

UTILIZACIÓN DE LOS TEST EN ESPAÑA: EL PUNTO DE VISTA DE LOS PSICÓLOGOS

Muñiz, José

UTILIZACIÓN DE LOS TEST EN ESPAÑA: EL PUNTO DE VISTA DE LOS PSICÓLOGOS

Papeles del Psicólogo, vol. 41, núm. 1, 2020

Consejo General de Colegios Oficiales de Psicólogos, España

Disponible en: <https://www.redalyc.org/articulo.oa?id=77862536001>

DOI: <https://doi.org/10.23923/pap.psicol2020.2921>

UTILIZACIÓN DE LOS TEST EN ESPAÑA: EL PUNTO DE VISTA DE LOS PSICÓLOGOS

Test Use in Spain: The psychologists' viewpoint

José Muñiz
Universidad de Oviedo, España

DOI: <https://doi.org/10.23923/pap.psicol2020.2921>
Redalyc: <https://www.redalyc.org/articulo.oa?id=77862536001>

Recepción: 15 Noviembre 2019
Aprobación: 02 Diciembre 2019

RESUMEN:

Los test son los instrumentos de medida más utilizados por los psicólogos para obtener datos que les ayuden a tomar decisiones sobre las personas. Conocer las opiniones de los profesionales de la psicología acerca de la práctica de los test es fundamental para plantear acciones encaminadas a mejorar su uso en nuestro país. El objetivo del presente trabajo es conocer con detalle lo que piensan los psicólogos españoles sobre el uso de los test. Se aplicó un cuestionario de 30 ítems sobre distintos aspectos de la práctica de los test. Respondieron 1.248 colegiados, 73,7% mujeres, con una edad media de 46,31 años y una desviación típica de 11,08. Los resultados muestran unas actitudes favorables de los psicólogos al uso de los test, si bien se ponen de manifiesto algunos aspectos que es necesario mejorar. Se analizan los resultados y se comentan algunas ideas sobre el futuro de los test.

PALABRAS CLAVE: Test, Uso de test, Opiniones de los psicólogos, Actitudes hacia los test.

ABSTRACT:

Tests are the measuring instruments most used by psychologists to obtain data that help them make decisions about people. It is essential to understand the opinions of psychology professionals regarding the practical implementation of tests, in order to propose and organize actions aimed at improving their use. The objective of this research is to understand in detail what Spanish psychologists think about the use of tests in our country. A 30-item questionnaire on the different aspects of testing practices was administered. The questionnaire was answered by 1,248 members of the Spanish Psychological Association, 73.7% women, with an average age of 46.31 years, and a standard deviation of 11.08. The results show favorable attitudes of Spanish psychologists to the use of tests in the exercise of their profession, although some aspects that need to be improved are presented. Results are discussed, and some ideas about the future of tests are analyzed.

KEYWORDS: Test, Use of tests, Psychologists. opinions, Attitudes towards tests.

PROCESO EVALUATIVO

Los psicólogos utilizan los test para obtener muestras de conducta que les ayuden a tomar decisiones sobre las personas, bien sea en el ámbito clínico, laboral, educativo, jurídico, u otros muchos en los que desarrollan su labor. Los test no tienen vida propia, no son autónomos, no toman las decisiones por sí mismos, las decisiones las toman los profesionales, eso sí, apoyados en datos objetivos, en evidencias, y es ahí cuando entran en escena los test. Como bien nos alertó tempranamente Paul Meehl (1954) en su inquietante librito, como él lo llamaba, los psicólogos deben abstenerse de hacer diagnósticos y tomar decisiones basándose únicamente en su juicio clínico, deben apoyarse en datos objetivos, y una fuente clave para obtenerlos, no la única, son los test. Un proceso de evaluación psicológica es más amplio y complejo que el mero hecho de utilizar los test de forma correcta (Fernández-Ballesteros, 2014; Fernández-Ballesteros et al., 2001; Groth-Marnat y Wright, 2016; Wright, 2011), exige dar respuesta a otras muchas cuestiones capitales, entre las que cabe destacar las siguientes: qué se evalúa y para qué, quién evalúa, cuáles son las partes legítimamente implicadas en la evaluación, cómo se evalúa, cómo se va a proporcionar la información a las partes implicadas (feedback), qué planes de mejora e intervenciones se derivan de la evaluación, y finalmente, cómo perciben la evaluación los distintos agentes implicados. Lo primero es dejar muy claro qué y para qué se evalúa, pues ello

