*, 7, 185-195. <https://doi.org/gg5wz2>
- Konishi, K., & Bohbot, V. D. (2013). Spatial navigational strategies correlate with gray matter in the hippocampus of healthy older adults tested in a virtual maze. *Frontiers in Aging Neuroscience*, 5, 1-8. <https://doi.org/ghjgmw>
- Konishi, K., Joobar, R., Poirier, J., MacDonald, K., Chakravarty, M., Patel, R., Breitner, J., & Bohbot, V. D. (2018). Healthy versus Entorhinal Cortical Atrophy Identification in Asymptomatic APOE4 Carriers at Risk for Alzheimer's Disease. *Journal of Alzheimer's Disease*, 61(4), 1493–1507. <https://doi.org/gcxzcq>
- Laczó, J., Anđel, R., Vyhnalek, M., Vlcek, K., Magerova, H., Varjassyova, A., Nedelska, Z., Gazova, I., Bojar, M., Sheardova, K., & Hort, J. (2012). From Morris Water Maze to Computer Tests in the Prediction of Alzheimer's Disease. *Neurodegenerative Diseases*, 10(1–4), 153–157. <https://doi.org/fxjx2g>
- Laczó, J., Anđel, R., Vyhnalek, M., Vlcek, K., Nedelska, Z., Matoska, V., Gazova, I., Mokrisova, I., Sheardova, K., & Hort, J. (2014). APOE and spatial navigation in amnesic MCI: Results from a computer-based test. *Neuropsychology*, 28(5), 676–684. <https://doi.org/f6f4tj>
- Migo, E., O'Daly, O., Mitterschiffthaler, M., Antonova, E., Dawson, G., Dourish, C., Craig, K., Simmons, A., Wilcock, G., McCulloch, E., Jackson, S., Kopelman, M., Williams, S., & Morris, R. (2015). Investigating virtual reality navigation in amnesic mild cognitive impairment using fMRI. *Aging, Neuropsychology, and Cognition*, 23(2), 196–217. <https://doi.org/gg64j2>
- Mohammadi, A., Kargar, M., & Hesami, E. (2018). Using virtual reality to distinguish subjects with multiple- but not single-domain amnesic mild cognitive impairment from normal elderly subjects. *Psychogeriatrics*, 18(2), 132–142. <https://doi.org/gmtdbg>

- Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, 4(1). <https://doi.org/gcsk56>
- Montana, J., Tuena, C., Serino, S., Cipresso, P., & Riva, G. (2019). Neurorehabilitation of Spatial Memory Using Virtual Environments: A Systematic Review. *Journal of Clinical Medicine*, 8(10), 1-25. <https://doi.org/gg64d8>
- National Library of Medicine. (s.f.). *MeSH Database*. Recuperado el 20 de mayo de 2021 de <https://bit.ly/3NiGBGk>
- Newman, T., & Kohn, M. (2020). *Evidence-Based Diagnosis: An Introduction to Clinical Epidemiology* (2nd Revised ed.). Cambridge University Press.
- O'Neill, M. E., & Mathews, K. L. (2002). Levene Tests of Homogeneity of Variance for General Block and Treatment Designs. *Biometrics*, 58(1), 216–224. <https://doi.org/fvrcp3>
- Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—a web and mobile app for systematic reviews. *Systematic Reviews*, 5(1), 1-10. <https://doi.org/gfkdzd>
- Panamerican Health Organization, World Health Organization & Centro Latinoamericano y del Caribe de Información en Ciencias de la Salud. (2017). *Health Sciences Descriptors: DeCS*. <https://bit.ly/3bIldlf>
- Parizkova, M., Lerch, O., Moffat, S. D., Andel, R., Mazancova, A. F., Nedelska, Z., Vyhnalek, M., Hort, J., & Laczó, J. (2018). The effect of Alzheimer's disease on spatial navigation strategies. *Neurobiology of Aging*, 64, 107–115. <https://doi.org/h2z5>
- Pengas, G., Patterson, K., Arnold, R. J., Bird, C. M., Burgess, N., & Nestor, P. J. (2010). Lost and Found: Bespoke Memory Testing for Alzheimer's Disease and Semantic Dementia. *Journal of Alzheimer's Disease*, 21(4), 1347–1365. <https://doi.org/f99tfg>
- Peñate Castro, W., Roca Sánchez, M. J., & del Pino Sedeño, T. (2014). Los nuevos desarrollos tecnológicos aplicados al tratamiento psicológico. *Acta Colombiana de Psicología*, 17(2), 91–101. <https://doi.org/h2z6>

