



Diversitas: Perspectivas en Psicología  
ISSN: 1794-9998  
ISSN: 2256-3067  
revistadiversitas@usantotomas.edu.co  
Universidad Santo Tomás  
Colombia

Brandwayn, Natalia; Restrepo, David; Macías-Bohórquez,  
Ricardo; López-López, Wilson; Acevedo-Triana, César A.  
Comparación del desempeño de jóvenes en pruebas  
neuropsicológicas en formato lápiz y papel e informatizadas  
Diversitas: Perspectivas en Psicología, vol. 16, núm. 2, 2020, Julio-, pp. 365-388  
Universidad Santo Tomás  
Bogotá, Colombia

DOI: <https://doi.org/10.15332/22563067.6319>

Disponible en: <https://www.redalyc.org/articulo.oa?id=67968203012>

- ▶ [Cómo citar el artículo](#)
- ▶ [Número completo](#)
- ▶ [Más información del artículo](#)
- ▶ [Página de la revista en redalyc.org](#)

[redalyc.org](https://www.redalyc.org)

Sistema de Información Científica Redalyc  
Red de Revistas Científicas de América Latina y el Caribe, España y Portugal  
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso  
abierto

# Comparación del desempeño de jóvenes en pruebas neuropsicológicas en formato lápiz y papel e informatizadas\*

## Comparison of the Performance of Young Adults in Neuropsychological Tests in Paper-and-Pencil and Computerized Formats

**Natalia Brandwayn**

ORCID: 0000-0001-5687-1778

Pontificia Universidad Javeriana

**David Restrepo**

ORCID: 0000-0001-7723-6503

Universität Ulm, Alemania

**Ricardo Macías-Bohórquez**

ORCID: 0000-0002-0217-8756

**Wilson López-López**

ORCID: 0000-0002-2964-0402

Pontificia Universidad Javeriana

**César A. Acevedo-Triana\*\***

ORCID: 0000-0002-1296-9957

Universidad Pedagógica y Tecnológica de Colombia

Recibido: 13 de enero de 2020

Revisado: 17 de marzo de 2020

Aceptado: 22 de mayo de 2020

### Resumen

Los campos de la aplicación de pruebas psicológicas y neuropsicológicas se han visto impactos por la tecnología específicamente con las pruebas informatizadas. Una de las preguntas frecuentes en versiones papel y lápiz comparadas con una versión informatizada de la misma prueba, es si éstas son equivalentes. El objetivo de este estudio fue comparar el desempeño de pruebas neuropsicológicas en formato lápiz y papel y formato computarizado. Participaron 41 estudiantes en dos aplicaciones de las mismas pruebas en las dos versiones. Se encontró que los desempeños fueron equivalentes en la mayoría de las pruebas. Sin embargo, existen variaciones en las modalidades dependiendo de la ejecución de la prueba y con alta variabilidad entre sujetos. Esto sugiere que la diferencia en las puntuaciones podría deberse a la diferencia en los procesos cognitivos subyacentes a la ejecución de las dos modalidades.

**Palabras clave:** Pruebas neuropsicológicas, lápiz y papel, pruebas informatizadas

\* Artículo de investigación. Esta publicación es resultado del proyecto ID 00006585, financiado por la Vicerrectoría de Investigación de la Pontificia Universidad Javeriana, Bogotá. Citar como: Brandwayn, N., Restrepo, D., Macías-Bohórquez, R., López-López, W. y Acevedo-Triana, C. (2020). Comparación del desempeño de jóvenes en pruebas neuropsicológicas en formato lápiz y papel e informatizadas. *Diversitas: Perspectivas en Psicología*, 16(2), 365-386. <https://doi.org/10.15332/22563067.6319>

\*\* Autor de correspondencia: César Acevedo-Triana. Escuela de Psicología, Facultad de Ciencias de la Salud. Universidad Pedagógica y Tecnológica de Colombia. Dirección: Calle 24 # 5-63 Antiguo Hospital San Rafael de Tunja, Tunja, Colombia. Correo electrónico: cesar.acevedo02@uptc.edu.co

## Abstract

The application of psychological and neuropsychological tests has been impacted by technology, specifically by computerized testing. One of the frequently asked questions is whether the version in pencil and paper and the computerized one can be considered equivalent tests. This study aimed to compare the performance on neuropsychological tests in pencil-and-paper and in computerized format. Forty-one students participated in two applications of the same tests in the two versions. It was found that the performances were equivalent in most of the tests. However, there are variations in the modalities depending on the execution of the test, and high intersubject variability. Results suggest that the difference in scores may be due to the difference in the underlying cognitive processes during the performance of the two modalities.

**Keywords:** neuropsychological test, paper-and-pencil format, computerized test.

## Introducción

La incorporación de la tecnología a la investigación en psicología y ciencias del comportamiento ha traído como consecuencia la adaptación de los instrumentos y métodos a versiones computacionales e instrumentos tecnológicos avanzados (Millsap, 2000; Piper *et al.*, 2015). En psicología esto ha significado por ejemplo, establecer nuevas formas de psicoterapia (Richards, 2013), realización de experimentos e investigaciones *on-line* (Garaizar, Vadillo, López-de-Ipiña y Matute, 2012), aplicación de pruebas psicométricas para selección de personal en versiones *on-line* (Joubert y Kriek, 2009) e incluso intentar mejorar el desempeño cognitivo mediante “entrenamiento asistido por computador” (Lampit, Hallock y Valenzuela, 2014; Prada Sarmiento, Pineda Garzón, Mejía Orduz y Conde Cotés, 2010; Toril, Reales, Mayas y Ballesteros, 2016; Vianin, 2016). En el campo de la psicometría las pruebas computarizadas (o informatizadas) (Olea, Abad y Barrada, 2010) permitieron un avance importante en el desarrollo de nuevas pruebas y tareas, mejoramiento en la velocidad de aplicación, el aumento de la información recolectada, la puntuación inmediata, la seguridad, la cobertura y el control de fuentes de error con respecto a las tradicionales aplicaciones de lápiz y papel (Bugbee, 1996; Jeong, 2014; Joubert y Kriek, 2009). Así mismo, ha

habido un esfuerzo para mantener estándares de validez y confiabilidad de las pruebas psicológicas y psicométricas, también importantes avances por adaptar, validar y normalizar las pruebas en sus diferentes versiones (Ison y Carrada, 2010; Leal-Soto & Alonso-Tapia, 2017).

Como un ejemplo de lo anterior, se ha mostrado que algunas pruebas en formato de lápiz y papel no logran medir adecuadamente los tiempos de reacción o sutiles habilidades de desempeño motor a comparación de las pruebas y simulaciones computarizadas (Bieri *et al.*, 2014). Este tipo de cuestionamientos sugiere indagar por las características de validez de las pruebas bajo diferentes formatos. Además, no en vano se encuentra que a la fecha son cientos las pruebas que han sido construidas en versiones informatizadas, dentro de varias razones metodológicas también se encuentran razones de facilidad en su uso, acceso abierto, replicabilidad de las investigaciones y limitaciones en el acceso a las mismas pruebas pero en formato de lápiz y papel y a nivel comercial (Mueller y Piper, 2014; Open Science Collaboration, 2012, 2015; Piper, Mueller, Talebzadeh y Ki, 2016; Reips, 2001). A este esfuerzo por mejorar la aplicación de las pruebas, también se ha generado el concepto de Psicología 2.0 para denotar todas aquellas herramientas que adoptan la tecnología y actualizan sus

procedimientos e instrumentos, así como el abordaje de cantidades de datos en la era de los grandes volúmenes de datos (*big data*) (Armayones Ruiz *et al.*, 2015; Markowitz, Błaszkiwicz, Montag, Switala y Schlaepfer, 2014).

Una de las preguntas clave con respecto a esta implementación es si existe una equivalencia entre las pruebas en versión de lápiz y papel e informatizadas (Bugbee, 1996; Noyes y Garland, 2008; Olea, Revuelta, Ximénez y Abad, 2000). Se han reportado estudios que identifican la necesidad de evaluación de esta característica de equivalencia, desde pruebas de personalidad hasta pruebas de habilidad o destreza (Baumer, Roded y Gafni, 2009; Collerton *et al.*, 2007; Holländare, Andersson y Engström, 2010). En algunos de estos estudios se ha encontrado una diferencia en la puntuación de los test que ha sido atribuida a los aspectos formales de la presentación (es decir, las instrucciones, la acción a realizar durante la prueba, la solución de inquietudes o componentes gráficos y visuales) (Baumer *et al.*, 2009; Noyes y Garland, 2008) o aspectos relacionados con la familiaridad tecnológica de los sujetos (Baumer *et al.*, 2009; Joubert y Kriek, 2009). Sin embargo, no todas las comparaciones resultan con puntuaciones diferentes, por ejemplo, algunas pruebas de medición de conocimiento o habilidades comunicativas muestran equivalencia a lo largo de comparar diferentes individuos por grados escolares de secundaria y universitarios (Baumer *et al.*, 2009; Millsap, 2000; Vallejo, Jordán, Díaz, Comeche y Ortega, 2007; Wang, Jiao, Young, Brooks y Olson, 2007). Incluso existen reportes de la medición de características cognitivas en adultos mayores que podrían ser más sensibles que las realizadas en formatos tradicionales (Bieri *et al.*, 2014; Collerton *et al.*, 2007; Mead y Drasgow, 1993).

Uno de los principales problemas de la equivalencia se ha relacionado con que el proceso (cognitivo o motor) que el individuo realiza para la solución o respuesta de la prueba difiere entre las versiones, por lo que pruebas incluso con una gran similitud podrían no tener una equivalencia psicométrica (Lozzia *et al.*, 2009; Noyes y Garland, 2008). Una explicación tentativa es suponer que las dos versiones requieren de procesos o funciones diferentes a

las que estarían involucradas en cada modalidad, por lo que la comparación de la prueba *per se* no es posible, antes de determinar la validez de cada versión (Muñiz, Hernández y Ponsoda, 2015).