determinará el tipo de instrumentos y estrategias a seguir. Como dejó dicho nuestro Séneca, nunca soplan vientos favorables para quienes no saben a qué puerto se dirigen. Naturalmente, lo que se evalúa y para qué se hace variará en cada situación aplicada, no hay recetas válidas para todos los casos, serán los profesionales quienes tengan que decidirlo en cada caso. Quién es el responsable de la evaluación es la segunda cuestión que debe quedar meridianamente clara, habrá situaciones en las resulte obvio, pero en otras pueden surgir dudas. Siempre que se trate de evaluar variables psicológicas el responsable debe ser un profesional de la psicología, aunque algunas labores auxiliares puedan delegarse, por ejemplo, la aplicación de los test, pero nunca la interpretación de los resultados o la elaboración de un diagnóstico. En tercer lugar, en un proceso evaluador hay que identificar a todas las partes legítimamente implicadas en la evaluación, que variarán según cada situación. El fallo en esta identificación puede arruinar la evaluación, independientemente de su diseño técnico. Así, por ejemplo, en un contexto educativo habrá que tener en cuenta, por supuesto al alumnado evaluado, pero también a los padres, profesores, dirección del centro, administración educativa, representantes legales de los profesores, entre otros. Todos juegan su papel legítimo, han de ser tenidos en cuenta, pues cualquiera de ellos puede hacer fracasar una evaluación, ya se sabe que es más fácil destruir que construir. En cuarto lugar, está el meollo central desde el punto de vista técnico, es decir, cómo se lleva a cabo la evaluación, qué instrumentos de medida y qué recursos se van a utilizar. Ni que decir tiene que los profesionales deben utilizar test que reúnan unas propiedades psicométricas adecuadas y usarlos de forma correcta. En la actualidad contamos con numerosas fuentes sobre la construcción y análisis de la calidad de los test (American Educational Research Association, American Psychological Association y National Council on Measurement in Education, 2014; Downing y Haladyna, 2006; Haladyna y Rodríguez, 2013; Irwing, Booth y Hughes, 2018; Lane, Raymond y Haladyna, 2016; Muñiz, 2018; Muñiz y Fonseca-Pedrero, 2019; Schmeiser y Welch, 2006). Muchas evaluaciones que llevan a cabo con éxito las cuatro fases descritas fallan de forma lamentable en la quinta, relativa a proporcionar la información adecuada (feedback) a todas las partes legítimamente implicadas. Cuando se diseña una evaluación hay que tener muy claro desde el principio el tipo de informe que se va a realizar para cada parte interviniente, siendo un error muy común el pensar que un modelo de informe es suficiente para todos, la talla única tampoco funciona aquí. No es infrecuente que una evaluación bien diseñada y ejecutada se eche a perder a la hora de comunicar los resultados. Hay que ser muy cuidadosos con los informes, y asegurarse de que los destinatarios los entienden, pues una buena evaluación mal comunicada se convierte automáticamente en una evaluación deficiente. Sobre el problema de la elaboración de informes pueden verse, por ejemplo, los trabajos de Goodman y Hambleton (2004), Ryan (2006), o Zenisky y Hambleton (2016). En sexto lugar, los profesionales deben ser conscientes de que no se evalúa por evaluar, lo más lógico es que la evaluación desemboque en el desarrollo de algún plan de mejora o intervención. Cuando se lleven a cabo planes de mejora estos deben ser viables, basados en evidencias y evaluables, para así poder comprobar empíricamente su eficacia. Finalmente, algo que se olvida con demasiada frecuencia es la necesidad de recabar de forma objetiva y sistemática la opinión de las partes implicadas sobre la evaluación. La percepción de los participantes es clave, pues va a modular de forma importante su éxito. Los participantes no tienen por qué ser expertos en cuestiones técnicas, casi nunca lo serán, pero no por ello su opinión es menos importante, no hace falta ser zapatero para opinar sobre unos zapatos. En suma, lo que se pretende subrayar en las líneas precedentes es que el desarrollo de un proceso evaluativo es mucho más amplio y complejo que la mera aplicación de test, si bien estos son cruciales en el proceso. Dicho lo cual, vamos a centrarnos ahora precisamente en la problemática implicada en la utilización de los test como instrumentos de evaluación y diagnóstico.