- Serino, S., Cipresso, P., Morganti, F., & Riva, G. (2014). The role of egocentric and allocentric abilities in Alzheimer's disease: A systematic review. *Ageing Research Reviews*, *16*, 32–44. <https://doi.org/h2z7>
- Serino, S., Morganti, F., di Stefano, F., & Riva, G. (2015). Detecting early egocentric and allocentric impairments deficits in Alzheimer's disease: an experimental study with virtual reality. *Frontiers in Aging Neuroscience*, *7*, 88-98. <https://doi.org/gg64gk>
- Serino, S., Morganti, F., Colombo, D., & Riva, G. (2018). The Contribution of Allocentric Impairments to the Cognitive Decline in Alzheimer's Disease. *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, 84–91. <https://doi.org/h2z9>
- Tarnanas, I., Laskaris, N., & Tsolaki, M. (2012). On the comparison of VR-responses, as performance measures in prospective memory, with auditory P300 responses in MCI detection. *Studies in health technology and informatics*, *181*, 156–161. <https://bit.ly/3Nkh67m>
- Terwee, C., Prinsen, C., Chiarotto, A., de Vet, H., Bouter, L., Alonso, J., Westerman, M., Patrick, D., & Mokkink, L. (2018). *COSMIN manual for systematic reviews of PROMs* (1.^a ed.). Department of Epidemiology and Biostatistics. <https://bit.ly/3OkcfED>
- Tuena, C., Mancuso, V., Stramba-Badiale, C., Pedroli, E., Stramba-Badiale, M., Riva, G., & Repetto, C. (2021). Egocentric and Allocentric Spatial Memory in Mild Cognitive Impairment with Real-World and Virtual Navigation Tasks: A Systematic Review. *Journal of Alzheimer's Disease*, *79*(1), 95–116. <https://doi.org/h22c>
- Valladares-Rodríguez, S., Pérez-Rodríguez, R., Anido-Rifón, L., & Fernández-Iglesias, M. (2016). Trends on the application of serious games to neuropsychological evaluation: A scoping review. *Journal of Biomedical Informatics*, *64*, 296–319. <https://doi.org/gbw5mm>
- Vasilyeva, M., & Lourenco, S. F. (2012). Development of spatial cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *3*(3), 349–362. <https://doi.org/gmb94j>

Weniger, G., Ruhleder, M., Lange, C., Wolf, S., & Irle, E. (2011). Egocentric and allocentric memory as assessed by virtual reality in individuals with amnesic mild cognitive impairment. *Neuropsychologia*, *49*(3), 518–527. <https://doi.org/frbsn4>

Widmann, C. N., Beinhoff, U., & Riepe, M. W. (2012). Everyday memory deficits in very mild Alzheimer's disease. *Neurobiology of Aging*, *33*(2), 297–303. <https://doi.org/c76m4q>

World Health Organization. (2018). *Towards a dementia plan: a WHO guide*. World Health Organization. <https://bit.ly/3HOtrj4>

Anexos

Anexo 1

Características de los estudios

Referencia	Muestra		Instrumento	Constructo	Tecnología de desarrollo / Interfaz	Otras técnicas	Correlatos neuropsicológicos	Biomarcador
	n	Edad Media (DE) Subgrupo						
Allison et al. (2016)	12	73 (9) EA Preclínica	Laberintos virtuales no inmersivos: Laberinto radial de pasillos interconectados con puntos de referencia	NE	Laberintos diseñados en <i>WorldViz Vizard</i> y <i>Autodesk 3ds Max</i>	Nivel de A β 42 en LCR	SRT-Free Recall	N.R.
	42	69 (9) Controles sanos		Alocéntrica: Aprendizaje del lugar				
	15	77 (10) EA Sintomática (DCL y demencia moderada)		Egocéntrica: Aprendizaje por respuesta				