En el campo de la neuropsicología, la evaluación de las características psicométricas de las pruebas, indican que existen variables de los instrumentos que requieren ser evaluadas, antes de suponer su funcionamiento estandarizado (García *et al.*, 2018). Esto sumado a la tendencia a aumentar el uso de las pruebas informatizadas debido a las ventajas relacionadas con la adaptación de ítems a diferentes condiciones y grupos de edad, medición de tiempos en escala más precisas, calificación automática y la versatilidad en la aplicación de múltiples pruebas de forma sencilla en un solo ordenador (Collerton *et al.*, 2007; Hurtado-Parrado *et al.*, 2016; Montoya-Arenas, Trujillo-Orrego y Pineda-Salazar, 2010; Ployhart, Weekley, Holtz y Kemp, 2003). Adicional a estas ventajas, algunos estudios han planteado que este tipo de implementaciones están relacionadas con favorecer y facilitar los estudios longitudinales, disminuyendo la deserción y aumentando la accesibilidad (Collerton *et al.*, 2007). También se han planteado resultados importantes evaluando la predicción de pruebas en desempeños cognitivos con resultados positivos (Ployhart *et al.*, 2003). Uno de los resultados del uso de imágenes cerebrales en pacientes que están desarrollando pruebas neuropsicológicas es el establecimiento de estructuras cerebrales implicadas durante la realización de las tareas (Alvarez & Emory, 2006; Johnson, Saykin, Flashman, McAllister y Sparling, 2001; Moll, Oliveira-Souza, Moll, Bramati y Andreiuolo, 2002; Rubia *et al.*, 2001; Zakzanis, Mraz y Graham, 2005). Ahora bien, la aplicación de dos modalidades de la misma prueba puede responder a estructuras diferentes por lo que la evaluación de las modalidades podría considerarse una forma de validación de los constructos.

Así, se ha señalado que la equivalencia de dos pruebas está dada por el cumplimiento de condiciones como la fiabilidad comparable de las pruebas; correlación entre los estimados de fiabilidad, correlación con otros test y subpruebas o criterios externos y medidas de tendencia central y dispersión similares (Joubert y Kriek, 2009). Con base en

lo anterior los objetivos de este artículo fueron: 1) Comparar el desempeño de los individuos en las pruebas de funciones ejecutivas en las modalidades de lápiz y papel e informatizadas; 2) Es establecer la equivalencia por medio de las correlaciones entre las pruebas aplicadas.

## Método

### Participantes

Participaron 41 estudiantes jóvenes (edad  $M = 17.95$ ;  $SD = 0.86$ ) de pregrado de psicología de la Pontificia Universidad Javeriana, Bogotá, Colombia. Fueron seleccionados por conveniencia e invitados por correo electrónico. Los sujetos no recibieron remuneración por la participación en el estudio, se controló que no conocieran las pruebas ni que tuvieran formación en áreas de psicometría.

### Instrumentos

#### Pruebas neuropsicológicas

Para la evaluación de la función de atención y velocidad de procesamiento se usó del test d2 (Brickenkamp y Zillmer, 2002), el cual permitió medir la velocidad de procesamiento de estímulos visuales en una tarea de cancelación de ítems y ofreció una medida de la capacidad de concentración del sujeto. La prueba consistió en tachar las letras “d” con 2 líneas a su alrededor en una hoja llena de elementos distractores. Se reportó la efectividad total en la prueba y el índice de concentración indicado como una medida de la atención (Budde, Voelcker-Rehage, Pietraszyk-Kendziorra, Ribeiro y Tidow, 2008; Jiménez *et al.*, 2012). Para la versión informatizada se utilizó una versión de la prueba programada en el software SuperLab (Cedrus, Inc., Versión 4.5).

La memoria de trabajo fue evaluada por la subprueba de Dígitos del Wechsler Adults Intelligence Scale (WAIS) en su versión IV, la cual, a partir de la repetición de ítems en orden directo, inverso y en secuencia evaluó la atención y la memoria de trabajo verbal (Pisoni y Cleary, 2003; Rosenthal, Riccio,

Gsanger y Jarratt, 2006; Wechsler, 1997). Para evaluar memoria de trabajo visual se utilizó la subprueba de localización espacial del Wechsler Memory Scale (WMS) en su versión III, que es una prueba equivalente a la versión visual de la prueba de dígitos del WAIS (Ardila, Rosselli y Strumwasser, 2009; Millis, Malina, Bowers y Ricker, 2010). Finalmente se aplicó la subprueba Números y Letras en el que se le pide a los sujetos ordenar números y letras según la instrucción que han sido presentados en orden, se piden en forma directa o inversa para evaluar memoria de trabajo verbal (Cadavid-Ruiz, Del Rio, Egido y Galindo, 2016; Cadavid y Del Rio, 2012). Las pruebas informatizadas para esta función fueron programadas en el software SuperLab (Cedrus, Inc., Versión 4.5) siguiendo la misma distribución de tareas que la versión en físico.

Para la función de planificación se usó el test Tower of London (ToL) en el cual los sujetos tenían que planear en orden el movimiento de unos anillos de colores para completar un patrón, se midió la cantidad de movimientos y el tiempo promedio de ejecución (Dockery, Hueckel-Weng, Birbaumer y Plewnia, 2009; Piper *et al.*, 2015). En la versión en papel y lápiz se utilizó la prueba de las Anillas realizada por Portellano, Díez y García (2007). Para la versión informatizada se utilizó el test automatizado del software de The Psychology Experiment Building Language (PEBL) (Mueller y Piper, 2014; Piper *et al.*, 2012, 2015).

Para la evaluación de la función de rotación mental y habilidades visuoconstructivas se usó la subprueba de “Cubos” del Wechsler Adults Intelligence Scale (WAIS) (Wechsler, 1997); que estuvo compuesta por nueve cubos de madera pintados de color rojo y blanco (dos caras rojas, dos caras blancas y dos caras divididas en diagonal, con una mitad roja y la otra blanca) y un cuadernillo con nueve diseños bidimensionales que el sujeto debía reproducir con los cubos. La prueba aumentó de complejidad a medida de su avance y evaluó variables de tiempo y ejecución (Rosas, 2013; Savage, 2016). Para la versión informatizada se utilizó una versión de la prueba programada en el software SuperLab (Cedrus, Inc., Versión 4.5).

Para la evaluación del efecto de re-aplicación y el reaprendizaje se hizo uso de las pruebas de Cubos del WAIS III, la prueba de Automatic Mirror Trace

donde los sujetos debían realizar la forma de una estrella pero con información visual obtenida desde un espejo y por lo tanto de forma invertida (Cellini, 2016) y el Test de rapidez motora (Clavijas) donde se pidió a los sujetos que completaran tareas de destreza motora fina de forma unimanual y bimanual (Roy y Park, 2010).

### Procedimiento

Se realizó una primera aplicación piloto a 5 estudiantes de pregrado de la Facultad de Psicología con el ánimo de evaluar la claridad de las instrucciones, la comprensión de las tareas y el tiempo de aplicación. Luego de este procedimiento se analizaron los resultados y se ajustaron los términos desconocidos y claridad en las aplicaciones. No se realizaron cambios sustanciales ni en la programación de las pruebas ni en las funciones evaluadas. Cada una de las pruebas computarizadas fue evaluada por un experto que emitió su concepto sobre la validez de las tareas.

Para la aplicación final, los participantes aceptaron y firmaron estar de acuerdo con la participación

en el estudio el día 1 de aplicación de pruebas, se recolectaron algunos datos sociodemográficos. El orden de aplicación de la pruebas fue contrabalanceado (virtual y físico vs físico y virtual) entre los sujetos. Esta aplicación tuvo una duración promedio de una hora y cuarenta minutos entre las 7 am y 5 pm. Luego de una semana se realizó la aplicación de las mismas pruebas realizadas en la aplicación inicial pero en la modalidad contraria a la inicialmente realizada. Todos los procedimientos y protocolos fueron realizados cumpliendo con la Resolución 8430/1993 de la ley Colombiana y respetando los principios internacionales de Helsinki para la investigación en humanos. Los resultados fueron analizados mediante pruebas *t* para muestras relacionadas para cada una de las modalidades y las pruebas y una prueba de ANOVA de una vía tomando como factor el orden de aplicación de las pruebas en SPSS.

### Resultados

En la tabla 1 se reportan los valores medios y desviación estándar de las pruebas realizadas.