UTILIZACIÓN DE LOS TEST

Un uso inadecuado de los test puede causar graves perjuicios a las personas evaluadas. Por ello distintas organizaciones nacionales e internacionales vienen desarrollando acciones y programas para tratar de mejorar

tres aspectos esenciales: la formación de los profesionales, la calidad de los test y la utilización que se hace de ellos. Las acciones llevadas a cabo por los Colegios Oficiales de Psicólogos (COP) y por organizaciones internacionales como la International Test Commission (ITC), o la Federación Europea de Asociaciones de Psicólogos (EFPA), pueden agruparse en dos grandes estrategias, denominadas restrictiva e informativa. Las acciones restrictivas se refieren a la limitación o restricción del uso de las pruebas a personas que tengan una formación adecuada. Esto varía mucho de unos países a otros (Bartram, 1996; Bartram y Coyne, 1998; Evers et al., 2012, 2017; Muñiz, Prieto, Almeida y Bartram, 1999), por ejemplo, los editores españoles clasifican los test en tres categorías (A, B, C) en función de su especialización, y solo venden los de las categorías B (test colectivos cognitivos y de personalidad) y C (escalas individuales y test proyectivos) a quienes acrediten el título de psicólogos. Esta y otras medidas, aunque adecuadas, no constituyen una garantía total, por lo que hay que complementarlas con otras de carácter informativo. La idea es que tanto los profesionales como los usuarios dispongan de toda la información necesaria sobre los test, para así minimizar su uso inadecuado. Entre estas acciones cabe destacar la publicación de códigos éticos, tales como el de la EFPA (2005), o las directrices de la asociación europea de evaluación psicológica (Fernández-Ballesteros et al., 2001). Un buen compendio en español puede verse en el número monográfico de la revista *Papeles del Psicólogo* (2009) dedicado al tema, o en autores como Knapp (2012), Koocher y Keith-Spiegel (2007), Lindsay, Koene, Ovreeide y Lang (2008), Leach y Oakland (2007), o Muñiz (1997). Aparte de estos códigos éticos hay todo un conjunto de directrices para orientar la construcción, análisis, uso y evaluación de la calidad de los test, entre las que cabe destacar los estándares técnicos desarrollados por la Asociación Americana de Psicología y otras dos organizaciones (AERA, APA y NCME, 2014), así como las directrices elaboradas por la ITC para la traducción y adaptación de los test de unas culturas a otras (Hambleton, Merenda y Spielberger, 2005; ITC, 2017; Muñiz, Elosua y Hambleton, 2013), o para su uso en investigación, control de calidad y seguridad (Muñiz, Hernández y Ponsoda, 2015). Para consultar otras directrices de interés, véase, por ejemplo, el trabajo de Muñiz y Bartram (2007), las páginas web de la ITC (www.intestcom.org), la EFPA (www.efpa.eu), o el COP (www.cop.es) en el apartado dedicado a la Comisión de Test.

Mención especial requiere la Norma ISO 10667 (AENOR, www.aenor.es), que regula todo lo relativo a la evaluación de personas en entornos laborales, y en cuya elaboración participó activamente el COP, formando parte de la comisión internacional que la desarrolló. En la norma se describen las competencias, obligaciones y responsabilidades de los clientes y de los proveedores del servicio de evaluación, antes, durante y después del proceso evaluativo. También proporciona directrices para todas las partes implicadas en el proceso evaluador, incluida la propia persona evaluada y quienes reciban los resultados de la evaluación (Muñiz y Fernández-Hermida, 2010). La norma está teniendo una gran influencia en la evaluación psicológica en contextos laborales, pues si bien no tiene rango de ley como tal, aporta un protocolo riguroso de actuación, prestigiando a las compañías que la adoptan. Las empresas e instituciones pueden certificarse en las agencias correspondientes si demuestran a los auditores que siguen los procesos de evaluación establecidos en la norma.

