

El Design Organization Test (DOT) en adultos peruanos: parámetros de validación y normativos

The Design Organization Test (DOT) in Peruvian adults: Validation and normative parameters

O Design Organization Test (DOT) em adultos peruanos: validação e parâmetros normativos

Le Design Organization Test (DOT) chez les adultes péruviens: paramètres de validation et normatifs

César Merino Soto^{1 *} , 0000-0002-1407-8306

Danitsa Alarcón Parco^{1 **} , 0000-0002-4556-0795

Verónica López-Fernández² , 0000-0003-0483-5884

Marcio Soto-Añari³ , 0000-0002-9121-3284

Karen González Restrepo⁴ , 0000-0002-6711-6249

¹ Universidad de San Martín de Porres - Perú

² Universidad Internacional de La Rioja - España, veronica.lopez@unir.net

³ Universidad Católica San Pablo - Perú ,
msoto@ucsp.edu.pe

⁴ Secretaría de Educación del Distrito de Bogotá - Colombia ,
kgonzalez1@educacionbogota.edu.co

Resumen

El presente estudio reporta la información normativa y psicométrica del DOT (forma A y B), una medida breve de organización perceptual-visual relevante para la investigación aplicada y la práctica clínica. Participó una muestra efectiva de 775 adultos universitarios, de Lima (Perú). La administración del test se hizo en grupo, con instrucciones estandarizadas. Se halló que la confiabilidad test-retest era moderadamente alta; la covariación con una prueba de inteligencia fluida fue pequeña, pero teóricamente consistente. La información clínico-métrica sobre el error de medición y el cambio confiable fue adecuada (mínimo cambio detectable). Las diferencias intra-grupos y entre-grupos (muestra americana) revelaron patrones de resultados esencialmente consistentes, detectándose un efecto de la práctica comparativamente mayor en la muestra peruana. Se discuten las implicaciones de los resultados.

Palabras clave: evaluación, percepción visual, validación, neuropsicología

Abstract

The present study reports the normative and psychometric in-

formation of DOT (form A and B), a brief measure of visuospatial organization relevant to applied research and clinical practice. A sample of 775 university adults participated, from Lima (Peru). The administration of the test was done in groups, with standardized instructions. We found moderately high test-retest reliability; and the covariation with a fluid intelligence test was small but theoretically consistent. The clinimetrics of the measurement error and the reliable change was adequate (minimum detectable change). The intra-group and between-group (American sample) differences revealed patterns of essentially consistent results, detecting a comparatively greater effect of practice in the Peruvian sample. The implications of the results are discussed.

Keywords: Assessment, visuospatial ability, validation, neuropsychology

Resumo

O estudo apresenta a informação normativa e psicométrica do DOT (formas A e B), uma medida breve da capacidade de organização perceptual-visual relevante para a pesquisa aplicada e para a prática clínica. Participou uma amostra efetiva de 775 adultos universitários de Lima (Peru). A administração do teste foi feita em grupos, com instruções padronizadas. Os resultados amostram uma confiabilidade teste-reteste moderadamente alta; a covariação com um teste de inteligência fluida foi pequena, mas teoricamente consistente. As informações clínicas sobre o erro de medição e a mudança confiável foram adequadas (mudança mínima detectável). As diferen-

ças intragrupos e entre grupos (amostra americana) revelaram padrões de resultados essencialmente consistentes, detectando-se um efeito comparativamente maior da prática na amostra peruana. As implicações dos resultados são discutidas.

Palavras-chave: avaliação, capacidade visuospacial, validação, neuropsicología

Résumé

La présente étude rapporte les informations normatives et psychométriques du DOT (formulaires A et B), une brève mesure de l'organisation capacité visuospaciale pertinente pour la recherche appliquée et la pratique clinique. Un échantillon effectif de 775 adultes universitaires ont participé, de Lima (Pérou). L'administration du test a été effectuée en groupes, avec des instructions normalisées. Nous avons constaté que la fiabilité test-retest était relativement élevée; la covariation avec un test d'intelligence fluide était petite mais théoriquement cohérente. Les informations clinométriques sur l'erreur de mesure et le changement fiable étaient adéquates (changement minimum détectable). Les différences entre groupes et entre groupes (échantillon américain) ont révélé des tendances de résultats essentiellement cohérents, détectant un effet comparativement plus important de la pratique dans l'échantillon péruvien. Les implications des résultats sont discutées.

Mots-clés: évaluation, capacité visuospaciale, validation, neuropsychologie

Dentro de las habilidades neuropsicológicas de tipo no verbal, destacan aquellas relacionadas con tareas visoespaciales, de hecho, las más estudiadas (Solovieva et al., 2017). Esta relevancia cognitiva y académica está justificada en el hecho de que los seres humanos procesamos constantemente información visoespacial para resolver un amplio abanico de actividades cotidianas, tales como el reconocimiento de objetos y lugares, el razonamiento, la navegación, el uso de herramientas y la construcción (Kosmidis et al., 2014; Maguire et al., 2006). Por ello, la habilidad visoespacial es inherente al ser humano, lo cual le ha permitido adaptarse a su medio y a su espacio, de manera que, le ha posibilitado percibir, reconocer y explorar construcciones físicas o intelectuales referentes a la posición o al espacio dentro de sistemas estáticos o dinámicos (Ness et al., 2017).

Cuando hablamos de habilidades visoespaciales es relevante mencionar la existencia de modelos de procesamiento basados en marcos de referencia (MR), entendiendo tal procesamiento como la capacidad para enfocar y discriminar a través del sentido de la vista el punto en el que se ubican los estímulos en base a referencias físicas facilitadas que promuevan la orientación en el espacio (Borsting, 1996). En este sentido, dicho procesamiento, posibilita la representación mental de los estímulos, las tareas de reconocimiento, así como el desenvolvimiento en función de las propiedades de esos estímulos (Merchán & Henao, 2011).

Dado que el procesamiento visoespacial consta de dos MR, uno basado en el objeto y el ambiente y otro en el observador (Zacks & Michelon, 2005), cabe resaltar que se asocian

a diferentes bases neuronales, las cuales forman circuitos córtico-subcorticales que abarcan a todos los lóbulos del cerebro (Tseng et al., 2010; De Benedictis, et al., 2014), permitiendo así la entrada de los estímulos visuales y su ubicación espacial (Von-Allmen et al., 2014). Es indudable que la envergadura de dicho proceso requiere de complejos procesos cognitivos y representaciones mentales espaciales (Mitolo et al., 2015).

En sujetos sanos el procesamiento visoespacial está asociado a una mayor organización estructural y funcional a nivel cerebral (López Velásquez et al., 2016), logrando así evidenciar la relación que se establece entre las habilidades visoespaciales con otros procesos cognitivos y su implicación en la adquisición o alteración de las mismas. En esta línea, se han planteado las habilidades visoespaciales como sinónimo de la inteligencia espacial, dada la capacidad de llevar a cabo adecuadas localizaciones en el espacio que implican relacionar, deducir e integrar aspectos que se relacionan con su entorno. De igual forma, la localización en el tiempo, que le permite al sujeto analizar, describir y transferir la información que percibe de dicho entorno, hace parte esencial del desarrollo de un individuo, de manera que puede equipararse con ejes de pensamiento como el lingüístico y lógico-matemático, como establece Calvo (2017), especialmente en lectoescritura y cálculo, respectivamente.

La importancia de la relación de las habilidades visoespaciales con otros procesos se explicita en estudios previos que han detectado deficiencias visoespaciales en: autismo (Seijas, 2015), síndrome de Down (Lanfranchi et al., 2015), déficit de atención (Pérez et al., 2016), dislexia (Chamberlain et al.,

2018), trastornos neurodegenerativos como la enfermedad de Alzheimer (Cronin-Golomb & Amick, 2001), esclerosis múltiple (Vleugels et al., 2000) y en pacientes con Parkinson (Chastan et al., 2019).