Tabla 1  
Estadísticos descriptivos de las pruebas aplicadas

Prueba/Subprueba	Media	Error de la Media	Desviación Estándar
<i>WAIS-IV - Subprueba Números y Letras</i>			
Manual total	21.19	0.47	3.03
Manual span	5.80	0.16	1.07
Virtual total	17.65	0.54	3.46
Virtual span	7.31	0.19	1.27
<i>Anillas</i>			
Tiempo	302.58	20.67	132.44
Movimientos	187.58	1.17	7.49
ToL PEBL tiempo	383.43	12.26	78.53
ToL PEBL movimientos	170.85	4.56	29.25
<i>d2 Test</i>			
Total de respuestas	443.73	10.14	64.93
Total de aciertos	171.00	4.31	27.60
Errores omisión	15.63	2.26	14.50

Prueba/Subprueba	Media	Error de la Media	Desviación Estándar
Errores comisión	0.80	0.18	1.20
Errores totales	427.29	9.69	62.07
Concentración	155.36	5.14	32.96
Virtual-Total de respuestas	344.22	12.20	78.16
Virtual-Total de Aciertos	134.43	7.81	50.06
Virtual-Errores omisión	16.04	4.61	29.54
Virtual-Errores comisión	23.36	6.21	39.80
Virtual-Errores totales	304.80	14.01	89.75
Virtual-Concentración	118.39	10.79	69.15
<b>WMS III- Subprueba Cubos</b>			
Manual directa	10.58	0.30	1.94
Manual directa span	6.87	0.16	1.0
Manual inversa	9.34	0.27	1.7
Manual inversa span	6.17	0.15	0.99
Manual total acumulado	19.92	0.48	3.10
Virtual directa	6.09	0.39	2.49
Virtual directa span	4.58	0.25	1.61
Virtual inversa	6.51	0.31	1.98
Virtual inversa span	5.04	0.20	1.28
Virtual total acumulado	12.61	0.51	3.27
<b>WAIS-IV - Subprueba Dígitos</b>			
Manual directos	9.78	0.29	1.90
Manual span directos	6.56	0.18	1.18
Manual inversa	9.04	0.37	2.40
Manual span inversos	5.07	0.23	1.50
Manual secuencia	9.63	0.34	2.22
Manual secuencia span	6.41	0.20	1.34
Manual total	28.46	0.80	5.16
Virtual directa	11.39	0.32	2.10
Virtual span	6.29	0.22	1.41
Virtual inversa	11.17	0.38	2.43
Virtual span inverso	5.70	0.22	1.45
Virtual secuencia	12.12	0.39	2.54
Virtual secuencia span	6.97	0.21	1.40
Virtual total	34.68	0.81	5.20

Fuente: elaboración propia.

## Atención

Para analizar la atención se utilizó la prueba d2 y la subprueba retención de dígitos. Se realizó un ANOVA de un factor, siendo el factor el orden de aplicación de las pruebas en sus diferentes modalidades (físico vs informatizado). Las variables que mostraron ser afectadas por el orden de la aplicación fueron únicamente de la prueba d2 en su formato virtual, siendo las variables: total de respuestas ( $F(1, 40) = 4.849; p = 0.034; \eta^2 = 0.11$ ), total de aciertos

( $F(1, 40) = 9.848; p = 0.003; \eta^2 = 0.201$ ), total de errores ( $F(1,40) = 6.654; p = 0.01; \eta^2 = 0.14$ ), índice de concentración ( $F(1, 40) = 7.011; p = 0.01; \eta^2 = 0.15$ ) e índice de variación ( $F(1, 40) = 35.02; p < 0.001; \eta^2 = 0.47$ ) (ver tabla 2). En este caso se observó un mejor rendimiento en los participantes que tuvieron el orden de aplicación primero en la modalidad a lápiz y papel luego modalidad virtual, no se observa este fenómeno en individuos con el orden contrario (figura 1 y 3).

Tabla 2.

Análisis de varianza para la prueba de atención (d2 test) comparando el tipo de aplicación manual versus virtual (grupo 1) y virtual versus manual (grupo 2).

Prueba-d2 test	F	p	$\eta^2$
Total de respuestas - manual	0.003	0.960	<0.001
Total de aciertos - manual	0.002	0.964	<0.001
Errores omisión-manual	0.228	0.636	0.006
Errores comisión-manual	1.198	0.280	0.029
Errores totales - manual	0.006	0.937	<0.001
Concentración - manual	0.062	0.805	0.001
Total de respuestas - virtual	4.849	0.034	0.11
Total de aciertos - virtual	9.848	0.003	0.201
Errores omisión-virtual	0.933	0.340	0.023
Errores comisión-virtual	0.350	0.558	0.008
Errores Totales - virtual	6.654	0.014	0.145

Fuente: elaboración propia.



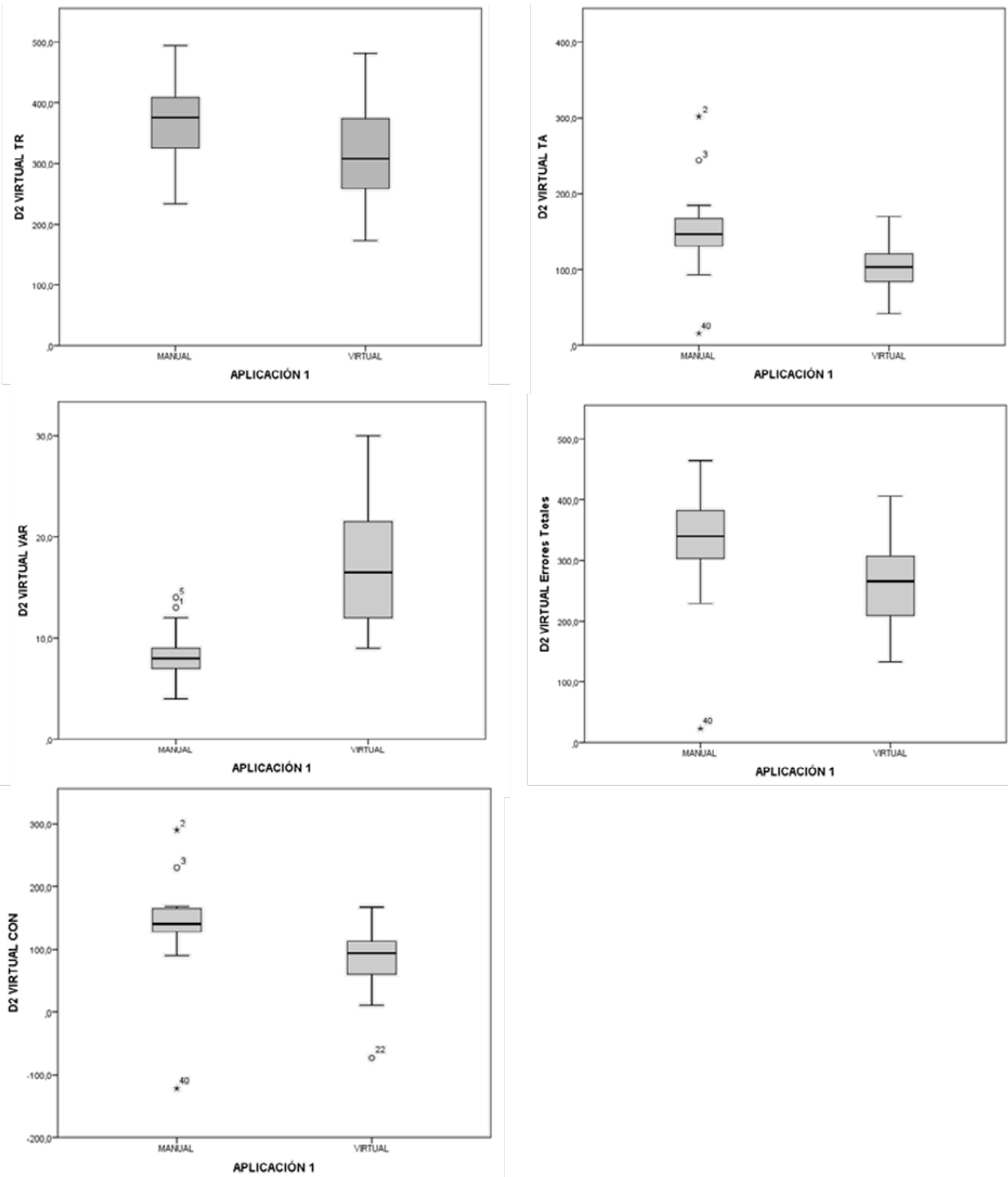


Figura 1. Diagramas de cajas y bigotes de variables que mostraron ser afectadas por el orden de la aplicación.

Fuente: elaboración propia

Así mismo, se realizó una prueba *t* para comparar las medias de las puntuaciones directas de la subprueba de dígitos. Se encontró que la puntuación directa tenía diferencias significativas, ( $t=-4.06$ ;  $p < .01$ ; *ic* 95% [-2.41, -0.81]) (tabla 3). En cuanto a la variable de span atencional directo, el cual hace

referencia a la cantidad de estímulos que lograron retener los sujetos, no presenta diferencia ( $t = 1.05$ ;  $p = 0.298$ ; *ic* 95% [-0.24, 0.78]) (tabla 3). Se presentó un aumento de errores en la modalidad manual pero una similitud en los alcances de las dos modalidades (figura 2).

Tabla 3.  
Diferencia entre los grupos en las aplicaciones de la prueba de dígitos mediante la prueba de diferencia de medias para muestras emparejadas

Dígitos	M	DE	SEM	95% de <i>ic</i> de la diferencia		t	Valor p
				Inferior	Superior		
Puntuación directa	-1.60	2.53	0.39	-2.41	-0.80	-4.06	< 0.001
Puntuación inversa	-2.12	2.66	0.41	-2.96	-1.28	-5.09	< 0.001
Secuencias	-2.48	2.73	0.42	-3.35	-1.62	-5.81	< 0.001
Total	-6.21	4.88	0.76	-7.76	-4.67	-8.15	< 0.001

Fuente: elaboración propia.