Otra actividad clave dentro de la estrategia de diseminar información rigurosa sobre los test es la publicación de los resultados de su evaluación por parte de expertos para ayudar a los profesionales a valorar los test más adecuados para sus fines. Hay dos grandes modelos de evaluación de la calidad de los test, uno de carácter más cualitativo, utilizado por el Buros Center for Testing en Estados Unidos, que viene a ser una revisión literaria en la que se comentan los puntos fuertes y débiles de la prueba evaluada (Carlson y Geisinger, 2012; Elosua y Geisinger, 2016; Geisinger, 2012, 2016), y otro el europeo, que tiene un carácter mixto cualitativo-cuantitativo. Se puntúan de 1 a 5 cada una de las características psicométricas del test y estas valoraciones numéricas se complementan con una evaluación cualitativa de los principales aspectos de la prueba. El modelo europeo puede consultarse en la página web de la EFPA (www.efpa.eu), y los detalles de su revisión pueden verse en Evers et al. (2013). La versión española del Cuestionario de Evaluación de Test Revisado (CET-R) puede consultarse en la página web del COP (www.cop.es) y el proceso de adaptación seguido en Hernández, Ponsoda, Muñiz, Prieto y Elosua (2016). Utilizando este modelo CET-R ya se han

llevado a cabo en España siete evaluaciones anuales de test, estando actualmente en marcha la octava (Elosua y Geisinger, 2016; Fonseca-Pedrero y Muñiz, 2017; Gómez-Sánchez, 2019; Hernández, Tomás, Ferreres y Lloret, 2015; Hidalgo y Hernández, 2019; Muñiz et al., 2011; Ponsoda y Hontagas, 2013). Hasta la fecha se han revisado un total de setenta y dos test, pudiendo consultarse libremente las revisiones en la página web del COP. Lo ideal sería que la mayoría de los test publicados en nuestro país estuviesen revisados, como ocurre en Holanda, no obstante, en el contexto europeo somos de los países que más avanzado lleva el proyecto de evaluación de test. Hay que destacar el papel de los editores españoles de test (CEPE, GiuntiEOS Psychometrics, Pearson y TEA), que en todo momento han colaborado en las evaluaciones, y ofrecido su ayuda y experiencia para ir mejorando la calidad de los test que editan, lo cual revertirá en beneficio de todos: profesionales, usuarios y las propias editoriales.

Esta evaluación ofrece dos grandes ventajas, por un lado, proporciona una información de primera mano a los profesionales sobre la calidad de los test, y, por otro, sirve de guía a constructores y editores, orientándoles sobre las características psicométricas exigibles a las pruebas. Un tercer aspecto nada desdeñable es su importancia en la docencia y formación de los futuros psicólogos (Hidalgo y Hernández, 2019).

OBJETIVO

Dentro del contexto de acciones encaminadas a la mejora de la práctica de los test, la EFPA ha impulsado la aplicación de una encuesta para conocer las opiniones de los psicólogos sobre el uso de los test en cada país. Conocer esta opinión es clave para promover acciones y proyectos sobre aquellos aspectos que sean percibidos como deficitarios por los profesionales. Los primeros datos en seis países europeos se obtuvieron en el año 2000 (Muñiz et al., 2001), y una buena panorámica sobre los resultados obtenidos a partir de entonces en Europa puede verse en Evers, Muñiz, Bartram et al. (2012). La misma encuesta fue utilizada también en otros países no europeos (Evers, McCormick, Hawley, Muñiz, et al. 2017). En España la primera encuesta para conocer las opiniones de los profesionales sobre los test se llevó a cabo en 1999 (Muñiz y Fernández-Hermida, 2000), y se volvió a repetir diez años después, actualizando el cuestionario utilizado inicialmente (Muñiz y Fernández-Hermida, 2010). Como se puede ver en los trabajos citados, en líneas generales las opiniones de los psicólogos españoles sobre el uso de los test en nuestro país son bastante positivas, si bien se identifican algunos aspectos que sería deseable mejorar.