La importancia de las habilidades visoespaciales y su estudio se justifica fundamentalmente en el hecho de que se manifiesta como un constructo flexible y dinámico, que constituye como hemos comentado, una base importante en otras habilidades, que se conectan y se desarrollan durante las diferentes etapas evolutivas del sujeto (Ness et al., 2017). Tanto es así, que la intervención en este campo ha demostrado que la estimulación cognitiva de la función visoespacial mejora las habilidades de navegación, orientación espacial y la interacción visomotora (Astrand et al., 2014; Hampstead et al., 2014; Prince & Daniel, 2014). Inclusive, utilizando técnicas de estimulación transcraneal de corriente directa en pacientes con demencia, se han hallado mejoras en las tareas de reconocimiento visual y en la memoria episódica visual (Boggio et al., 2009, 2012).

Un paso previo a la intervención, es sin duda contar con adecuados instrumentos de evaluación de un constructo. No obstante, uno de los principales problemas de la evaluación del procesamiento visoespacial es el tiempo que toma su administración. Las baterías neuropsicológicas donde se incluye su evaluación, como las escalas de Weschler, el Test de Barcelona (Peña-Casanova et al., 1997); o que la batería Birmingham de reconocimiento de objetos o el Visual Object and Space Perception (VOSP de Warrington & James, 1991) suelen tomar, en condiciones óptimas, entre 3 y 4 horas, lo cual hace que el paciente se fatigue o “renuncie” a la evaluación y

los resultados de la misma pueden estar sesgados afectados por esas variables. Además, muchos de los pacientes descritos pueden tener alguna afectación motora, lo cual dificultaría completar con éxito la batería. Por lo tanto, se hace pertinente contar con pruebas tipo *screening* que nos permitan tener un análisis preciso y rápido sobre el nivel de afectación del procesamiento visuoespacial, tanto en contextos clínicos como educativos por tratarse de unas habilidades sin duda importantes para la vida (Burggraaf et al., 2016).

Por ello, Killgore et al. (2005) y Killgore y Gogel (2014) desarrollaron The Design Organization Test (DOT) con el objetivo de tener una prueba breve y sensible para el análisis del procesamiento visuoespacial. Aprovechando su eficiencia respecto al tiempo que toma aplicarlo y la especificidad sobre la habilidad que mide, otros estudios lo han usado como parte de una batería de pruebas neuropsicológicas, para observar el efecto de variables que podrían producir alguna varianza en el mismo (Hofman et al., 2011; Koppelmans et al., 2012).

En estudios más recientes, Killgore y Gogel (2014) destacan la importancia de la validación adicional de esta prueba, puesto que, se requiere presentar utilidad desde el ámbito clínico. El primer estudio se realizó con una muestra de estudiante universitarios y los resultados podían haberse sobrestimado, y la segunda muestra se basaba en pacientes con algún tipo de lesión cerebral, lo que podía subestimar los resultados, por lo que, se han de realizar nuevos estudios con el fin de corroborar la confiabilidad y validez, administrando así la prueba a participantes sanos que presentaban diversas capacidades o funciones intelectuales entre sí. Los hallazgos

para este nuevo estudio revelan y reiteran la validez y confiabilidad de la prueba, mostrando un alto potencial frente a la estimación del funcionamiento intelectual, funcionando incluso sobre pruebas más extensas que generan desgaste y falta de motivación en su ejecución (Killgore et al., 2014). Así, esta prueba ha mostrado buenos indicadores psicométricos en sus primeros estudios, tanto a nivel de confiabilidad como de validez.

En base a ello, el objetivo del presente estudio es obtener evidencias de validez del DOT para una muestra de universitarios de Lima Metropolitana; los objetivos secundarios son explorar las diferencias normativas, y obtener parámetros normativos de referencia para el grupo de participantes. Estos objetivos fueron cumplidos mediante la identificación de la información descriptiva (piso y techo de los puntajes, distribución teórica) una vez analizado y removido valores extremos (outliers) y patrones irrelevantes de respuesta; comparación de estos datos con los del estudio original, confiabilidad de formas paralelas, y obtención de parámetros clinicométricos.

Debido al potencial uso clínico y de investigación científica con el DOT, esta herramienta puede integrarse en las evaluaciones de despistaje en la comunidad o en grupos de trabajadores con factores psicosociales de riesgo. Por otro lado, aparentemente estos objetivos permiten definir la primera información normativa en participantes latinoamericanos en general, y peruanos en particular, debido que, hasta la fecha del presente manuscrito, no hay estudios que describan su uso. La información sobre las posibles diferencias o similitudes normativas en el DOT no están exploradas más allá de los grupos anglosajones participantes,

y se requiere indagarlas en otros contextos socioculturales de uso del DOT para lograr descripciones relevantes de la habilidad de procesamiento visuoespacial. Como con otros instrumentos de evaluación psicológica, las propiedades psicométricas e información normativa no son propiedades estáticas, ni pueden inducirse en forma segura desde sus estudios de construcción, más aún si estos se realizaron en contextos inequívocamente diferentes respecto al nuevo contexto de validación.

Método

Participantes

La muestra inicial estuvo conformada por 785 estudiantes de psicología entre 18 y 30 años, de una universidad privada ubicada en Lima Metropolitana, matriculados durante el semestre 2013-1. La participación fue voluntaria y sin incentivos, y fueron elegidos por su disponibilidad para aceptar participar en el estudio, y la firma del consentimiento informado; el criterio de exclusión fue la potencial detección de patrones estadísticamente anormales. Los datos fueron examinados inicialmente, como se describe en la sección de resultados, en el que 10 participantes que representaron al 1.27% de la muestra total, fueron excluidos del análisis adicional, por presentar patrones de respuesta estadísticamente atípicos.

La muestra efectiva después de este proceso fue de 775 participantes de diferentes distritos de Lima Metropolitana, por

lo que se puede considerar una población heterogénea respecto a la distribución de cada una. Se evaluaron a los estudiantes de todos los semestres académicos (1ro al 10mo), siendo predominante en los primeros ciclos (1ro al 4to), que más de la mitad vive con ambos padres, la mayoría de ellos no trabaja, y son solteros. La información completa aparece en la Tabla 1.

Tabla 1.- Distribución de los participantes en las variables

demográficas (n=785)

Distribución de los participantes en las variables demográficas (n=785)

	n	%
Trabajo actual		
Sí	164	21.2
No	555	71.6
Sin datos	56	7.2
Convivencia familiar		
Vivo con ambos padres	413	53.3
Vivo con la madre	185	23.9
Vivo con el padre	37	4.8
No vivo con ninguno	79	10.2
Sin datos	61	7.9
Lugar de Nacimiento		
Lima	612	79.0
Fuera de Lima	108	13.9
Sin datos	55	7.1
Estado civil		
Soltero	707	91.2
Casado/conviviente	17	2.2
Sin datos	51	6.6
Semestre		
1ro	108	13.9
2do	98	12.6
3ro	109	14.1
4to	118	15.2
5to	87	11.2
6to	75	9.7
7to	92	11.9
8vo	4	.5
9no	24	3.1
10mo	7	.9
Sin datos	53	6.8
Ultimo nivel de estudio de la madre		
Primaria	24	3.1
Secundaria	180	23.2
Técnico	295	38.1
Universitario	214	27.6
Sin datos	62	8.0
Ultimo nivel de estudio del padre		
Primaria	20	2.6
Secundaria	154	19.9
Técnico	244	31.5
Universitario	289	37.3
Sin datos	68	8.8

Medición

Prueba de Organización de Diseños (Design Organization Test, DOT; Killgore et al., 2005). Se trata de un instrumento breve de papel y lápiz, que evalúa la habilidad visoespacial, y que ha sido diseñado como una alternativa eficiente a la Prueba de Diseños de Cubos (Block Design Test, BDT) que es una subescala de Wechsler Adult Intelligence Scale (WAIS Wechsler, 1981). Dicha eficiencia se basa en que el DOT es más rápida y de fácil aplicación, por ejemplo, comparado con el BDT (Burggraaf et al., 2018). Entre las evidencias de validez, se encontraron correlaciones estadísticamente significativas ($r = .92$) entre ambas pruebas (Killgore et al., 2005).