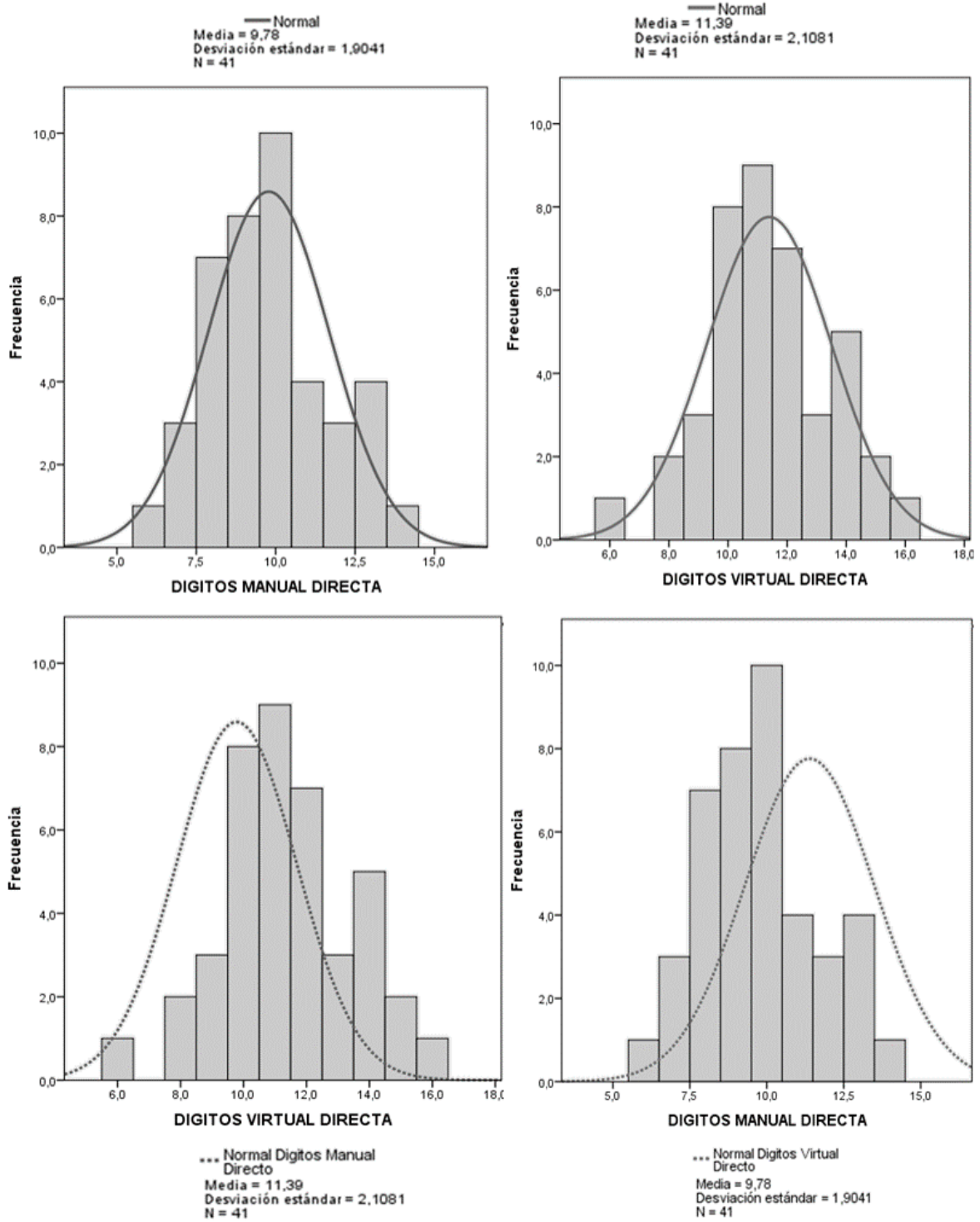


Figura 2. Histogramas de variables que mostraron diferencias significativas por la modalidad de aplicación.

Fuente: elaboración propia.

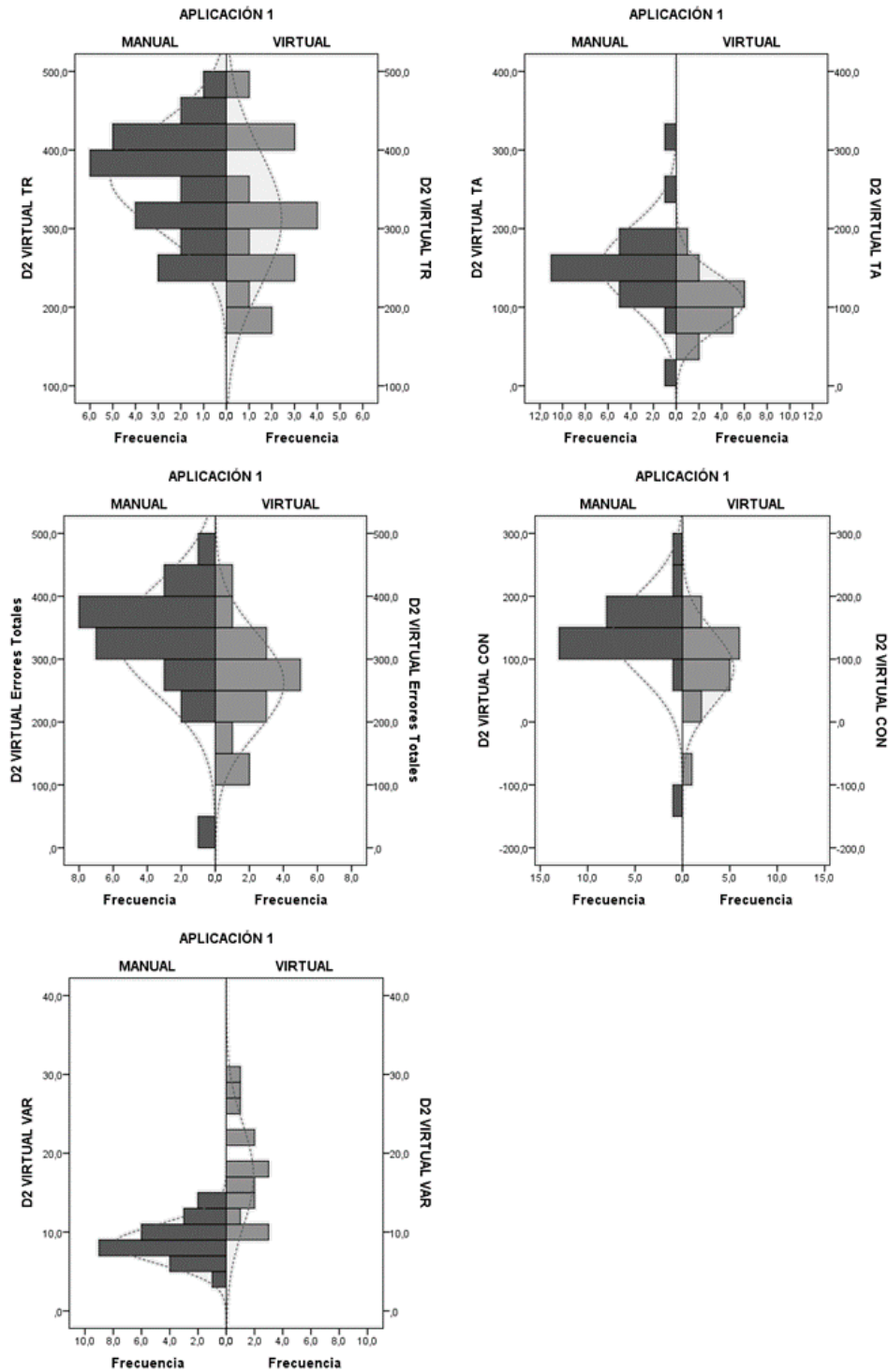


Figura 3. Histogramas comparados de variables de la prueba d2 que mostraron ser afectadas respecto al orden de la aplicación

Fuente: elaboración propia.

### Memoria de trabajo

Para analizar la memoria de trabajo se utilizaron las subprueba de números y letras, retención de dígitos inversos, retención de dígitos en secuencia y localización espacial inversa. A continuación se presentan los datos que mostraron una diferencia significativa entre las aplicaciones. En cuanto al puntaje directo obtenido en la prueba números y letras, se encontró una diferencia entre las dos versiones en el puntaje directo ( $t = -5.585$ ;

$p < 0.001$ ;  $ic\ 95\% [-2.06, -0.96]$ ), teniendo en la modalidad manual un promedio de 21.2, mientras que la modalidad informatizada el promedio fue de 17.2. Pese a esto, el span atencional fue mayor en el formato virtual ( $M = 7.32$ ) con respecto a la versión en lápiz y papel ( $M = 5.8$ ) (figura 4), lo que muestra en la modalidad virtual una mayor cantidad de estímulos que se pueden recordar, pero una mayor posibilidad de cometer un error en la prueba (tabla 4).

Tabla 4. Prueba t para muestras relacionadas para la subprueba números y letras.

Números y letras	M	DE	SEM	95% de ic de la diferencia		t	Valor p
				Inferior	Superior		
Total	3.53	4.02	0.62	2.26	4.81	5.62	< 0.001
Total CNL	-1.51	1.73	0.27	-2.05	-0.96	-5.58	< 0.001

Fuente: elaboración propia.

### Planificación

Para evaluar la planificación se compararon las pruebas ToL en PEBL y la prueba de anillas. El desempeño en el componente de “movimientos” se encontró cercano en las dos versiones y sin diferencia estadística ( $M_{(manual)} = 187.58$ ; PEBL:  $M_{(informatizada)} = 170.85$ ). Sin embargo, las desviaciones estándar

del componente manual y virtual ( $DE_{(manual)} = 7.5$  y  $DE_{(informatizada)} = 29.25$ ) muestran que la desviación estándar de la modalidad informatizada contiene 4 desviaciones estándar de la modalidad manual presentando una diferencia significativa ( $t = 3.55$ ;  $p = 0.001$ ;  $ic\ 95\% [7.22, 26.24]$ ) (tabla 5 y figura 6).

Tabla 5. Prueba t para muestras relacionadas de la prueba anillas y su versión virtual en PEBL.