Transcurrida una década desde que se aplicase la última encuesta, el objetivo central del presente trabajo es recoger las opiniones de los psicólogos españoles sobre el uso de los test en nuestro país y su evolución. Conocer estas opiniones es un primer paso fundamental para poder llevar a cabo acciones y proyectos encaminados a mejorar todo lo relativo a la práctica de los test. Se trata, en suma, de dar respuestas a preguntas tales como ¿Qué opinión tienen los psicólogos sobre los test? ¿Creen que se utilizan adecuadamente? ¿Les parecen caros? ¿Cuáles son los problemas más importantes? ¿Conocen la evaluación anual de test que lleva a cabo el COP? ¿Qué opinan sobre ella?

ENCUESTA

El cuestionario utilizado en 2019 es el mismo que el ya utilizado diez años atrás con algunas modificaciones, unas dirigidas a matizar algunos aspectos sobre los test que no se habían incluido entonces (ítems 25, 26, 27, 28-9, 28-10, 28-11), y otras a recabar la opinión de los psicólogos acerca del proceso de evaluación anual de test que lleva a cabo el COP desde 2011 (ítems 29-1, 29-2, 29-3, 29-4, 29-5). Todos los ítems de respuesta cerrada tienen un formato tipo Likert con cinco categorías, excepto el ítem 29-1 que es dicotómico. Precisamente en este ítem, si la persona responde que no conoce la evaluación anual de test que lleva a cabo el COP, la aplicación informática salta las cuatro preguntas siguientes (29.2, 29.3, 29.4, 29.5), relativas a

dicha evaluación, y pasa directamente al ítem 30, en el que se pregunta por los tres test más utilizados por los encuestados. La encuesta utilizada puede verse en el Anexo 1.

PARTICIPANTES

Respondieron a la encuesta 1.248 colegiados de un total de 75.135 que componen el COP. Las características principales de los participantes aparecen en las tablas 1 y 2. La muestra está formada por un 73,7% de mujeres y un 26,3% de hombres, y dado que el COP está compuesto por un 80,57% de mujeres y un 19,43% de hombres, se observa una mayor participación de los hombres en la encuesta. En cuanto a la edad, el 6,3% tienen entre 23 y 29 años, el 24,1% entre 30 y 39, el 27,4% entre 40 y 49, el 29,1% entre 50 y 59, el 12,4% entre 60 y 69, y el 0,7% tiene 70 años o más, siendo el promedio de edad de los participantes de 46,31 años, con una desviación típica de 11,08 años. Este promedio es similar a la media de las edades del COP que está en 44,19 años con desviación típica de 12,2. Estos datos parecen indicar que la muestra utilizada representa razonablemente bien a la población de colegiados del COP, aunque hubiese sido deseable un mayor número de participantes. En cuanto a las especialidades predomina Clínica-Salud (67,6%), seguida de Educativa (12,7%) y Trabajo (5,5%). Otras especialidades, como deporte, jurídica, tráfico, servicios sociales, etc., constituyen el 14,1% restante. Resulta interesante notar que en Clínica y Educativa el porcentaje de hombres es del 24,6% y el 25,8%, respectivamente, mientras que en el área de Trabajo dicho porcentaje aumenta hasta el 48%. La mayoría trabaja en el sector privado (69,9%), independientemente de la especialidad: Clínica-Salud (74,5%), Educativa (66,7%) y Trabajo (62,3%). Un 91,2% de los encuestados trabajan como psicólogos, y sólo un 8,8% ejercen otro tipo de trabajos.