El DOT cuenta con dos formas (A y B) además de una parte donde el participante puede practicar lo que debe hacer en la prueba. Cada forma incluye nueve diseños: 5 diseños pequeños (cuadrículas de 2×2) y 4 diseños grandes (3×3), todos compuestos de cuadrados y triángulos en blanco y negro. Debajo de cada diseño se tiene una cuadrícula de respuesta del tamaño y forma idénticos. El objetivo es que los sujetos cumplimenten tantos cuadros en blanco como les sea posible en el tiempo establecido (dos minutos por cada forma), empleando el código numérico asociado (del 1 al 6) al inicio de la prueba en función de la forma (cuadrado o triángulo) y color (blanco o negro) que se muestra en la prueba. En total, los sujetos pueden cumplimentar hasta 56 espacios en blanco

según el modelo dado y el código numérico asociado. La puntuación total de la prueba se obtiene sumando los aciertos en los diseños completados, como es usual en este tipo de pruebas (Hoogendam, 2014).

Dado que las habilidades visoespaciales se consideran un importante predictor de la inteligencia general (Shea, Lubinski & Benbow, 2001), se aplicó una prueba de inteligencia en este estudio con el fin de tener en cuenta esta medida en los análisis.

Test de Factor g, Escala 3 (Cattell & Cattell, 2001). Tiene como finalidad evaluar la inteligencia fluida mediante un formato de prueba de velocidad mediante estímulos gráficos. Está conformada por 4 subtest: *Series* (13 ítems de series incompletas y de progresión de sus estímulos), *Clasificación* (14 de ítems para identificar figuras que no guardan relación con las demás), *Matrices* (13 de ítems, de completamiento de matrices eligiendo la más adecuada) y *Condiciones* (10 de ítems, en los que se debe de identificar los criterios correctos para representar figuras). El tiempo de aplicación efectivo es de 12´ 30" para responder a los 50 elementos totales de la escala. La puntuación total de la prueba se obtiene sumando el número de aciertos en los elementos realizados. De acuerdo a la adaptación española (Cattell & Cattell, 2001), se reportaron resultados favorables de validez convergente con otras pruebas multifactoriales de inteligencia verbal y no verbal, y con puntajes de habilidad intelectual desde en pruebas de personalidad (por ejemplo, el factor B del 16 PF-A)

Procedimiento

La aplicación se hizo de manera colectiva, en el horario de la mañana o tarde, y dentro del horario regular de clases. Para la administración se tuvo el apoyo de 15 profesionales psicólogos previamente capacitados (todos con experiencia en aplicación de pruebas psicológicas y que trabajaban en la misma institución de la muestra de participantes), quienes ingresaron a los salones de capacidad de 50 alumnos por aula aproximadamente. Se aplicó el formulario de consentimiento informado, una hoja demográfica, la prueba DOT y la Escala 3 de Cattell, manteniéndose constante este orden de aplicación en cada grupo evaluados. Ambos instrumentos se administraron siguiendo las recomendaciones e instrucciones disponibles en las publicaciones, conllevando aproximadamente unos 6 minutos en total para la prueba DOT y unos 20 minutos para la Escala 3 de Cattell.

Análisis de datos

Como análisis preliminar, se redujo la influencia de los valores extremos en los estadísticos descriptivos, mediante su identificación por medio de la *desviación absoluta de la mediana* (Median Absolute Deviation, MAD; Pham-Gia & Hung, 2001), porque es menos afectado que los puntajes estandarizados construidos con la media y la desviación estándar (Leys et al., 2013). Los sujetos con $MAD \leq 3.0$ fueron removidos,

pero se hizo especial énfasis en los puntajes muy bajos, debido que son más probables que ocurran consecuencia de patrones aleatorios de respuesta, respuesta descuidada o no comprometida con el proceso de evaluación (Merino & López, 2010). Esto parece razonable en el marco de la población muestreada, en que todos tienen estudios superiores en progreso y experiencia en exámenes de selección previos.

El análisis descriptivo de los datos se hizo mediante la obtención de estadísticos básicos, la observación del efecto piso y techo de los puntajes, y la evaluación de la distribución de probabilidad teórica que podría describir mejor la densidad de los datos en ambas formas del DOT. El efecto piso y/o techo se evaluó comparando el porcentaje de sujetos en el puntaje mínimo y máximo posible frente al criterio porcentual de 15% (Terwee et al., 2007), en que deberían ser menor al criterio. Respecto al ajuste distribucional, se aplicó el programa *-EasyFit* (Mathwave, 2010), que facilita explorar el ajuste estadístico de varias distribuciones teóricas a los datos; se usaron las pruebas Kolmogorov-Smirnov (KS; Smirnov, 1939), y Anderson y Darling (1954), proporcionadas por el mismo programa estadístico.

Al estimar varios parámetros psicométricos, la estrategia de análisis general fue tomar sus valores y compararlos con los previos estudios de validación (Killgore et al., 2005; Killgore & Gogel, 2014) para evaluar la generalización de los mismos. Se estimó la confiabilidad de formas paralelas mediante la correlación lineal de Pearson, y se lo comparó con lo obtenido en estudios previos de los autores (Killgore et al., 2005; Killgore & Gogel, 2014), mediante una prueba z de comparaciones de correlaciones para dos muestras independientes

(Howell, 2011); la magnitud de la diferencia correlacional fue estimada mediante el coeficiente, q (Cohen, 1992). Derivado de esta correlación, se construyó el error estándar de medición (Nunnally & Bernstein, 1995) para evaluar la precisión del puntaje, el mismo que debería ser menor que 0.5(DE) para una tolerable variación del puntaje observado (Norman et al., 2003; Wyrwich et al., 1999).

Luego, el efecto de la práctica se verificó comparando las medias de ambas formas mediante la prueba t de muestras dependientes y d como estimador de la magnitud de esta diferencia (Cohen, 1992); su intervalo de confianza fue obtenido usando una aproximación al error estándar asintótico:

$$\sqrt{\frac{d^2}{2(n-1)} + \frac{2(1-r_{12})}{n}} \quad (\text{Kline, 2004}).$$

Por otro lado, la comparación normativa de medias entre nuestros datos y los publicados por Killgore et al., siguió la misma estrategia (pruebas de significancia estadística y magnitud del efecto). Finalmente, se obtuvieron indicadores que puedan ayudar a la interpretación del cambio individual intra-sujeto entre las formas A y B. Primero, se evaluó la heterocedasticidad en las dos mediciones test-retest fue estimada mediante la correlación Pearson entre la diferencia absoluta intra-sujeto y la media intrasujeto de cada medición (Atkinson & Nevill, 1998); un coeficiente Pearson mayor a .30 sugiere que la diferencia no se mantiene similar a lo largo de la puntuación del DOT, y por lo tanto existe heterocedasticidad (Atkinson & Nevill, 1998). Complementariamente, el *coeficiente de variación entre medidas* (CV) fue computado como una medida de error de método (Portney & Watkins, 2015), estimado como

$CV = 100 \times 2(DE_d / \sqrt{2}) / (\bar{X}_1 + \bar{X}_2)$, donde DE_d es la desviación estándar de las diferencias (entre el pre-test y post-test), y \bar{X}_i es la media en ambas mediciones. Parece no existir algún estándar relevante para interpretar el CV.