Allison et al. (2019)	30	73 (4)	Controles sanos	Tareas de navegación espacial: mapeo cognitivo, vinculación, toma de perspectiva visual.	NE Alocéntrica	Software <i>WorldViz Vizard and Autodesk 3ds Max.</i>	Nivel de A β 42 y pTau 182 en LCR	WAIS-III; Wechsler Memory Scale-III; Wechsler Memory Scale-R; Wechsler Memory Scale Associate; Learning task y Free Recall; Shipley Vocabulary Test; Auditory Consonant Trigrams y TMT B	N.R.
	60	67 (5)	Altos niveles de A β 42 y pTau 182 en LCR		Perspectiva visual	<i>Interfaz: Portátil Alienware, monitor de 17 pulgadas y joystick.</i>		Cuestionario de capacidades de navegación	
Bellasen et al. (2012)	20	28.2 (5.7)	Controles sanos	<i>Starmaze task:</i> Entorno de realidad virtual pentagonal.	Memoria espaciotemporal.	Entorno de realidad virtual <i>Starmaze</i> , diseñado con 3D <i>StudioMax</i> y <i>Virtools 3.5.</i>	N/R	MMSE; FCSRT; ROCF; FAB; CBT	N/R
	19	48.3 (6.8)	Controles sanos						
	24	72 (4.2)	Controles sanos						
	14	70,8 (8.8)	DCL-a						

	16	70 (9.0)	EA						
	11	68,2 (7.8)	DFT						
Caffò et al. (2012)	28	69.89(5.17)	DCL-a Ud	<i>Virtual Reorientation Test (VReoT):</i> Entorno de RV ambientada en una habitación cuadrada.	Reorientación espacial.	Entorno de Realidad virtual no inmersivo, presentado en un monitor de escritorio.	N/R	MMSE; ADL; CDR.	N/R
	23	73.91(4.72)	DCL-a Md						
	53	68.06(5.96)	Controles sanos						
Caffò et al. (2018)	203	68.6(7.1)	Controles sanos	<i>Virtual Reorientation Test (VReoT):</i> Entorno de RV ambientada en una habitación de dimensiones variables.	Reorientación espacial	Entorno no inmersivo de Realidad virtual, presentado en un monitor de escritorio.	N/R	CDR, MMSE, ADL, Digit Span Forward.	N/R
	14	70.9(8)	DCL-a Ud						
	22	73(4.7)	DCL-na Ud						
	16	74.9(8)	DCL-a Md						

	10	75.3(6.7)	DCL-na Md						
	21	74.3(6.6)	Prob. Dem.						
Coughlan et al. (2020)	33	64.38 (5.6)	Portadores e3e3	<i>Virtual Supermarket Task</i>	Memoria espacial	Ambiente virtual en un <i>iPad</i> de 9,7 pulgadas.	Mapa en papel Toma de muestras APOE 2, 3 y 4	CCI, ACEversión B (primera aplicación), ACE version C (en el test-retest), ROCF en la primera aplicación, TCFT, FMT	N/R
	31	63.74 (6.4)	Portadores e3e4	<i>Sea Hero Quest</i>					
Da Costa et al. (2021)	19	73 (6.50)	DCL	<i>SOIVET Maze</i>	Orientación espacial	Las tareas se desarrollaron en la plataforma <i>Unity</i> ®. Para su administración se utilizó:	N/R	ACE – R; MRMT; CBT; LT.	N/R

	29	70 (5.28)	Controles sanos	<i>SOIVET Route immediate & delayed</i>		- Un kit <i>Oculus Rift CV1</i> (pantalla para cabeza, dos controladores táctiles, dos sensores infrarrojos) - Un ordenador compatible con <i>Microsoft Windows 10</i>			
Davis y Sikorskii, (2020)	7	76.57 (5.03)	EA Leve	<i>Virtual reality simulation of a large senior residence</i>	Atención selectiva	Pantalla de 12 pies <i>Joystick</i>	Eye-tracking	N/R	N/R
	8	75 (1.20)	Controles sanos			Gafas de seguimiento <i>Applied Science Industries Mobile Eye-XG.</i>			
Kerbler et al. (2015)	22	74.7 (8.1)	EA	<i>Human Morris Water Maze</i>	NE alocéntrica NE egocéntrica	Ordenador con pantalla táctil y mouse	Espacio real: una arena rodeada de una cortina de terciopelo azul oscuro y manejado por	RAVLT; RAVLT 30-min Delayed Recall Trial; ECRT; ROCF; TMT - B; COWAT; DS forward &	RM de Cerebro basal anterior: hipocampo, protuberancia, materia gris y LCR
	27	71.4 (8.1)	DCL – a						