Pruebas Anillas/ToL	M	DE	SEM	95% de ic de la diferencia		t	Valor p
				Inferior	Superior		
Tiempo	-80.85	132.65	20.71	-122.72	-38.98	-3.9	< 0.001
Movimientos	16.73	30.13	4.70	7.21	26.24	3.55	0.001

Fuente: elaboración propia.

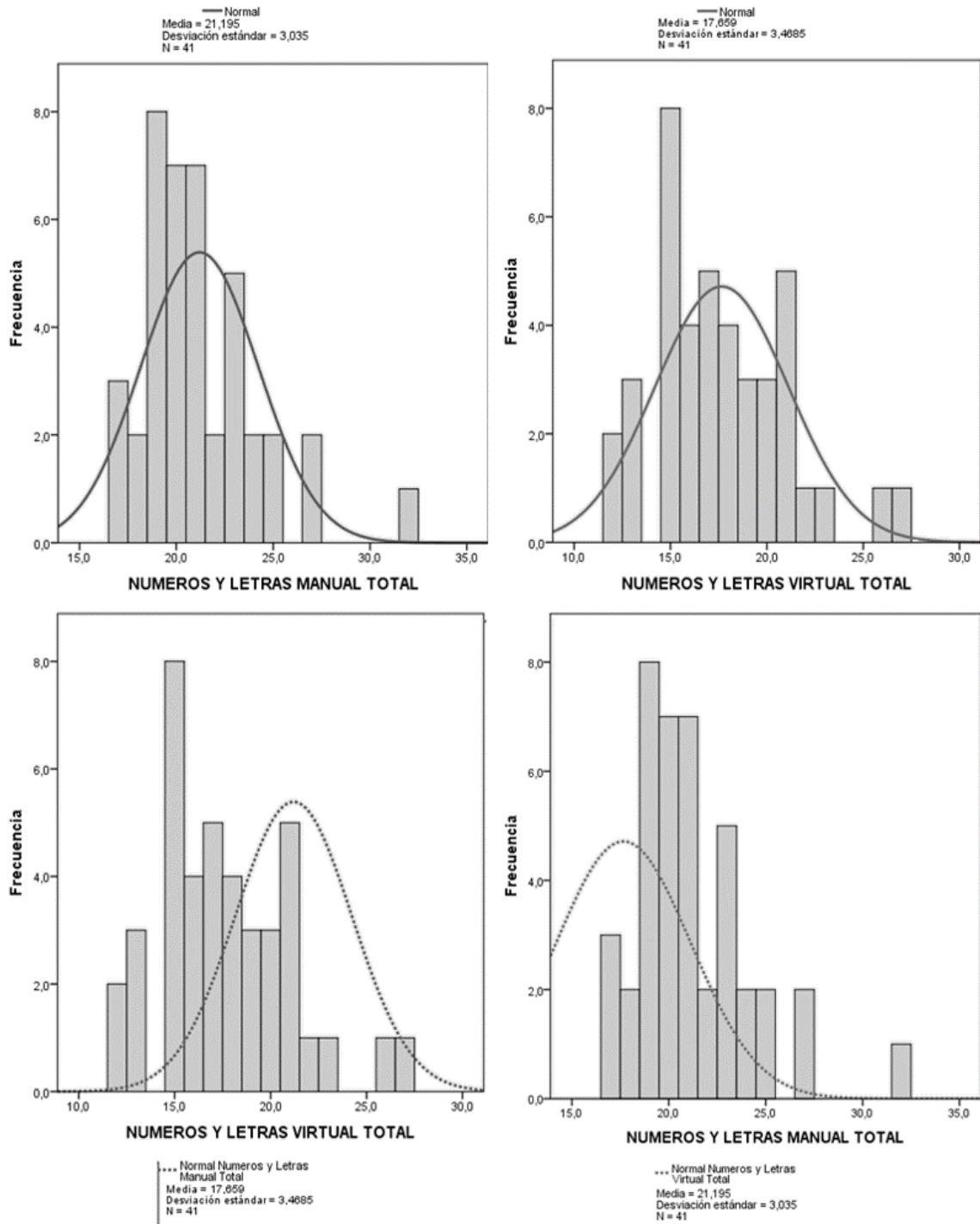


Figura 4. Histogramas que muestran las diferencias entre los puntajes de la prueba números y letras.

Fuente: elaboración propia.

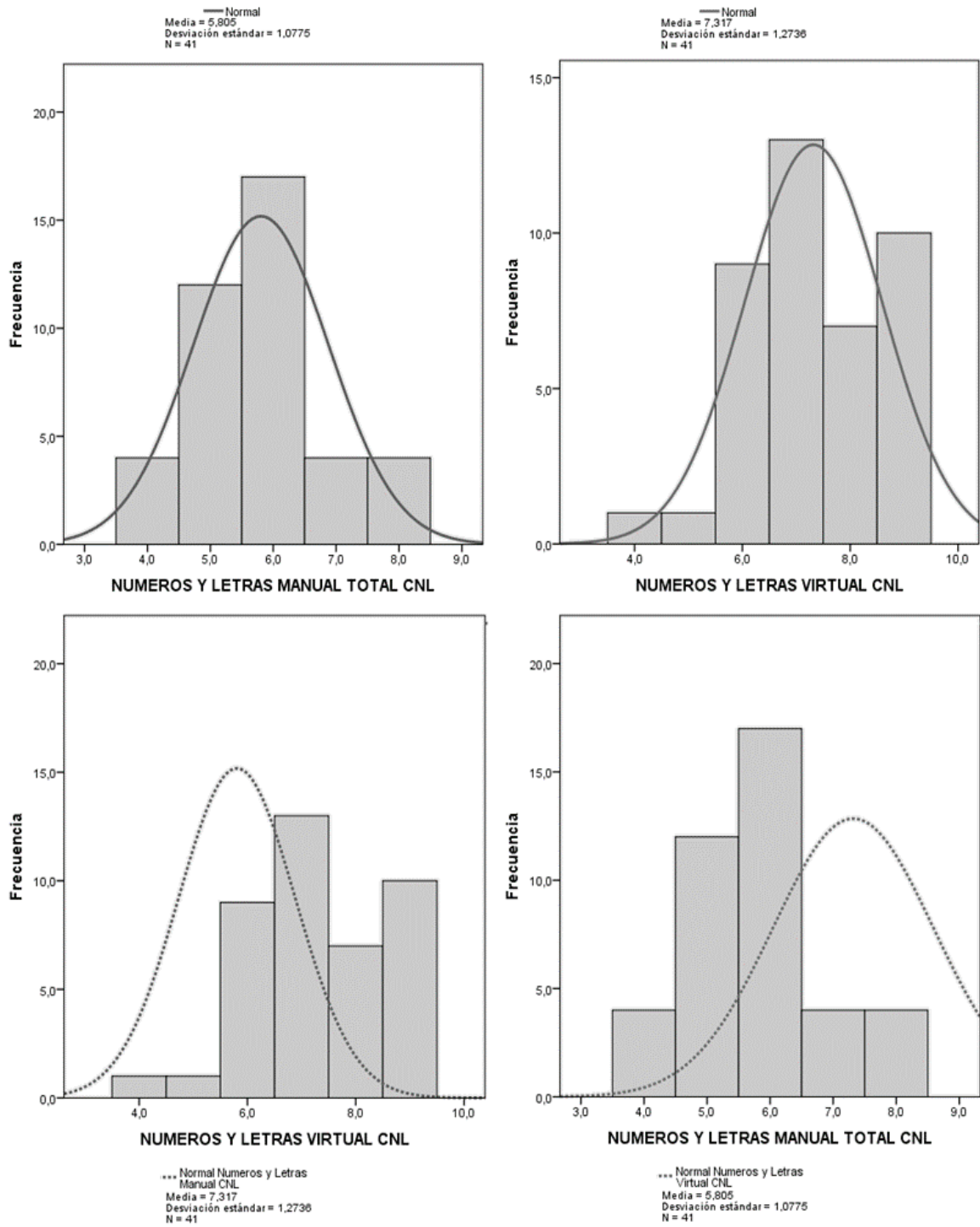


Figura 5. Histogramas que muestran las diferencias entre modalidades del Span atencional de la prueba números y letras.

Fuente: elaboración propia.