Finalmente, se calculó un indicador de uso clínico (Beaton et al., 2001), el *mínimo cambio detectable* (MDC, llamado también diferencia mínima detectable, el cambio detectable más pequeño o cambio confiable; Portney & Watkins, 2015), que sirve como estimador del cambio significativo e independiente del error de medición (Beckerman et al., 2001; Portney & Watkins, 2015; Weir, 2005). Fue calculado como $z \times SEM \times \sqrt{2}$, donde z es el valor estandarizado proveniente de la curva normal, del porcentaje de confianza (90% = 1.645, 95% = 1.960, o 99% = 2.576), y SEM es el error estándar de medición ($DE\sqrt{1-ICC}$); este último usa la desviación estándar (DE) del puntaje directo y la correlación intraclase (ICC, dos vías, una medida, acuerdo absoluto). El ICC es un estimador de la confiabilidad (Beckerman et al., 2001; Portney & Watkins, 2015; Weir, 2005). El porcentaje de error relativo a la media de las dos medias de la forma A y B se estimó mediante SEM / M_c (Nair, Hornby & Behrman, 2012). Esta variación relativa a la media conjunta (M_c) de las evaluaciones en la forma A y B puede ser representada porcentualmente, mediante $MDC\% = (MDC / M_c) 100$ (Smidt et al., 2002). Un monto de error relativo de $MDC\% \leq 30\%$ es aceptable, y $MDC\% \leq 10\%$, es un nivel excelente (Smidt et al., 2002).

Resultados

Análisis preliminar

Se removieron 5 sujetos con valores extremos identificados en $z = -3.74$ ($Z_{MAD} = -3.50$). Luego, fueron removidos 5 sujetos más, quienes puntuaron cero en la forma B ($z = -4.29$), pero que lograron algún puntaje interpretable en la forma A. A menos que el sujeto intencionalmente no desee responder, el puntaje de cero en la forma B no es coherente con lo que se puede esperar luego de responder primero a la forma A, pues el sujeto debería obtener un puntaje por lo menos igual al obtenido en la forma A (efectivamente, estos sujetos obtuvieron puntajes en la forma A de alrededor de la media hasta ± 1 desviación estándar).

Indicadores distribucionales y descriptivos

Estadísticos descriptivos. En la identificación de la distribución teórica que mejor pueda describir la densidad estadística de los puntajes de la forma A del DOT, las pruebas de ajuste Anderson-Darling (AD) y KS arrojaron que la distribución Burr de 4 parámetros (Burr-4P; AD = 3.120, $p > .05$; KS = .054, $p > .02$) y Log-Logístico de 3 parámetros (LL-3P; AD = 2.50, $p > .05$; KS = .055, $p > .02$) pueden ser las mejores opciones de ajuste. Los parámetros para Burr-4P son $k = .587$ (forma), $\alpha = 2.9137E+5$ (forma), $\beta = 1.0491E+6$ (escala) y $g =$

-1.0491E+6 (localización); los parámetros para LL-3P son: $\alpha = 16.379$ (forma), $\beta = 74.0$ (escala) y $g = -37.97$ (localización). Respecto a la forma B, las pruebas de ajuste coincidieron en que la distribución de error generalizado (Coin, 2017) representa mejor a los datos ($AD = 4.356$, $p > .20$; $KS = .064$, $p > .20$).

Efecto de piso y techo. El puntaje mínimo posible (0) en las formas A y B no fue alcanzado en la presente muestra, sino los puntajes 8 y 15, respectivamente. El porcentaje de la muestra en ambos puntajes fue 0.1% ($n = 1$). El máximo puntaje posible (56) fue alcanzado por la muestra, con porcentajes de 3.9% ($n = 30$) y 4.5% ($n = 35$) respectivamente para la forma A y B. Estos porcentajes están por debajo del criterio de Terwee et al. (2007), y por lo tanto, adecuados; adicionalmente, son similares a los reportados por los resultados previos de validación de Killgore y colaboradores.

Confiabilidad de formas paralelas

La correlación Pearson entre ambas formas (A y B) fue .66 ($p < 0.01$), y generalmente puede considerarse de alta magnitud. Esta magnitud es ligeramente menor ($z = 4.39$, $p < 0.01$, $q = 0.31$) que lo reportado en el primer estudio de validación del DOT ($r = .81$, $n = 259$; Killgore et al., 2005); pero frente al segundo de estudio ($r = .90$, $n = 31$; Killgore & Gogel, 2014), la diferencia puede considerarse grande, $z = 3.38$, $p < 0.01$, $q = .66$. La correlación intraclase de dos vías aleatorias, estimada como el acuerdo absoluto para medidas únicas, fue .53

(IC 95%), un nivel que puede considerarse entre aceptable (Fleiss, 1986).

Mínimo cambio detectable

Heterocedasticidad. La correlación Pearson fue .07 ($n = 775$, IC 95%) sugiriendo que la distribución de errores no se diferencia de la producida por efecto aleatorio en los niveles del puntaje del DOT. Esta magnitud no fue estadísticamente diferente ($r = -.01$, $n = 785$, IC 95%) con los datos completos que incluían los 10 sujetos extremos más bajos. Por otro lado, en la métrica porcentual, el CV fue 12.37%.

Mínimo cambio detectable. Usando la estimación ICC, el error estándar de medición fue 5.64 (SEM% = 15.43 relativo a la media) y el MDC fue 7.98. En los niveles de confianza de 90%, 95% y 99%, el MDC fue 13.13 (~ 13), 15.64 (~ 16) y 20.56 (~ 21), respectivamente. Porcentualmente, el monto de error aleatorio relativo (MDC%) fue, para cada una de las estimaciones en los niveles de confianza (90%, 95% y 99%, respectivamente), 33.13%, 39.47% y 51.88%. En el mejor escenario del ICC, esto es, el límite superior de confianza del ICC calculado ($\text{sup}_{\text{ICC}} = .73$; SEM = 4.21; SEM% = 11.51%), el MDC fue 9.79 (~ 10), 11.67 (~ 12) y 15.34 (~ 15); y sus correspondientes MDC% fueron: 24.71%, 29.44% y 38.70%.

Efecto de la práctica

Debido a la fuerte asimetría de las distribuciones, se usó el método bootstrap (1000 replicaciones) para obtener valores de significancia estadística similarmente comparables similares respecto al error Tipo I, frente al método asintótico para la *t* Student dependiente. Los resultados entre ambos métodos fueron indistinguibles, así que se presentarán únicamente los resultados asintóticos. En el análisis comparativo intra-grupos, la diferencia de medias ($M_{dif} = 6.118$, $DE_{dif} = 6.93$, $r = .66$) entre la forma A y B fue estadísticamente significativa, $t(774) = 24.56$; los puntajes en la forma B fueron moderadamente más altos ($d = 0.58$), como era de esperar al introducirse el efecto de la práctica luego de aplicar la forma A primero. Al comparar este efecto de la práctica con el que fue reportado por Killgore et al. (2005) ($M_{dif} = 3.99$, $DE_{dif} = 4.62$, $n = 253$), la media de nuestro grupo se halló que es moderadamente mayor que lo reportado en ese estudio ($t[1191] = 7.44$, $p < 0.001$; $d = .72$). Cuando se lo compara con el efecto de la práctica ($M_{dif} = 5.13$, $DE_{dif} = 4.05$, $n = 31$) reportado en el segundo estudio de validación de Killgore y Gogel (2014), ambos pueden considerarse similares aunque levemente superior en la muestra del presente estudio ($t[804] = 1.17$, $p > .05$; $d = .25$).