TABLA 1 DESCRIPCIÓN DE LA MUESTRA Y DE LA POBLACIÓN ENCUESTADA		
	Muestra	Colegio Oficial de Psicólogos
Participantes	1.248	75.135
Mujeres	73,7%	80,57%
Hombres	26,3%	19,43%
Media de Edades (DT)	46,31 (11,08)	44,19 (12,20)
Años que lleva colegiado (DT)	17,45 (10,96)	12,80 (10,38)
DT: desviación típica		

TABLA 1
DESCRIPCIÓN DE LA MUESTRA Y DE LA POBLACIÓN ENCUESTADA

TABLA 2 DESCRIPCIÓN DE LA MUESTRA ATENDIENDO A SU DISTRIBUCIÓN POR CAMPO PROFESIONAL Y SECTOR		
		%
Campo profesional	Clínica y Salud	67,6
	Educativa	12,7
	Trabajo	5,5
	Otras	14,1
Sector	Público	26,6
	Privado	69,9
	Desempleado	3,5

TABLA 2
DESCRIPCIÓN DE LA MUESTRA ATENDIENDO A SU DISTRIBUCIÓN POR CAMPO PROFESIONAL Y SECTOR

PROCEDIMIENTO

La encuesta se aplicó online, a través de un enlace en la página web del COP. El sistema informático utilizado verificaba que la persona estaba colegiada, y no daba la posibilidad de responder la encuesta más de una vez. Para difundirla entre los colegiados se utilizaron distintos medios, tales como Infocop, la Agenda, o el Congreso Nacional de Psicología, explicando los objetivos de la misma y solicitando la participación. También se envió un correo electrónico a los colegiados, informándoles de la encuesta y animándoles a responderla. Al entrar a la encuesta, los participantes daban su consentimiento informado y se garantizó el tratamiento confidencial de los datos.

ANÁLISIS DE LOS DATOS

Se calcularon los estadísticos descriptivos de los ítems y de las dimensiones factoriales obtenidas. La fiabilidad se evaluó mediante el coeficiente Alfa (Cronbach, 1951). Para el estudio de las diferencias entre las medias de las especialidades se utilizó el análisis de varianza. La estructura dimensional se estimó mediante un análisis de Componentes Principales, utilizando el Análisis Paralelo sobre una matriz de correlaciones policóricas, con rotación oblicua (Calderón, Navarro, Lorenzo-Seva y Ferrando, 2019; Ferrando y Lorenzo-Seva, 2017; Ferrando y Anguiano, 2010; Izquierdo, Olea y Abad, 2014; Lloret-Segura, Ferreres-Traver, Hernández-Baeza y Tomás-Marco, 2014). Los análisis se llevaron a cabo con los programas SPSS (IBM, 2016) y FACTOR (Ferrando y Lorenzo-Seva, 2017; Lorenzo-Seva y Ferrando, 2013).

DIMENSIONES EVALUADAS POR EL CUESTIONARIO

Se obtuvieron cinco dimensiones principales (tabla 3) que explican el 51,63% de la varianza total.

La primera dimensión se refiere a las actitudes de los psicólogos hacia los test por considerarlos una herramienta útil (ítems 21-26), más el ítem 27 que presenta una saturación factorial débil. La segunda agrupa los ítems relativos a los problemas de uso de los test, recoge los 11 problemas presentados en el cuestionario (ítems 28-1 a 28-11). La tercera recoge los ítems relativos a la regulación de los test y su uso. Como cabría esperar los ítems referidos a la necesidad de legislación y regulación del uso de los test presentaron saturaciones positivas, mientras que los referidos a la permisividad en su uso presentaron saturaciones negativas. La cuarta dimensión se refiere a la formación y conocimientos sobre los test (ítems 1, 2, 4 y 6). Finalmente, la quinta recoge el uso de internet, la informatización y automatización (ítems 5, 7, 10, 13, 15, 17 y 20).