	17	70.71 (4.8)	Controles sanos				un programa informático	backward; TMT- A; BNT; MMSE.	
							Resonancia magnética		
Konishi et al. (2018)	15	65.6 (4.5)	Portadores e3e4	<i>The Concurrent Spatial Discrimination Learning Task (CSDLT)</i>	NE aloce3trica	<i>Unreal Tournament 2003 Editor.</i>	Resonancia Magnética.	MoCa; MMSE; RAVLT; ROCF; ST; DST.	RM del Núcleo caudado e hipocampo
	49	66.5 (4.6)	Controles sanos		NE egoc3trica		Control con cuatro botones para el movimiento.		
	21	75.9 (5.6)	EA						
Laczó et al. (2012)	10	77.3 (10.8)	DCL-ah	<i>Hidden Goal Task (HGT)</i>	NE aloce3trica	Versión computarizada de 2 dimensiones proyectada en un portátil de 17' sobre el cual el participante usaba el puntero del ratón para desplazarse.	<i>Hidden Goal Task</i> en espacio real.	N/R	N/R
	32	72.7 (9.2)	DCL – anh		NE egoc3trica				
	28	68.9 (7.2)	Controles sanos						

	33	74.4 (10.8)	DCL-a e4 -/-						
Laczó et al. (2014)	26	74.9 (7.3)	DCL-a e4 +/-	<i>Hidden Goal Task (HGT)</i>	NE allocéntrica NE egocéntrica	Versión computarizada de 2 dimensiones proyectada en un portátil de 17' con pantalla táctil.	Toma de muestra de genotipo APOE aislado de ADN sanguíneo.	MMSE, RAVLT, TMT, COWAT, DS- forward & backward, BVRT	RM del Volumen del Hipocampo
	15	71.7 (7.3)	DCL-a e4 +/+						
	20	70.8 (5.2)	Controles sanos						
Lee et al. (2014)	20	72.4 (5.6)	EA	<i>VRAM Task:</i> Laberinto radial virtual de seis brazos.	Memoria visoespacial de corto plazo	Computador de escritorio, entorno no inmersivo	N/R	Korean MMSE, CDR, ROCF, SSFB	N/R
	20	70.7 (5.0)	DCL-a						
	22	69.77 (4.60)	Controles sanos	VREAD: Entorno de RV experimental compuesto por tres módulos que ofrecen una puntuación general.	Memoria topográfica	Entorno no inmersivo, computador utilizando las flechas del teclado.	Cuestionario al finalizar la tarea para evaluar la usabilidad del entorno de realidad virtual.	GNT, CANTAB, MMSE, Word recall	N/R
Lesk et al. (2014)	9	72.67 (6.56)	DCL						

Migo et al. (2016)	10	70.3 (6.5)	Controles sanos	<i>Platform task</i>	NE alocéntrica NE egocéntrica	Realidad virtual proyectada en una pantalla <i>Perspex</i> . Se utiliza una <i>trackerball</i> para moverse en el entorno de RV	FRMI	WASI; CVLT; Logical Memory and Visual Reproduction Subtests; Wechsler-III; HBT; NART	FRMI del Hipocampo, áreas parahipocampales, núcleo caudado
	8	69.6 (5.8)	DCL-a						
Mohammedi et al. (2018)	30	69.8 (1.43)	Controles sanos	<i>Virtual Reality Navigation Task (VRNT):</i>	NE alocéntrica	Entornos virtuales en primera persona de 3D, controlado por <i>Joystick</i>	No reportan	RAVLT; ROCF; MMSE	N/R
	20	73.6 (2.47)	EA leve	<i>Virtual Neighborhood</i>					
	30	70.0 (1.63)	DCL-a ud	<i>Virtual Reality Navigation Task (VRNT):</i>					
	30	70.0 (1.68)	DCL-a md	<i>Virtual Maze</i>					
Morganti et al. (2013)	26	77.23 (5.25)	Controles sanos	<i>VR-Maze spatial task (VRMT)</i>	NE alocéntrica	Dispositivo de realidad virtual inmersivo manejado con <i>joypad</i>	Versión de lápiz y papel de la tarea para marcar la ruta	MMSE; CBT; TMT; ToL; BLOT.	N/R
	26	80.96 (6.3)	EA						