Comparaciones normativas

En una comparación entre-grupos, los estadísticos presentados en la Tabla 2 para la forma A se compararon con los resultados de Killgore et al. (2005) en su primer estudio ($M = 44.25$, $DE = 7.9$, $n = 259$), observándose que el desempeño

de la presente muestra fue levemente menor, $t(1032) = 2.40$, $p < .05$, $d = .36$). Luego, comparado con la forma A del estudio de Killgore y Gogel (2014) ($M = 35.9$, $DE = 8.06$, $n = 30$), se hallaron diferencias triviales, $t(813) = .44$, $p > .05$, $d = .08$.

Por otro lado, con respecto a la forma B ($M = 43.61$, $DE = 7.83$, $n = 259$) de Killgore et al. (2005), las diferencias no fueron estadísticamente significativas, $t(1042) = 1.494$, $d = .11$. Igualmente, la diferencia con el desempeño en la pequeña muestra de Killgore y Gogel (2014; $M = 41.03$, $DE = 9.1$, $n = 30$) fue trivial, $t(813) = 1.02$, $p > .10$, $d = .19$.

Tabla 2.- Resultados descriptivos del DOT (Forma A y B)

	Forma A	Forma B
Media	36.576	42.695
Mediana	36	43
Desviación estándar	8.243	8.765
Mínimo	8	15
Máximo	56	56
Asimetría (cc = 0.08)	.26	-.20
Curtosis (cc = 0.17)	.61	-.58

Relación con inteligencia fluida y edad

Con la edad ($n = 716$), la relación de esta con la forma A ($r = -.07$, IC 95%) fue estadísticamente significativa ($p = .041$) pero de magnitud trivial; con la forma B, la relación también fue

-estadísticamente significativa ($p < .01$) pero baja ($r = -.13$, IC 95%). La diferencia entre ambas correlaciones dependientes, que contienen un elemento común (la edad), fue estadísticamente significativa ($z = .34$; $p = .046$; Steiger, 1980), pero de magnitud trivial ($q = .05$). Por otro lado, la edad covarió en con el DOT en un rango de trivial a pequeño efecto, sugiriendo que el incremento (o decremento) de la edad apenas está involucrado en el incremento.

Discusión

Los parámetros presentados para uso clínico del DOT, proporcionan referencias relevantes para aplicarlos en la comparación del desempeño visoespacial entre las dos versiones (A y B). En este marco clinicométrico, de especial utilidad es el MDC obtenido, porque la interpretación de las diferencias en la forma A y B reduce el efecto del error aleatorio, y tiene el MDC un valor clínico porque se enfoca en la variabilidad individual en lugar de la variabilidad entre grupos (Beaton et al., 2001). En el presente estudio, este parámetro puede considerarse apropiado debido que la heterocedasticidad no fue estadísticamente diferente de la producida por efecto aleatorio, y por lo tanto el MDC puede ser interpretado con confianza en los diferentes niveles de diferencia de puntuación del DOT.

La variación hallada en el MDC (medida por el MDC%) fue superior al máximo sugerido para el monto aceptable de variación entre medidas repetidas (Smidt et al., 2002). Debido

que la variabilidad registrada por el MDC incluye dos tipos de variabilidad (Beckerman et al., 2000), esto es, la variabilidad por el desempeño (variación sistemática) y por error de medición (variabilidad aleatoria), entonces la interpretación del cambio del puntaje entre la forma A y B debe interpretarse en el contexto de los resultados. Específicamente, nuestros resultados en ambas medidas aplicadas, inmediatamente una después de la otra, representan un cambio sistemático en el desempeño (mejor en B comparado con A), lo cual está asociado a la menor confiabilidad test-retest hallada en nuestro estudio ($r = .66$), comparada con el estudio principal de Killgore et al. ($r = .81$), y al acuerdo absoluto estimado por el ICC, esto es, .53. La variación de la medición entre ambas formas, una aplicada inmediatamente después de la otra, fue de 12%. Esta magnitud parece pequeña, pero no hay estándar absoluto de referencia para interpretar el monto de variación; esto es condicional al juicio del investigador para identificar el máximo monto aceptable (Portney & Watkins, 2015), o a la evidencia acumulativa sistematizada por un meta-análisis.

Asociado con el cambio de los puntajes en la forma A y B del DOT, las diferencias establecidas por las diferencias de medias (prueba t para muestras dependientes) es un indicador de sesgo sistemático (Kazis et al., 1989), pero necesariamente no indica un problema para la interpretación de la diferencia forma A – forma B del DOT, porque la mejora del desempeño en la forma B es esperable debido al efecto de la práctica, y es una tendencia similar a lo hallado en los estudios del autor del DOT (Killgore et al., 2005; Killgore & Gogel, 2014). Las diferencias halladas pueden interpretarse como de significancia con valor clínico, dado que la magnitud de esta

diferencia no fue pequeña ni trivial, sino moderada. Aunque la magnitud de las diferencias se comprende mejor en el contexto específico del uso del instrumento, y los estándares de interpretación están condicionadas a ellas (Merino & Cópez-Lonzoy, 2018), la conocida taxonomía de Cohen (1992) puede ser aplicada al presente contexto como en otros estudios relacionados con la neuropsicología (e.g., Cohen et al., 1998; Lee et al., 2017).

Al examinar la forma distribucional de los puntajes directos de la forma A y B del DOT, es claro que se requieren más parámetros para definir con mayor precisión la distribución de sus puntajes, y se requiere modelar la asimetría y la dispersión del mismo. Las distribuciones que parecen mejor ajustarse (Burr-4P y Log-logística-3P para la forma A, y Error y Valor Extremo Generalizado) se caracterizan por reducir el modelamiento de la simetría distribucional. Los puntajes de la forma B parece solo requerir 3 parámetros (forma, escala y localización), sugiriendo que la larga cola hacia la izquierda y la mayor densidad hacia la derecha de la distribución es más clara y con menor incertidumbre.

En resumen, mientras las distribuciones sugeridas proponen mayor flexibilidad en el ajuste (por ejemplo, la distribución Burr-4P; Choi & Min, 2011) o son incluidas con alguna regularidad en psicología (por ejemplo, log-logística; Aryal, 2013), en conjunto se requieren más parámetros para un mejor ajuste y estimación de percentiles de desempeño en el DOT; comparados con los dos parámetros requeridos para ajustar a una distribución normal, esta puede ser insuficiente. Dado que requieren más parámetros comparados con la distribución normal, en las distribuciones teóricas detectadas como

posibles modelos de distribución para los puntajes del DOT, los percentiles pueden ser estimados con más exactitud. Nuestros resultados sobre la forma distribucional también apuntan a confirmar que la no normalidad de los puntajes del DOT parece ser una tendencia “normal”, y no la excepción en relación a la tendencia distribucional de los datos conductuales (Blanca, et al. 2013; Cain, Zhang & Yuan, 2017; Ho & Yu, 2014; Micceri, 1989).

En lo referente a los análisis efectuados al comparar los resultados obtenidos en el presente trabajo con estudios previos respecto al efecto de la práctica, encontramos que al comparar con el estudio de Killgore et al. (2005) la muestra peruana obtiene menor puntuación en su conjunto. Este resultado es esperable dado que Killgore afirma que los 418 participantes de su estudio inicial pertenecían a una universidad privada muy selectiva donde las admisiones están marcadas por gran competitividad, lo que podría indicar que la muestra de ese estudio presenta un sesgo hacia el extremo superior de la capacidad cognitiva (Killgore y Gogel, 2014), por lo que proponen que se comparen con el estudio de 2013 al tratarse del resultado normativo. Con esta sugerencia, al comparar las diferencias forma A-forma B en la muestra peruana con la del estudio posterior, las del presente estudio fueron más grandes que las diferencias forma A-forma B en los americanos, lo cual parece indicar un mayor efecto de la práctica que el obtenido en la muestra americana. Estas diferencias podrían deberse a los rangos de edad empleados en ambos estudios y al tamaño muestral de los mismos, ya que en el estudio de Killgore y Gogel (2014) que sirve como grupo normativo de 18 a 45 años, solo contaban con 61 sujetos, y el

presente estudio cuenta con más de 700 sujetos con una edad de 18 a 30 años. Este aspecto es relevante, dado que existe aún controversia en cuanto a la relación entre la edad y el rendimiento en habilidades visoespaciales en la prueba DOT, como veremos más adelante.