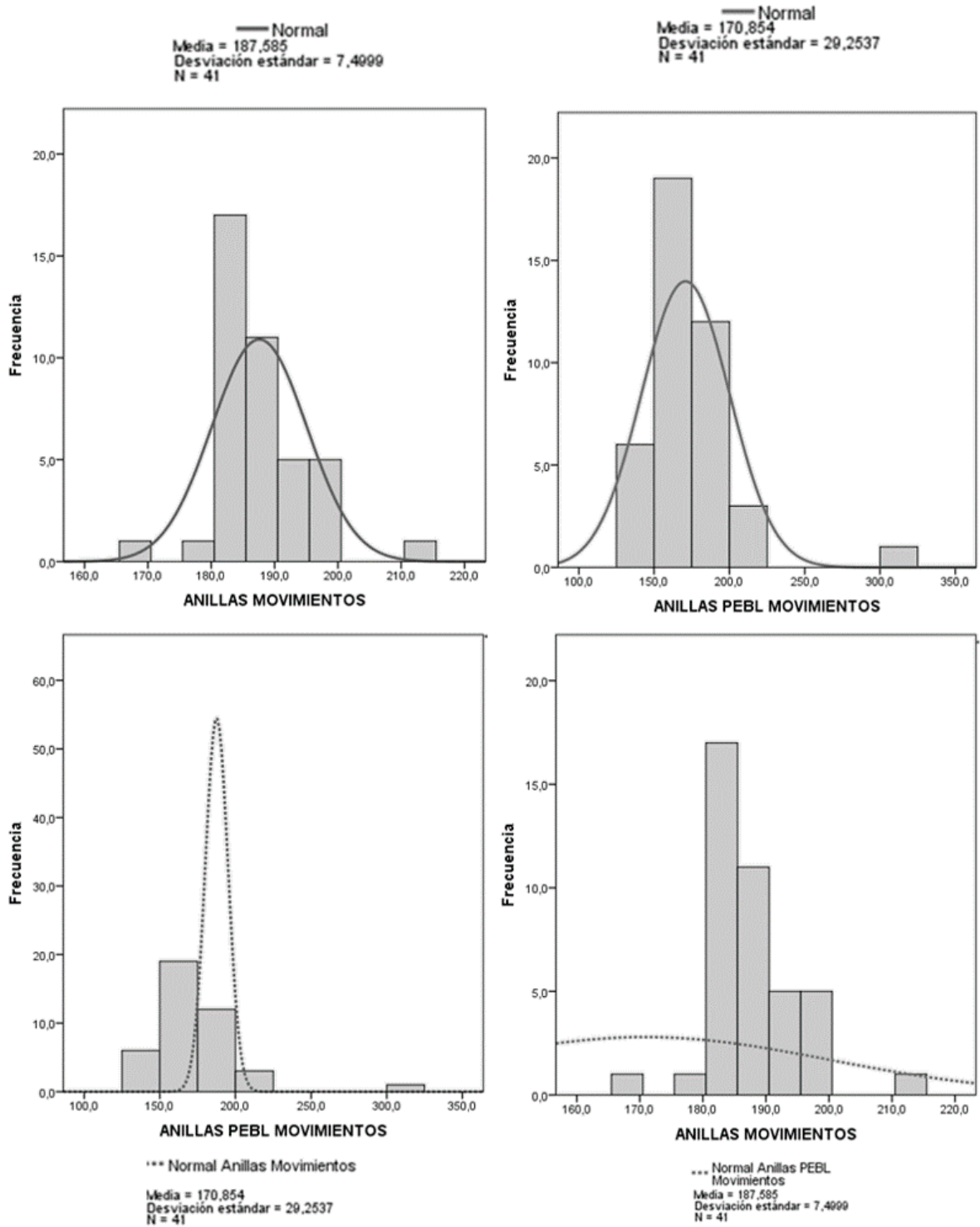


Figura 6. Histogramas que muestran las diferencias entre puntajes directos de la prueba anillas y su versión virtual por medio de la plataforma PEBL.

Fuente: elaboración propia.



## Discusión

El objetivo principal de este estudio fue comparar las aplicaciones entre las pruebas neuropsicológicas en modalidad de lápiz y papel e informatizadas. Los resultados mostraron en general que a lo largo de las pruebas no es posible establecer un desempeño similar en la ejecución de las dos versiones, probablemente debido a factores externos que son implementados en las versiones informatizadas y que pueden servir de claves adicionales para la ejecución de las mismas (Sands, Waters y Mcbride, 1999). En cuanto a la aplicación en físico se han denominado algunas claves importantes que facilitan la identificación de palabras clave por parte de los sujetos por variaciones en la entonación (Campos y Ameijide, 2014). En este mismo sentido, diferencias en cuanto al lugar de aplicación puede influir en la variación debido a la naturaleza misma de las pruebas y que se ha sustentado bajo estudios sobre ergonomía (Salmerón y Cañas, 2004).

En el caso de la atención medida en la prueba d2 se plantea que la influencia del tipo de tarea que se debe realizar en cada modalidad (tachar manualmente la letra en un papel o elegir una respuesta con el teclado en la versión informatizada) puede explicar las diferencias. Así, el hecho de tener un estímulo visual que se debe leer linealmente y convertir en respuesta motora (tachar o no tachar el estímulo) inhibiendo estímulos irrelevantes como primera aproximación a la tarea, puede facilitar el proceso de respuesta de la tarea virtual. Esto, debido a que en la modalidad informatizada se presenta un estímulo de manera aislada y se debe responder a este por medio de un teclado. La diferencia principal radica en que en la modalidad virtual no se deben inhibir los demás estímulos que están al lado y la instrucción se facilita en este aspecto, esto es coherente con los componentes de la atención expresados por Sohlberg y Mateer, (2001) ya que la atención focalizada se considera como un proceso mucho más básico y menos complejo que la selectiva, la cual sería necesaria para inhibir el estímulo relevante dentro de una lista larga.

En cuanto a la prueba de números y letras se presentó una diferencia en la puntuación directa según

el orden de aplicación, en donde, aunque se evidenció un puntaje inferior en las pruebas virtuales, se encontró que estas contaban con un span atencional mayor; por lo que se plantea la posibilidad de que este cambio se deba a una mayor distractibilidad en el formato virtual al no tener una interacción personal con el evaluador. Esto, debido a que pese a cometer más errores en la ejecución virtual en algunos reactivos se lograba avanzar más recordando y ordenando una mayor cantidad de ítems. Esta característica podría indicarse como un costo de las pruebas informatizadas que previamente ha sido señalado (Baumer *et al.*, 2009; Noyes y Garland, 2008). Seguido a esto, el span atencional en la modalidad informatizada podría haber sido mayor debido a que los elementos físicos de la prueba virtual, tales como el teclado, pueden brindar claves mnemotécnicas de aprendizaje que significa una ayuda en la tarea que involucra el span atencional como las ejemplificadas en los estudios de Campos y Ameijide (2014).

Basados en los resultados adquiridos en la comparación de la prueba de Anillas, se plantea que la versión informatizada de Anillas "ToL" (PEBL) favorece la planeación, debido a que aumenta la cantidad de tiempo con una reducción significativa de movimientos (Piper *et al.*, 2012, 2015). Así mismo, se plantea que el aumento en el tiempo se puede deber a una dificultad en la realización de dicha tarea puesto que su instrumento es el cursor táctil de un computador. Esto significa que en la modalidad virtual se da un proceso motriz diferente al movimiento del brazo y el agarre, acciones necesarias para realizar la tarea de manera física (Benjumea, 2009), por lo tanto se vería la influencia de la actividad motora para mejorar el desempeño de planeación.

En cuanto a la prueba de localización espacial, al notar que los resultados no muestran una diferencia en relación a la modalidad de aplicación, se plantea que el instrumento virtual puede contar con una consistencia con respecto a la modalidad manual, sin que esto implique que ambos instrumentos realicen la medición de la misma manera. Esto, debido a que la diferencia se puede ver reflejada en los resultados en la modalidad virtual. Se sugiere que entre las versiones puede presentarse

una diferencia entre los componentes práxico y la necesidad de motricidad que podría ser un indicador de diferencia en la evaluación de las pruebas (Vecchi y Richardson, 2001).

Por otro lado, otros factores que pudieron tener efecto sobre los resultados fueron el agotamiento de los participantes afectando los procesos de atención focalizada y los tipos jerarquizados de la atención explicarían las facilidades o dificultades de un evaluado para realizar una tarea en concordancia con la implicaciones en términos atencionales que estas requieran (Sohlberg y Mateer, 2001).

En la prueba dígitos, al observar que el span atencional y de memoria de trabajo es similar en ambas modalidades, a pesar que la puntuación directa muestra una diferencia evidente, se plantea la hipótesis de que la diferencia entre la puntuación se debe a que el participante puede llegar a perder interés o atención en la prueba debido a la facilidad o repetitividad de esta, por lo tanto, la dificultad de la tarea inicial de dígitos directos puede generar que el participante no tenga motivación frente a la tarea, mientras que en la tarea inversa que implica memoria de trabajo, se presentó mayor dificultad, lo cual puede llevar a que los sujetos tengan un esfuerzo mayor y un mejor desempeño. Esto se puede ver mencionado en estudios como los realizados por Herczeg y Lapegna, (2012) en donde afirman que en términos de aprendizaje, la motivación es determinante sobre todo en el modo de pensar y actuar de manera intrínseca y extrínseca. En este mismo sentido, Pinrich (1991, citado en Herczeg y Lapegna, 2012) afirma que una valoración positiva de la tarea conlleva a un compromiso más intenso con las mismas, es decir, que la considerar la tarea como novedosa o útil el sujeto tendrá mayor motivación y esta a su vez influenciará en términos positivos su puntaje en las pruebas.

A partir de lo anterior y en forma de conclusión, es posible afirmar que la comparación del desempeño y las diferencias de cada proceso de aplicación de las pruebas neuropsicológicas implementadas no son equiparables, debido a que los procesos implicados para la resolución de pruebas manuales y virtuales son diferentes, tanto en la recepción

del estímulo como en codificación y evocación de respuestas (Baumer *et al.*, 2009; Jeong, 2014; Noyes y Garland, 2008).

Se observó que el uso de pruebas informatizadas presentó ventajas y desventajas. Dentro de las ventajas se encontró que las pruebas virtuales pueden presentar más versatilidad en la investigación (Garaizar *et al.*, 2012; Matute, Vadillo y Bárcena, 2007; Sands *et al.*, 1999) destacándose la cantidad de tiempo empleado, aumento en la población que se evalúa y un mayor control de variables al permitir eliminar sesgos producidos por el evaluador. Ahora bien, es necesario establecer la validez predictiva, de constructo, de criterio y de contenido en las pruebas y tareas virtuales para favorecer, de esta forma, que instrumentos virtuales lleguen a ser equiparables con los instrumentos manuales y puedan ser utilizados y estandarizados.