Los ítems que presentan saturaciones positivas son los que ponen en duda la validez y potencial de estos avances tecnológicos, mientras que los que presentan saturaciones negativas, plasman un reconocimiento a las ventajas de los test informatizados y el uso de internet. El cuarto componente, referido a formación y conocimientos sobre los test, es el que presenta una fiabilidad más baja, pudiendo atribuirse en gran medida al reducido número de ítems que conforman esta dimensión. En suma, teniendo en cuenta los resultados obtenidos parece razonable pensar que los ítems del cuestionario se organizan en torno a cinco grandes dimensiones, que podrían describirse como: actitudes hacia los test, problemas de uso, regulación del uso de los test, formación-conocimientos, y el uso de internet, la informatización y automatización. Cabe señalar que en las ediciones anteriores de la encuesta se obtuvieron ocho dimensiones en vez de cinco como ahora, lo cual es debido a que el método de extracción de dimensiones utilizado entonces (regla de Kaiser) tiende a generar estructuras factoriales más fragmentadas. Aunque no se incluye aquí por parsimonia, se han analizado también los datos utilizando dicho método y los resultados confirman en general los obtenidos en ediciones anteriores.

TABLA 3 ANÁLISIS FACTORIAL DE LOS ÍTEMS DEL CUESTIONARIO					
Dimensiones					
Ítems	I	II	III	IV	V
21	0,86				
22	0,85				
23	0,85				
24	0,61				
25	0,72				
26	0,81				
27	0,27				
28-1		0,65			
28-2		0,84			
28-3		0,84			
28-4		0,78			
28-5		0,82			
28-6		0,76			
28-7		0,78			
28-8		0,83			
28-9		0,84			
28-10		0,80			
28-11		0,35			
3			0,55		
8			0,61		
9			0,52		
11			0,66		
12			0,78		
14			-0,28		
16			-0,52		
18			-0,54		
19			0,72		
1				0,89	
2				0,73	
4				0,45	
6				0,74	
5					-0,43
7					-0,76
10					0,48
13					0,66
15					-0,72
17					0,64
20					0,69
% varianza	17,53	14,64	8,18	5,87	5,41
% acumulado	17,53	32,17	40,35	46,22	51,63
Coefficiente Alfa	0,84	0,91	0,69	0,62	0,70
Nota. Se eliminaron los pesos inferiores a 0,45, excepto cuando la variable no alcanzaba dicho peso.					

TABLA 3
ANÁLISIS FACTORIAL DE LOS ÍTEMS DEL CUESTIONARIO

OPINIONES SOBRE EL USO DE LOS TEST

En la Tabla 4 se muestran las medias y las desviaciones típicas de las respuestas de los participantes a los ítems del cuestionario. Se ofrecen los datos de la muestra total y dividida por especialidades profesionales de Clínica-Salud, Educativa y Trabajo. Las diferencias entre las medias de las especialidades se estudiaron mediante el análisis de varianza. Cuando no se cumplía el supuesto de homogeneidad de las varianzas (ítems 5, 11, 15, 17, 19, 21, 22, 23, 24, 26 y 28-2) se empleó la prueba robusta de Brown-Forsythe para las comparaciones. El asterisco detrás del texto del ítem indica que existieron diferencias significativas entre especialidades ($p < 0,05$), y entre paréntesis figuran los pares de comparaciones que resultaron estadísticamente significativas (La C se refiere a Clínica-Salud, la E a Educación y la T a Trabajo).

Los datos de la tabla hablan por sí mismos, pero cabe mencionar algunos aspectos reseñables. El primero (ítems 1 y 6) es que los colegiados consideran, con muy buen criterio, que los conocimientos obtenidos en la carrera de Psicología no son suficientes para el uso adecuado de la mayoría de los test, por tanto, es necesario seguir formándose. También alertan los profesionales de que los test informatizados (ítem 5) van reemplazando progresivamente a los clásicos de papel y lápiz, y ello con mayor intensidad en el ámbito educativo, y sobre todo en el de trabajo. Hay un consenso claro sobre la necesidad de que sean los psicólogos quienes utilicen e interpreten los resultados de los test (ítems 8 y 9). Los psicólogos son muy conscientes de las posibilidades de fraude que abre la aplicación de test por internet (ítems 17 y 20), y reclaman que el COP ejerza un papel más activo para mejorar el uso de los test (ítem 19). Los psicólogos utilizan habitualmente los test en el ejercicio de su profesión (ítem 21), especialmente en el ámbito educativo, seguido de clínica y trabajo, y no tienen dudas sobre su utilidad si se usan adecuadamente (ítems 22 y 23). Tienen además la percepción de que en la última década el uso de los test ha mejorado en España (ítem 24). En cuanto a los problemas en el uso de los test en nuestro país (ítems 28-1 al 28-11) su percepción es bastante positiva, quejándose, eso sí, sobre el elevado precio de las pruebas, especialmente en clínica y educativa.