Al comparar el rendimiento en ambas formas del DOT del presente estudio y las investigaciones previas, encontramos que, en la forma A, el desempeño de la presente muestra fue levemente menor que en el estudio de 2005 y se hallaron diferencias triviales levemente favorables a la muestra de este estudio en comparación al de 2013. Estos resultados podrían interpretarse igual que las diferencias forma A y B descritas previamente. Analizando la forma B cabe reseñar que no hay diferencias entre las muestras estudiadas, probablemente debido a la práctica en la resolución de la tarea, teniendo en cuenta además que en la prueba DOT se cometen pocos errores (Killgore y Gogel, 2014). Alguna comparación normativa adicional entre nuestros resultados y los de otros estudios (por ejemplo, Burggraaf et al., 2017) ya que en este no se emplearon los dos minutos de la prueba por cada forma, como establece Killgore et al. (2005).

Finalmente, en relación a la edad, los resultados de este estudio sugieren que el incremento (o decremento) de la edad apenas está involucrado en el incremento en la puntuación del DOT. Estos hallazgos están en consonancia con los encontrados por Killgore y Gogel (2013) en los que no se encontró correlación entre ambas variables. Sin embargo, Killgore et al. (2005) halló correlación estadísticamente significativa en una muestra de 41 pacientes neurológicos entre ambas variables, incluyendo edades de 18 a 76 años; y en un estudio con

adolescentes de 11 a 20 años sí se encontró correlación positiva entre la edad y la puntuación del DOT (Burggraaf et al., 2017). Estas diferencias podrían sugerir una relación compleja entre la edad y las habilidades visoespaciales, asociadas al desarrollo cerebral en el caso de los adolescentes y juventud, y por ello representando una relación directa, y a partir de ahí, dado que el cerebro alcanzaría un buen desarrollo de las estructuras cerebrales implicadas, la edad no correlacionaría con el rendimiento, a excepción de que hubiesen patologías neurológicas que intermedien en esa relación. Futuros estudios deberían explorar esta cuestión, con más poblaciones y tamaño muestral, ya que resulta interesante esclarecer el tipo de relación entre ambas variables. Esto es así porque conocer cómo es la ejecución en las pruebas de tipo neuropsicológico, como la DOT, en la población normal, y en diferentes grupos de edad, nos permitirá entender e instaurar los perfiles disonantes con esa normalidad, para poder diagnosticar e intervenir con mayor eficacia (Cruz, et al. 2018).

Respecto a las limitaciones del estudio, una de ellas es la muestra del estudio, que no garantiza ser representativa de la población normal peruana. Aunque la muestra no puede ser considerada representativa de la población peruana en general, ni de la población universitaria en particular, los estadísticos normativos obtenidos pueden servir en tres tipos de representaciones: primero, la muestra de estudio puede representar un grupo particular, regional y temporalmente, y sus características pueden ser comparadas con las obtenidas en otros estudios para verificar si los resultados son invariantes o producen variabilidad en los puntajes del DOT. Segundo, los resultados representan la variabilidad del constructo que plausiblemente

varíen en intensidad (en rango de bajo hasta alto habilidad de procesamiento visuoespacial) en esta muestra. Y tercero, la muestra universitaria de nuestro estudio representa también al grupo de adultos que se dispersan en actividades principalmente académicas, y/o que contemporáneamente trabajan, y por lo tanto representa aproximadamente a un grupo activo predominante en Perú (Instituto Nacional de Estadística e Informática, 2007; Tavera, Oré & Málaga, 2017). Finalmente, otra principal limitación es que se requieren establecer más evidencias de validez no indagadas en este estudio, como la asociación con otros constructos, la predictividad respecto al funcionamiento cognitivo o la memoria, entre otros. Estudios futuros han de seguir recabando propiedades psicométricas del DOT en diferentes contextos.

Referencias Bibliográficas

1. Aryal, G. R. (2013). Transmuted Log-Logistic distribution. *Journal of Statistics Applications & Probability*, 2(1), 11–20. <https://doi.org/10.12785/jsap/020102>
2. Astrand, E., Wardak, C. & Ben Hamed, S. (2014). Selective visual attention to drive cognitive brain-machine interfaces from concepts to neurofeedback and rehabilitation applications. *Frontiers in Systems Neuroscience*, 8, 144. <https://doi.org/10.3389/fnsys.2014.00144>
3. Atkinson, G. & Nevill, A. M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine*,

26, 217–238. <https://doi.org/10.2165/00007256-199826040-00002>

4. Beaton, D. E., Bombardier, C., Katz, J. N. & Wright, J. G. (2001). A taxonomy for responsiveness. *Journal of Clinical Epidemiology*, 54(12), 1204–1217. [https://doi.org/10.1016/s0895-4356\(01\)00407-3](https://doi.org/10.1016/s0895-4356(01)00407-3)
5. Beckerman, H., Roebroeck, M. E., Lankhorst, G. J., Becher, J. G., Bezemer, P. D. & Verbeek, A. L. (2001). Smallest real difference, a link between reproducibility and responsiveness. *Quality Life Research*, 10(7), 571–578. <https://doi.org/10.15200/winn.146857.76572>
6. Blanca, M. J., Arnau, J., Bono, R., López-Montiel, D. & Bendayan, R. (2013). Skewness and kurtosis in real data samples. *Methodology. European Journal of Research Methods for the Behavioral and Social Sciences*, 9(2), 78–84. <https://doi.org/10.1027/1614-2241/a000057>
7. Boggio, P. S., Ferrucci, R., Mameli, F., Martinis, D., Martinis, O., Vergari, M. & Priori, A. (2012). Prolonged visual memory enhancement after direct current stimulation in Alzheimer's disease. *Brain Stimulation*, 5(3), 223–230. <https://doi.org/10.1016/j.brs.2011.06.006>
8. Boggio, P. S., Khoury, L. P., Martinis, O. E., de Macedo, E. C. & Fregni, F. (2009). Temporal cortex direct current stimulation enhances performance on a visual recognition memory task in Alzheimer disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, 80(4), 444–447. <https://doi.org/10.1136/jnnp.2007.141853>
9. Borsting, E. (1996). *Visual Perception and Reading. Vision Reading*, 149–176.