Es de resaltar que el formato virtual de las pruebas introduce una serie de variables novedosas que diferencian las pruebas virtuales de la forma tradicional. Dentro de estas variables, se reconoce como una limitación el hecho de perder en las pruebas informatizadas los datos cualitativos, que solo se pueden obtener por medio de la interacción directa con el evaluado y se pierde el rol del evaluador dentro de las pruebas (Chan, Shum, Toulopoulou y Chen, 2008; Rabin, Barr y Burton, 2005; Reynolds y Fletcher-Janzen, 2013). Esto sugiere una consideración en torno a la pertinencia del uso de este tipo de instrumentos en ámbitos clínicos, educativos o de evaluación en general, donde no se considera que la modalidad virtual sea la más apropiada puesto que se correría el riesgo de perder información cualitativa que complementa y ayuda a profundizar la información obtenida.

A pesar de no ser representativa la población del estudio, se encuentra que los resultados aportan a la evaluación de las modalidades de pruebas aplicadas tradicionalmente en diferentes campos en neuropsicología y que se suman al esfuerzo de optimizar y automatizar procesos de evaluación que superen limitaciones de espacio, acceso y aplicación a gran escala (Mueller y Piper, 2014; Piper *et al.*, 2012, 2015; Prada Sarmiento *et al.*, 2010).

## Referencias

- Alvarez, J. A. y Emory, E. (2006). Executive Function and the Frontal Lobes: A Meta-Analytic Review. *Neuropsychology Review*, 16(1), 17-42. <http://doi.org/10.1007/s11065-006-9002-x>
- Ardila, A., Rosselli, M. y Strumwasser, S. (2009). Neuropsychological Deficits in Chronic Cocaine Abusers. *International Journal of Neuroscience*, 57(1-2), 73-79. <http://doi.org/10.3109/00207459109150348>
- Armayones Ruiz, M., Boixadós, M., Gómez Zúñiga, B., Guillamón, N., Hernández, E., Nieto, R., ... Sara, B. (2015). Psicología 2.0: oportunidades y retos para el profesional de la psicología en el ámbito de la salud. *Papeles Del Psicólogo*, 36(2), 153-160.
- Baumer, M., Roded, K. y Gafni, N. (2009). Assessing the Equivalence of Psychometric Tests. *Proceedings of the 2009 GMAC Conference on Computerized Adaptive Testing*.
- Benjumea, M. (2009). *Elementos constitutivos de la Motricidad como dimensión humana*. Medellín: Universidad de Antioquia.
- Bieri, R., Jäger, M., Gruber, N., Nef, T., Müri, R. M. y Mosimann, U. P. (2014). A novel computer test to assess driving-relevant cognitive functions - a pilot study. *International Psychogeriatrics*, 26(2), 229-238. <http://doi.org/10.1017/S104161021300183X>
- Brickenkamp, R. y Zillmer, E. (2002). *Test de Atención d2*. Madrid: TEA Ediciones.
- Budde, H., Voelcker-Rehage, C., Pietraßyk-Kendziorra, S., Ribeiro, P. y Tidow, G. (2008). Acute coordinative exercise improves attentional performance in adolescents. *Neuroscience Letters*, 441(2), 219-223. <http://doi.org/http://dx.doi.org/10.1016/j.neulet.2008.06.024>
- Bugbee, A. C. (1996). The Equivalence of Paper-and-Pencil and Computer-Based Testing. *Journal of Research on Computing in Education*, 28(3), 282-299. <http://doi.org/10.1080/08886504.1996.10782166>
- Cadavid-Ruiz, N., Del Rio, P., Egado, J. y Galindo, P. (2016). Age Related Changes in the Executive Function of Colombian Children. *Universitas Psychologica*, 15(5), 1-10. <http://doi.org/10.11144/Javeriana.upsy15-5.arce>
- Cadavid, N. R. y Del Rio, P. (2012). Memoria de trabajo verbal y su relación con variables sociodemográficas en niños colombianos. *Acta Colombiana de Psicología*, 15(1), 99-109.
- Campos, A. y Ameijide, L. (2014). Mnemotecnia y metamemoria en las personas mayores. *Universitas Psychologica*, 14(1), 57-66. <http://doi.org/10.11144/Javeriana.upsy14-1.mmpm>
- Cellini, N. (2016). Memory Consolidation in Sleep Disorders. *Sleep Medicine Reviews*, 35, 101-112. <http://dx.doi.org/10.1016/j.smrv.2016.09.003>
- Chan, R. C. K., Shum, D., Touloupoulou, T. y Chen, E. Y. H. (2008). Assessment of executive functions: review of instruments and identification of critical issues. *Archives of Clinical Neuropsychology*, 23(2), 201-16. <http://doi.org/10.1016/j.acn.2007.08.010>
- Collerton, J., Collerton, D., Arai, Y., Barrass, K., Eccles, M., Jagger, C., ... Von Zglinicki, T. (2007). A comparison of computerized and pencil-and-paper tasks in assessing cognitive function in community-dwelling older people in the Newcastle 85+ pilot study. *Journal of the American Geriatrics Society*, 55(10), 1630-1635. <http://doi.org/10.1111/j.1532-5415.2007.01379.x>
- Dockery, C. A., Hueckel-Weng, R., Birbaumer, N. y Plewnia, C. (2009). Enhancement of Planning Ability by Transcranial Direct Current Stimulation. *The Journal of Neuroscience*, 29(22), 7271-7277. <http://doi.org/10.1523/JNEUROSCI.0065-09.2009>
- Garaizar, P., Vellido, M. A., López-de-Ipiña, D. y Matute, H. (2012). The Web as a Platform for e-Research in the Social and Behavioral Sciences. In A. A. Juan, T. Daradoumis, M. Roca, S. E. Grasman y J. Faulin (Eds.), *Collaborative and Distributed E-Research* (pp. 34-61). Hershey: IGI Global. <http://doi.org/10.4018/978-1-4666-0125-3>

- García, T., Álvarez-García, D., Areces, D., Segurola, L. A., González-Castro, P. y Rodríguez, C. (2018). Psychometric characteristics of the executive functioning scale for families (EFS-F). *Revista Iberoamericana de Diagnostico y Evaluacion Psicologica*, 2(47), 71-82. <http://doi.org/10.21865/RIDEP47.2.05>
- Herczeg, C. y Lapegna, M. (2012). Autorregulación, estrategias y motivación en el aprendizaje. *Autorregulation, Estratégias e Motivación En El Aprendizaje*, 35(37), 9-19.
- Holländare, F., Andersson, G. y Engström, I. (2010). A Comparison of Psychometric Properties Between Internet and Paper Versions of Two Depression Instruments (BDI-II and MADRS-S) Administered to Clinic Patients. *Journal of Medical Internet Research*, 12(5), e49. <http://doi.org/10.2196/jmir.1392>
- Hurtado-Parrado, C., Arias-Higuera, M., González, C. A., Bohórquez, M. C., Henao, K., García, E. y Zapata, A. (2016). Verbal-nonverbal interactions during a computerized adaptation of Catania *et al.*'s (1982) experimental task: The effects of embedding an aversive contingency in the nonverbal component. *Learning and Motivation*, 55, 13-30. <http://doi.org/10.1016/j.lmot.2016.05.006>
- Ison, M. S. y Carrada, M. (2010). Assessment of attentional efficiency: Preliminary normative study carried out with students in Argentina. *Revista Iberoamericana de Diagnostico y Evaluacion Psicologica*, 1(29), 129-146.
- Jeong, H. (2014). A comparative study of scores on computer-based tests and paper-based tests. *Behaviour & Information Technology*, 33(4), 410-422. <http://doi.org/10.1080/0144929X.2012.710647>
- Jiménez, J., Hernández, S., García, E., Díaz, A., Rodríguez, C. y Martín, R. (2012). Test de Atención d2 datos normativos y desarrollo evolutivo de la atención en educación primaria. *European Journal of Education and Psychology*, 5(1), 93-106. <https://doi.org/10.30552/ejep.v5i1.79>
- Johnson, S., Saykin, A., Flashman, L., McAllister, T. y Sparling, M. (2001). Brain activation on fMRI and verbal memory ability: Functional neuroanatomic correlates of CVLT performance. *Journal of the International Neuropsychological Society*, 7(01), 55-62.
- Joubert, T. y Kriek, H. J. (2009). Psychometric comparison of paper-and-pencil and online personality assessments in a selection setting. *SA Journal of Industrial Psychology*, 35(1), 78-88. <http://doi.org/10.4102/sajip.v35i1.727>
- Lampit, A., Hallock, H. y Valenzuela, M. (2014). Computerized Cognitive Training in Cognitively Healthy Older Adults: A Systematic Review and Meta-Analysis of Effect Modifiers. *PLoS Medicine*, 11(11), e1001756. <http://doi.org/10.1371/journal.pmed.1001756>
- Leal-Soto, F. y Alonso-Tapia, J. (2017). Classroom motivational climate questionnaire: Intercultural, inter-gender, developmental and predictive validity. *Revista Iberoamericana de Diagnostico y Evaluacion Psicologica*, 3(45), 57-70. <http://doi.org/10.21865/RIDEP45.3.05>
- Lozzia, G. S., Abal, F. J. P., Blum, D., Aguerri, M. E., Galibert, M. S. y Attorresi, H. F. (2009). Tests Informatizados: nuevos desafíos prácticos y éticos para la Evaluación Psicológica. *Summa Psicológica UST*, 6(1), 135-148.
- Markowetz, A., Błaszkiwicz, K., Montag, C., Switala, C. y Schlaepfer, T. E. (2014). Psychoinformatics: Big Data shaping modern psychometrics. *Medical Hypotheses*, 82(4), 405-411. <http://doi.org/10.1016/j.mehy.2013.11.030>
- Matute, H., Vadillo, M. A. y Bárcena, R. (2007). Web-based experiment control software for research and teaching on human learning. *Behavior Research Methods*, 39(3), 689-693. <https://doi.org/10.3758/bf03193041>
- Mead, D. A. y Drasgow, F. (1993). Equivalence of Computerized and Paper- and- Pencil Cognitive Ability Tests: A Meta- Analysis. *Psychological Bulletin*, 114(3), 449-458.