	26	77.23 (5.25)	Controles sanos						
	26	80.96 (6.3)	EA	<i>VR-Road Map task (VR-RMT)</i>	NE egocéntrica	Dispositivo de realidad virtual inmersivo manejado con <i>joypad</i>	Versión de lápiz y papel de la tarea para marcar la ruta		
	20	67.50 (7.13)	Controles sanos						
Parizkova et al. (2018)	28	74.46 (5.87)	DCL	Ve & Vysa	NE allocéntrica NE egocéntrica	Monitor de 24' a aprox. 50cm de la cara del evaluado, manejado con <i>joystick</i>	Análisis de neuroimagen funcional	RAVLT; ROCFT (Delayed Recall Copy condition); COWAT; SSFB; TMT A; MMSE.	MRI del Volumen del Hipocampo
	21	73.24 (6.93)	EA						
	32	71 (5.1)	DCL		Memoria Topográfica	Ordenador portátil Dell Latitude D820 con un procesador Intel Core Duo (2,6 GHz), 1 GB de RAM, una pantalla de 15,4".	Modified Way-finding Effectiveness Scale: cuestionario con preguntas sobre	MMSE; ACE-R; NART IQ; ROCF; RAVLT A30; PAL; VOSP; BORB; GNT; GDS; HADS; FMT (Place perception, PP,	N/R
Pengas et al. (2010)	22	68 (6.9)	Prob. EA	<i>Virtual Route Learning Test (VRLT)</i>	(Capacidad para navegar hacia nuevos lugares y recordar el				

	15	65 (7.7)	DS		camino hacia lugares conocidos)	<i>Joystick Logitech Attack</i>	desorientación espacial en participantes.	Place memory, MP)	
	<hr/>			<i>Heading-Orientation Test (HOT)</i>					
	35	69 (5.6)	Controles sanos						
	<hr/>								
	15	82.93 (5.61)	EA	Sala virtual con dos objetos (planta, piedra) y una flecha dibujada en el suelo apuntando hacia el norte que representa el inicio de la navegación.					
	<hr/>								
Serino et al. (2015)	15	77.53 (5.52)	DCL-a		NE aloccéntrica	Entorno de RV creado en NeuroVirtual 3D. Ordenador portátil (Acer Aspire con CPU Intel Core i5) con un procesador gráfico NVidia GeForce GT 540M, resolución 1024 x 768, Gamepad (Logitech Rumble F510) y un monitor.		CBT Span; CBT Supraspan; MRMT; Manikin's Test; JLO.	N/R
				Fase de codificación, memorizar entorno	NE egocéntrica				
	<hr/>								
	15	73.87 (7.38)	Controles sanos	Fase de recuperación, mapa real (tarea 1), contraposición (tarea 2)					
	<hr/>								

Serino et al. (2018)	15	86,73 (3.97)	Prob. EA	Fase de codificación, memorizar entorno	NE alocéntrica	Entorno de realidad virtual creado en NeuroVirtual 3D (extensión de Neuro VR), Ordenador portátil (Acer Aspire con CPU Intel Core i5) con un procesador gráfico NVidia GeForce GT 540M, resolución 1024 x 768, Gamepad (Logitech Rumble F510) y un monitor.	N/R	CBT Span; CBT Supraspan; MRMT; Manikin's Test; JLO	N/R
	15	83,53 (5.62)	Controles sanos	Fase de recuperación, mapa real					
Tarnanas et al. (2012)	25	78 (4.7)	DCL-a	VAP-M	Cognición espacial.	Entorno virtual desarrollado en Photo Modeler y Blaxxun. Ordenador Pentium (4MB RAM, Intel Quad Core, memoria 512MB, tarjetas gráficas NVIDIA).	Registro electrofisiológico ERP auditivo (Nihon Kohden - Neuropack M1 MEB-9200). Registrados	CDS; WLT.	N/R

	25	77,2 (4.9)	Controles sanos			Videobeam 1080p de Epson. Tarjeta de sonido Sound Blaster. <i>Joystick</i> inalámbrico Logitech. Mouse	durante un Electroencefalograma con 15 electrodos		
Weniger et al. (2011)	29	59 (9)	DCL-a	Virtual park. Entorno virtual de un parque con 9 puntos de intersección de dos direcciones y 11 callejones sin salida.	NE alocéntrica		Análisis volumétrico de regiones cerebrales de interés a través de Resonancia magnética nuclear (RMN) 3D.	MMSE; BDI; WAIS-R (Similarities); WCST; WMS-R (Verbal Memory, Visual Memory, Delayed Recall, Attention / Concentration); WAIS-R (Block Design; VOSP (dot counting, position discrimination, number location, cube analysis)	RMN del Hipocampo, precuneus, corteza parietal inferior, lesiones lacunares estriatales y WMHs
	29	59 (8)	Controles sanos	<i>Virtual Maze</i> . Entorno virtual de un laberinto con 6 puntos de intersección de dos direcciones y 7 callejones sin salida.	NE egocéntrica	Pantalla 21" y un <i>Joystick</i>			
Widman et al. (2012)	15	73 (9)	EA Leve	VE: Modelo virtual fotorrealista de la ciudad de Filadelfia	Memoria episódica verbal y espacial	Modelo fotorrealista proporcionado por GeoSim.	NR	CVLT; TMT; WMS (visual reproduction); Blockspan forward and backward	N/R