10. Burggraaf, R., Frens, M., Hooge, I. & van der Geest, J. (2016). A quick assessment of visuospatial abilities in adolescents using the design organization test (DOT), *Applied Neuropsychology: Child*, 5(1), 44–49. <https://doi.org/10.1080/21622965.2014.945114>
11. Burggraaf, R., Frens, M. A., Hooge, I. T. C., & van der Geest, J. N. (2018). Performance on tasks of visuospatial memory and ability: A cross-sectional study in 330 adolescents aged 11 to 20. *Applied Neuropsychology: Child*, 7(2), 129–142. <https://doi.org/10.1080/21622965.2016.1268960>
12. Cain, M. K., Zhang, Z. & Yuan, K. H. (2017). Univariate and multivariate skewness and kurtosis for measuring nonnormality: Prevalence, influence and estimation. *Behavior Research Methods*, 49(5), 1716–1735. <https://doi.org/10.3758/s13428-016-0814-1>
13. Calvo, P. (2017). Modelo de interacción entre el dibujo y las actividades sensomotrices (MIDAS): Fomentando la inteligencia espacial. *Innovaciones Educativas*, 19(26), 50–64. <https://doi.org/10.22458/ie.v19i26.1853>
14. Cattell, R. B. & Cattell, A. K. S. (2001). *Tests de Factor «g», Escalas 2 y 3*. Madrid: TEA Ediciones.
15. Chamberlain, R., Brunswick, N., Siev, J. & McManus, I. C. (2018). Meta-analytic findings reveal lower means but higher variances in visuospatial ability in dyslexia. *British Journal of Psychology*, 109(4), 897–916. <https://doi.org/10.1111/bjop.12321>
16. Charter, R. A. (1996). Revisiting the standard errors of measurement, estimate, and prediction and their application to test scores. *Perceptual and Motor*

- Skills*, 82, 1139–1144.
<https://doi.org/10.2466/pms.1996.82.3c.1139>
17. Chastan, N., Bair, W. N., Resnick, S. M., Studenski, S. A. & Decker, L. M. (2019). Prediagnostic markers of idiopathic Parkinson's disease: Gait, visuospatial ability and executive function. *Gait & Posture*, 68, 500–505. <https://doi.org/10.1016/j.gaitpost.2018.12.039>
 18. Choi, P. & Min, I. (2011). A comparison of conditional and unconditional approaches in value-at-risk estimation. *Journal of Japanese Economic Association*, 62, 99–115. <https://doi:10.1111/j.1468-5876.2010.00456.x>
 19. Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159.
 20. Cohen, R. M., Nordahl, T. E., Semple, W. E., Andreason, P. & Pickar, D. (1998). Abnormalities in the distributed network of sustained attention predict neuroleptic treatment response in schizophrenia. *Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology*, 19, 36–47. [https://doi.org/10.1016/s0893-133x\(97\)00201-7](https://doi.org/10.1016/s0893-133x(97)00201-7)
 21. Coin, D. (2017) A goodness-of-fit test for generalized error distribution. *Communications in Statistics - Theory and Methods*, 46, 23, 11485–11499. <https://doi.org/10.2139/ssrn.2946894>
 22. Cronin-Golomb, A. & Amick, M. (2001). Spatial abilities in aging, Alzheimer's disease, and Parkinson's disease. En P. R. Hof & C. V. Mobbs (Eds.), *Functional neurobiology of aging* (pp. 517–529). San Diego, CA:

Academic Press.

23. Cruz, J. S., Cruz, M. E. B., Anacleto, A. G., Amaya, L. I. L., Aguillón, C., Zamorano, E. R. & Martínez, J. B. B. (2018). Efectos de la edad sobre el tipo de error en la copia directa e inmediata de la Figura de Taylor. *Cuadernos de Neuropsicología*, 12(3), 5.
24. Fleiss, J. L. (1986). *The Design and Analysis of Clinical Experiments*. Nueva York: Wiley. <https://doi.org/10.1002/9781118032923.ch1>
25. Gogel, H. & Killgore, W. D. (2014, Febrero). *A psychometric validation of the Design Organization Test (DOT) in a healthy sample*. Forty Second Annual Meeting International Neuropsychological Society, February 12–15, Seattle, Washington, USA.
26. Hampstead. B. M, Brown. G. S. & Hartley. F. (2014). Transcranial direct current stimulation modulates activation and effective connectivity during spatial navigation. *Brain Stimulation*, 7(2), 314–324. <https://doi.org/10.1016/j.brs.2013.12.006>
27. Ho, A. D. & Yu, C. C. (2014). Descriptive statistics for modern test score distributions: skewness, kurtosis, discreteness, and ceiling effects. *Educational and Psychological Measurement*, 75(3), 365-388. <https://doi.org/10.1177/0013164414548576>
28. Hofman, A., van Duijn, C. M., Franco, O. H., Ikram, M. A., Janssen, H. L., Klaver, C. C. & Witteman, J. C. (2011). The Rotterdam Study: 2012 objectives and design update. *European Journal of Epidemiology*, 26, 657–686. <https://doi.org/10.1007/s10654-011-9610-5>
29. Hoogendam, Y. (2014). *The role of brain aging in*

cognition and motor function. Netherland: Erasmus University Rotterdam.

30. Howell, D. C. (2011). *Statistical methods for Psychology* (7th ed.). Belmont, CA: Wadsworth, Cengage Learning.
31. Instituto Nacional de Estadística e Informática, INEI. (2007). *Censo IX de población*. Lima: INEI.
32. Kazis, L. E., Anderson, J. J. & Meenan, R. F. (1989). Effect sizes for interpreting changes in health status. *Medical Care*, 27, S178–S189.
33. Killgore, W. D. & Gogel, H. (2014). The Design Organization Test: Further demonstration of reliability and validity as a Brief Measure of Visuospatial Ability. *Applied Neuropsychology Adult*, 21(4), 1–13. <https://doi.org/10.1080/23279095.2013.811671>
34. Killgore, W., Glahn, D. & Casasanto, D. (2005), Development and validation of the design organization test (dot): a rapid screening instrument for assessing visuospatial ability. *Journal of Clinical and Experimental Neuropsychology*, 27, 449–459. <https://doi.org/10.1080/13803390490520436>
35. Kline, R. B. (2004). *Beyond significance testing: Reforming data analysis methods in behavioral research*. Washington, DC: American Psychological Association.
36. Kolmogorov, A. N. (1933). Sulla determinazione empirica di una legge di distribuzione. *Giornale Dell' Istituto Italiani degli Attuari*, 4, 83–91.
37. Koppelmans, V., Breteler, M. M., Boogerd, W., Seynaeve, C., Gundy, C. & Schagen, S. B. (2012).

- Neuropsychological performance in survivors of breast cancer more than 20 years after adjuvant chemotherapy. *Journal of Clinical Oncology*, 30, 1080–1086. <https://doi.org/10.1200/JCO.2011.37.0189>
38. Kosmidis, M., Economou, A., Liozidou, A. & Yiannis, G. (2014). B-88: Neurocognitive Correlates of Driving Behavior. *Archives of Clinical Neuropsychology*, 29(6), 569-570. <https://doi.org/10.1093/arclin/acu038.176>
39. Lanfranchi, S., Mammarella, I. C. & Carretti, B. (2015). Spatial-simultaneous working memory and selective interference in Down syndrome. *Child Neuropsychology*, 21(4), 481–489. <https://doi.org/10.1080/09297049.2014.913557>
40. Lee, P., Lu, W.-S., Liu, C.-H., Lin, H.-Y. & Hsieh, C.-L. (2017). Test-Retest reliability and minimal detectable change of the D2 Test of attention in patients with schizophrenia. *Archives of Clinical Neuropsychology*, 1, 1–9. <https://doi.org/10.1093/arclin/acx123>
41. Leys, C., Ley, C., Klein, O., Bernard, P. & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, 49, 764-766. <https://doi.org/10.1016/j.jesp.2013.03.013>
42. López-Velásquez, N. D., Coronado López, J., Herrera-Pino, J., Soto-Añari, M. & Ferrel Ortega, F. R. (2016). Efecto del aprendizaje de morfología en el procesamiento visoespacial de estudiantes universitarios chilenos. *Revista Interamericana de Psicología/Interamerican Journal of Psychology*, 50(3), 441–448.