- Millis, S. R., Malina, A. C., Bowers, D. A. y Ricker, J. H. (2010). Confirmatory Factor Analysis of the Wechsler Memory Scale-III. *Journal of Clinical and Experimental Neuropsychology*, 21(1), 87-93. <https://doi.org/10.1076/jcen.21.1.87.937>
- Millsap, C. M. (2000). *Comparing of Computer Testing Versus Traditional Paper and Pencil Testing*. Denton: University of North Texas.
- Moll, J., Oliveira-Souza, R. de, Moll, F. T., Bramati, I. E. y Andreiuolo, P. A. (2002). The cerebral correlates of set-shifting: an fMRI study of the trail making test. *Arquivos de Neuro-Psiquiatria*, 60(4), 900-905. <http://doi.org/10.1590/S0004-282X2002000600002>
- Montoya-Arenas, D. A., Trujillo-Orrego, N. y Pineda-Salazar, D. A. (2010). Capacidad intelectual y función ejecutiva en niños intelectualmente talentosos y en niños con inteligencia promedio. *Universitas Psychologica*, 9(3), 737-748.
- Mueller, S. T. y Piper, B. J. (2014). The Psychology Experiment Building Language (PEBL) and PEBL Test Battery. *Journal of Neuroscience Methods*, 222, 250-259. <http://doi.org/10.1016/j.jneumeth.2013.10.024>
- Muñiz, J., Hernández, A. y Ponsoda, V. (2015). Nuevas directrices sobre el uso de los tests: investigación, control de calidad y seguridad. *Papeles Del Psicólogo*, 36(3), 161-173.
- Noyes, J. M. y Garland, K. J. (2008). Computer- vs. paper-based tasks: are they equivalent? *Ergonomics*, 51(9), 1352-1375. <http://doi.org/10.1080/00140130802170387>
- Olea, J., Abad, F. J. y Barrada, J. R. (2010). Tests informatizados y otros nuevos tipos de tests. *Papeles Del Psicólogo*, 31(1), 97-107. <http://doi.org/10.1017/CBO9781107415324.004>
- Olea, J., Revuelta, J., Ximénez, M. y Abad, F. (2000). Psychometric and psychological effects of review on computerized fixed and adaptive tests. *Psicológica*, 21, 157-173.
- Open Science Collaboration. (2012). An Open, Large-Scale, Collaborative Effort to Estimate the Reproducibility of Psychological Science. *Perspectives on Psychological Science*, 7(6), 657-660. <http://doi.org/10.1177/1745691612462588>
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), 4716-8. <http://doi.org/10.1126/science.aac4716>
- Piper, B. J., Li, V., Eiwaz, M. A., Kobel, Y. V., Benice, T. S., Chu, A. M., ... Mueller, S. T. (2012). Executive function on the Psychology Experiment Building Language tests. *Behavior Research Methods*, 44(1), 110-123. <http://doi.org/10.3758/s13428-011-0096-6>
- Piper, B. J., Mueller, S. T., Geerken, A. R., Dixon, K. L., Kroliczak, G., Olsen, R. H. J. y Miller, J. K. (2015). Reliability and validity of neurobehavioral function on the Psychology Experiment Building Language test battery in young adults. *PeerJ*, 3, e1460. <http://doi.org/10.7717/peerj.1460>
- Piper, B. J., Mueller, S. T., Talebzadeh, S. y Ki, M. J. (2016). Evaluation of the validity of the Psychology Experiment Building Language tests of vigilance, auditory memory, and decision making. *PeerJ*, 2016(3). <http://doi.org/10.7717/peerj.1772>
- Pisoni, D. B. y Cleary, M. (2003). Measures of working memory span and verbal rehearsal speed in deaf children after cochlear implantation. *Ear Hear*, 24(1 Suppl), 106S-120S.
- Ployhart, R. E., Weekley, J. A., Holtz, B. C. y Kemp, C. (2003). Web-based and paper-and-pencil testing of applicants in a proctored setting: are personality, biodata, and situational judgment tests comparable? *Personnel Psychology*, 56(3), 733-752. <http://doi.org/10.1111/j.1744-6570.2003.tb00757.x>
- Portellano, J. ., Díez, A. y García, J. (2007). El test de las anillas (TA), un nuevo instrumento para la evaluación de las Funciones Ejecutivas The test of the rings , a new instrument. *Mapfre Medicina*, 18(1), 54-63.

- Prada Sarmiento, E. L., Pineda Garzón, G. E., Mejía Orduz, M. A. y Conde Cotés, C. A. (2010). Prueba computarizada Memonum: Efecto de intervalos y distractores sobre la memoria de trabajo en mujeres mayores de 50 años. *Universitas Psychologica*, 9(3), 893-906.
- Rabin, L. A., Barr, W. B. y Burton, L. A. (2005). Assessment practices of clinical neuropsychologists in the United States and Canada: a survey of INS, NAN, and APA Division 40 members. *Archives of Clinical Neuropsychology*, 20(1), 33-65. <http://doi.org/10.1016/j.acn.2004.02.005>
- Reips, U. D. (2001). The Web Experimental Psychology Lab: five years of data collection on the Internet. *Behavior Research Methods, Instruments y Computers*, 33(2), 201-211.
- Reynolds, C. y Fletcher-Janzen, E. (2013). *Handbook of Clinical Child Neuropsychology*. New York: Springer.
- Richards, D. (2013). Developments in Technology-Delivered Psychological Interventions. *Universitas Psychologica*, 12(2), 571-579. <http://doi.org/10.11144/Javeriana.UPSY12-2.dtdp>
- Rosas, R. (2013). Evidencia de validez convergente entre las versiones chilenas de WAIS-IV y WISC-III. *Papeles de Investigación*, 1, 1-11.
- Rosenthal, E. N., Riccio, C. A., Gsanger, K. M. y Jarratt, K. P. (2006). Digit Span components as predictors of attention problems and executive functioning in children. *Archives of Clinical Neuropsychology*, 21(2), 131-139. <http://doi.org/10.1016/j.acn.2005.08.004>
- Roy, S. y Park, N. W. (2010). Dissociating the memory systems mediating complex tool knowledge and skills. *Neuropsychologia*, 48(10), 3026-3036. <http://dx.doi.org/10.1016/j.neuropsychologia.2010.06.012>
- Rubia, K., Russell, T., Overmeyer, S., Brammer, M. J., Bullmore, E. T., Sharma, T., ... Taylor, E. (2001). Mapping Motor Inhibition: Conjunctive Brain Activations across Different Versions of Go/No-Go and Stop Tasks. *NeuroImage*, 13(2), 250-261. <http://doi.org/10.1006/nimg.2000.0685>
- Salmerón, L. y Cañas, J. J. (2004). Sobre la relación entre la ergonomía y psicología cognitivas. *Anuario de Psicología*, 35(4), 507-519.
- Sands, W. A., Waters, B. K. y McBride, J. R. (1999). *CATBOOK Computerized Adaptive Testing: From Inquiry to Operation*. Alexandria, VA: United States Army Research Institute for the Behavioral and Social Sciences.
- Savage, G. (2016). Cognitive Neuropsychological Formulation. In *Neuropsychological Formulation* (pp. 221-239). Cham: Springer International Publishing. [http://doi.org/10.1007/978-3-319-18338-1\\_13](http://doi.org/10.1007/978-3-319-18338-1_13)
- Sohlberg, M. M. y Mateer, C. (2001). *Cognitive Rehabilitation an interactive Neuropsychological Approach*. New York: The Guilford Press.
- Toril, P., Reales, J. M., Mayas, J. y Ballesteros, S. (2016). Video Game Training Enhances Visuospatial Working Memory and Episodic Memory in Older Adults. *Frontiers in Human Neuroscience*, 10(5), 1-14. <http://doi.org/10.3389/fnhum.2016.00206>
- Vallejo, M. A., Jordán, C. M., Díaz, M. I., Comeche, M. I. y Ortega, J. (2007). Psychological assessment via the internet: A reliability and validity study of online (vs paper-and-pencil) versions of the General Health Questionnaire-28 (GHQ-28) and the Symptoms Check-List-90-Revised (SCL-90-R). *Journal of Medical Internet Research*, 9(1), 1-10. <http://doi.org/10.2196/jmir.9.1.e2>
- Vecchi, T. y Richardson, J. T. (2001). Measures of visuospatial short-term memory: the Knox Cube Imitation Test and the Corsi Blocks Test compared. *Brain and Cognition*, 46(1-2), 291-295. [http://doi.org/10.1016/S0278-2626\(01\)80086-5](http://doi.org/10.1016/S0278-2626(01)80086-5)
- Vianin, P. (2016). Computerized Exercises to Promote Transfer of Cognitive Skills to Everyday

Life. *Frontiers in Psychiatry*, 7(4), 1-5. <http://doi.org/10.3389/fpsyt.2016.00056>

Wang, S., Jiao, H., Young, M. J., Brooks, T. y Olson, J. (2007). Comparability of Computer-Based and Paper-and-Pencil Testing in K 12 Reading Assessments: A Meta-Analysis of Testing Mode Effects. *Educational and Psychological Measurement*, 68(1), 5-24. <http://doi.org/10.1177/0013164407305592>

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale* (Third edit). San Antonio, TX: The Psychological Corporation.

Zakzanis, K. K., Mraz, R. y Graham, S. J. (2005). An fMRI study of the Trail Making Test. *Neuropsychologia*, 43(13), 1878-1886. <http://doi.org/10.1016/j.neuropsychologia.2005.03.013>