31	64 (9)	Controles sanos	Tarea 1 Recuerdo libre; Tarea 2 Dibujo de la ruta; Tarea 3 Colocación de tarjetas; Tarea 4 Reconocimiento ; Tarea 5 Nombreamiento.	<i>Videobeam</i> (Epson EMP 7600/5600) y un <i>Joystick</i>
----	--------	-----------------	--	---

Nota. A β 42: péptido beta-amiloide de 42 aminoácidos; ACE-R: addenbrooke's cognitive examination – revised; ACE versión B: addenbrooke's cognitive examination versión B; ACE versión C: addenbrooke's cognitive examination versión C; ADL: scores of the activities daily living; APOE: apolipoproteína; BDI: Beck's depression inventory; BLOT: Benton line orientation test; BNT: Boston naming test; BORB: Birmingham object recognition battery; BVRT: Benton's visual retention test; CBT: Corsi block test; CCI: cognitive change index; CDR: clinical dementia rating; CDS: cognitive difficulties scale; COWAT: controlled oral word association test; CVLT: California verbal learning test; DCL: deterioro cognitivo leve; DCL- a: DCL subtipo amnésico; DCL- a e4 -/-: DCL amnésico con gen APOE e4 negativo; DCLa e4 +/-: DCL amnésico con gen APOE e4 heterocigótico; DCLa e4 +/+ : DCL amnésico con gen APOE e4 homocigótico; DCL- ah: DCL subtipo amnésico hipocampal; DCL – anh: DCL subtipo amnésico no hipocampal; DCL – amd: DCL amnésico múltiple dominio; DCL – aud: DCL amnésico único dominio; DCL – na: DCL no amnésico; DCL – naud: DCL no amnésico único dominio; DFT: demencia frontotemporal; DFT-vc: demencia frontotemporal variante conductual; DS: demencia senil; DS- forward & backward: digit span (forward & backward); DST: digit symbol test; EA: enfermedad de Alzheimer; EA leve: enfermedad de Alzheimer Leve; EA preclínica: enfermedad de Alzheimer preclínica; EA sintomática: etapa sintomática sin información de la gravedad o los autores combinan grupos incluyendo etapa prodrómica y DCL; ECRT: enhanced cued recall test; FAB: frontal assessment battery; FCSRT: free and cued selective reminding test; FMT: four mountain test; FRMI: análisis de neuroimagen funcional; GDS:geriatric depression scale; GNT: graded naming test; HADS: hospital anxiety and depression scale; HBT: Hayling and Brixton tasks; JLO: judgment of line orientation; LCR: líquido cefalorraquídeo; LT: london tower; MMSE: mini mental state examination; MoCA: Montreal cognitive assessment; MRMT: money road-map test of direction sense; NART: national adult reading test; NE: navegación espacial; PAL: paired-associates learning test; Portador e3e3: portador gen APOE homocigótico; Portador e3e4: portador gen APOE heterocigótico; Prob. Dem.: probable demencia; Prob. EA: probable EA; RAVLT trials 1-5: Rey auditory verbal learning test, intentos 1-5; RAVLT 30-min Delayed Recall Trial: Rey auditory verbal Learning Test 30-min Delayed Recall Trial; RM: resonancia magnética; ROCF: Rey osterrieth complex figure; SRT- Free Recall: selective reminding test, subprueba Free Recall; p. Tau 182: proteína Tau; SSFB: spatial span forward and backward tasks; ST: Stroop Test; TCFT: Taylor complex figure task; TMT- A: trial making test- A; TMT- B: trial making test- B; VOSP: visual object and space perception battery; WAIS-III: Wechsler adult Intelligence scale- III; WASI: Wechsler abbreviated scale of intelligence; WCST: Wisconsin card sorting test; WLT: word learning test; WMS-R: Wechsler memory scale-revised.