43. Maguire, E. A., Woollett, K. & Spiers, H. J. (2006). London taxi drivers and bus drivers: a structural MRI and neuropsychological analysis. *Hippocampus*, 16(12), 1091–1101. <https://doi.org/10.1002/hipo.20233>
44. Mathwave (2010). *EasyFit Professional 5.5* [Computer software]. Westminster, CA: Psytek Services
45. Merino, C. & López, J. (2010). RAN: un programa informático para estimar patrones de respuesta aleatorios en test objetivos. -*Psicothema*, 22(2), 334–339.
46. Merino-Soto, C. & Cópez-Lonzoy, A. (2018). Un problema con diferentes rostros: arbitrarios niveles de tamaño del efecto. *Enfermería Clínica*, 28(6), 402–403. <https://doi.org/10.1016/j.enfcli.2018.05.00>
47. Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin*, 105(1), 156–166. <https://doi.org/10.1037//0033-2909.105.1.156>
48. Mitolo, M., Gardini, S., Caffarra, P., Ronconi, L. & Venneri, A. (2015) Relationship between spatial ability, visuospatial working memory and self-assessed spatial orientation ability: a study in older adults. *Cognitive Processing*, 16(2), 165–176. <https://doi.org/10.1007/s10339-015-0647-3>
49. Nair, P. M., Hornby T, G. & Behrman, A. L. (2012). Minimal detectable change for spatial and temporal measurements of gait after incomplete spinal cord injury. *Topics in spinal cord injury rehabilitation*, 18(3), 273–281. <https://doi.org/10.1310/sci1803-273>
50. Ness, D., Farenga, S. & Garofalo, S. (2017). *Spatial*

Intelligence. Why It Matters from Birth through the Lifespan. Nueva York, NY: Routledge.

51. Norman, G. R., Sloan, J. A. & Wyrwich, K. W. (2003). Interpretation of changes in health-related quality of life: The remarkable universality of half a standard deviation. *Medical Care*, 41, 582–592. <https://doi.org/10.1097/01.MLR.0000062554.74615.4>
52. Nunnally, J. C. & Bernstein, I. J. (1995). *Teoría Psicométrica* (3ª ed.). México, D.F.: McGraw-Hill Latinamericana.
53. Pérez, M. A., Molina, D. & Gómez, M. (2016). La intervención neuropsicológica en el tratamiento interdisciplinar para el TDAH. *Neuropsicología*, 1(2), 2.
54. Peña-Casanova, J., Guardia, J, Bertran- Serra, I. & Jarne, A. (1997). Versión abreviada del Test Barcelona (I): subtests y perfiles normales. *Neurología*, 12,19–23.
55. Pham-Gia, T. & Hung, T. L. (2001). The mean and median absolute deviations. *Mathematical and Computer Modelling*, 34(7-8), 921–936. [https://doi:10.1016/S0895-7177\(01\)00109-1](https://doi:10.1016/S0895-7177(01)00109-1)
56. Portney, L. G. & Watkins, M. P. (2015). *Foundations of clinical research: Applications to practice (3th Ed.)*. Philadelphia, PA: F. A. Davis Company.
57. Merchán, M. S., Henao, J. L. (2011). Influencia de la percepción visual en el aprendizaje. *Ciencia & Tecnología para la Salud Visual y Ocular*, 9(1), 93–101. <https://doi.org/10.19052/sv.221>
58. Prince, C. & Daniel, M. (2014). A-59: Prompting for

- delayed visuospatial memory: effect on memory test performance. *Archives of clinical neuropsychology*, 29(6), 525. <https://doi.org/10.1093/arclin/acu038.59>
59. Seijas, R. (2015). Atención, memoria y funciones ejecutivas en los trastornos del espectro autista: ¿cuánto hemos avanzado desde Leo Kanner? *Revista Asociación Española de Neuropsiquiatría*, 35(127), 573–586. <https://doi.org/10.4321/S0211-57352015000300009>
60. Shea, D. L., Lubinski, D. & Benbow, C. P. (2001). Importance of assessing spatial ability in intellectually talented young adolescents: A 20-year longitudinal study. *Journal of Educational Psychology*, 93(3), 604–614. <https://doi.org/10.1037/0022-0663.93.3.604>
61. Smidt, N., Van der Windt, D. A., Assendelft, W. J., Mourits, A. J., Devill, W. L. & de Winter, A. F. (2002). Interobserver reproducibility of the assessment of severity of complaints, grip strength, and pressure pain threshold in patients with lateral epicondylitis. *Archives of Physical Medicine and Rehabilitation*, 83, 1145–1150. <https://doi.org/10.1053/apmr.2002.33728>
62. Smirnov, N. V. (1939). Estimation of deviation between empirical distribution functions in two independent samples (en ruso). *Bulletin of Moscow University*, 2, 3–16.
63. Solovieva, Y., Abundis, A. C. H., Balbuena, I. M. & Gonzalez, H. J. P. (2017). Cambios en las funciones viso-espaciales como efecto de método del dibujo en preescolares menores de 4 a 6 años de edad/changes in visual spatial functions as an effect of the drawing method in preschoolers under 4 to 6 years of age.

- European Journal of Education Studies*, 3(11), 523–541. <https://doi.org/10.5281/zenodo.1067844>
64. Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, 87, 245–251. <https://doi.org/10.1037//0033-2909.87.2.245>
65. Tavera, J. A, Oré, T. & Málaga, R. (2017). La dinámica de la población que no estudia ni trabaja en el Perú: quiénes son, cómo son y cómo han cambiado. *Apuntes*, 44(80), 5–49. <https://doi.org/10.21678/apuntes.80.903>
66. Terwee, C. B., Bot, S. D. M., de Boer, M. R., van der Windt, D. A. W. M., Knol, D. L., Dekker, J., Bouter, L. M. & de Vet, H. C. W. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, 60(1), 34–42. <https://doi.org/10.1016/j.jclinepi.2006.03.012>
67. Tseng, P., Hsu, T. Y., Muggleton, N. G., Tzeng, O. J., Hung, D. L. & Juan, C. H. (2010). Posterior parietal cortex mediates encoding and maintenance processes in change blindness. *Neuropsychologia*, 48(4), 1063–1070. <https://doi.org/10.1016/j.neuropsychologia.2009.12.005>
68. Vleugels, L., Christophe, L., Nunen, A., Nachtergaele, S., Ketealer, P., Charlier, M. & Vandebussche, E. (2000). Visuo-perceptual impairment in multiple sclerosis patients diagnosed with neuropsychological task. *Multiple Sclerosis*, 6, 241–54. <https://doi.org/10.1191/135245800678827879>
69. Von Allmen, D.Y., Wurmitzer, K. & Klaver, P. (2014).

Hippocampal and posterior parietal contributions to developmental increases in visual short-term memory capacity. *Cortex*, 59, 95–102. <https://doi.org/10.1016/j.cortex.2014.07.010>

70. Warrington, E. K. & James, M. (1991). *Visual Object and Space Perception Battery*. Suffolk, UK: Thames Valley Test Co.
71. Weir, J. P. (2005). Quantifying Test-Retest Reliability Using the Intraclass Correlation Coefficient and the SEM. *Journal of Strength and Conditioning Research*, 19(1), 231–240. <https://doi.org/10.1519/15184.1>
72. Wyrwich, K., Nienaber, N., Tierney, W. & Wolinsky, F. (1999). Linking clinical relevance and statistical significance in evaluating intra-individual changes in health-related quality of life. *Medical Care*, 37(4), 469–478. <https://doi.org/10.1097/00005650-199905000-00006>
73. Zacks, J. M. & Michelon, P. (2005). Transformations of visuospatial images. *Behavioral and Cognitive Neuroscience Reviews*, 4(2), 96–118. <https://doi.org/10.1177/1534582305281085>

Autor de correspondencia: *sikayax@yahoo.com.ar

Autor de correspondencia: **dalarcon@usmp.pe

Recibido: 21 de marzo 2019

Revisado: 10 de setiembre 2019

Aceptado: 13 de setiembre 2019