En la tabla 5 se muestran las medias y las desviaciones típicas de las respuestas de los participantes a las cinco dimensiones del cuestionario. Los ítems que presentaban saturaciones negativas en las dimensiones tres y cinco fueron invertidos antes de calcular la puntuación total. De esta manera una mayor puntuación indica una opinión más positiva hacia el control y legislación sobre los test y su uso, y una opinión más negativa hacia el uso de la tecnología (uso de internet, test informatizados, informes automatizados). Se ofrecen los resultados globales y para las tres especialidades, cuyas medias son comparadas mediante un ANOVA, con excepción de la primera dimensión, para la que se utilizó la prueba de Brown-Forsythe por no cumplirse el supuesto de homogeneidad de varianzas.

Como se puede ver en la tabla 5, las actitudes de los psicólogos españoles hacia los test son bastante favorables, con una media global de 4 puntos en una escala que va de 1 a 5. Los especialistas en educativa y trabajo son algo más favorables que los clínicos, lo cual no deja de tener su lógica, dada la naturaleza del ejercicio profesional de las especialidades. Como ya se señaló al comentar los resultados desglosados por ítems, los psicólogos españoles no encuentran problemas graves de uso de los test, con una media de 2,93. Manifiestan una clara posición a favor de un mayor control y regulación del uso de los test, con una media de 4,07 en esa dimensión. Claramente demandan una necesidad de formación continua, como muestra una media de 2,72 en la dimensión 4. Y no se muestran muy entusiastas con el uso de las nuevas tecnologías (internet, informatización, automatización); en concreto los especialistas en clínica son algo más escépticos respecto a la utilidad de los avances tecnológicos y perciben más problemas en su uso (2,93), que los especialistas en Educativa (2,79) y Trabajo (2,63).

TABLA 4 MEDIA Y DESVIACIÓN TÍPICA DE CADA UNO DE LOS ÍTEMS DE LA ENCUESTA POR ESPECIALIDADES (CLÍNICA-SALUD, EDUCATIVA, TRABAJO) Y GLOBAL								
Ítems	Clínica-Salud N=844		Educativa N=159		Trabajo N=69		Global N=1.248	
	Media	DT	Media	DT	Media	DT	Media	DT
1. La formación recibida en la carrera de Psicología es suficiente para la utilización correcta de la mayoría de los test	2,46	1,16	2,62	1,14	2,64	1,19	2,47	1,16
2. La formación recibida en cursos y Máster es suficiente para el uso correcto de la mayoría de los test	3,05	1,07	3,13	0,96	2,94	1,11	3,05	1,05
3. La Federación Europea de Asociaciones de Psicólogos (EFPA) debería establecer un sistema para acreditar la competencia de los usuarios de test* (C-E)	3,34	1,29	3,60	1,24	3,58	1,36	3,39	1,30
4. Los profesionales disponen de suficiente información (revisiones independientes, investigaciones, documentación...) sobre la calidad de los test editados en nuestro país	3,01	1,10	3,21	,96	2,91	1,10	3,04	1,09
5. En mi campo profesional los test informatizados están reemplazando progresivamente a los test de papel y lápiz* (C-E, C-T)	3,27	1,30	3,62	1,10	3,90	1,07	3,36	1,29
6. Mis conocimientos actuales en relación con los test los recibí fundamentalmente durante la carrera de Psicología	2,30	1,20	2,42	1,21	2,57	1,25	2,33	1,21
7. La aplicación de los test por Internet tiene muchas ventajas en comparación con la aplicación clásica de papel y lápiz* (C-T, E-T)	3,14	1,18	3,20	1,08	3,59	1,10	3,16	1,18
8. El uso de los test psicológicos debería restringirse a psicólogos cualificados	4,23	1,02	4,28	0,98	4,41	1,12	4,24	1,03
9. Aunque los no psicólogos podrían aplicar y puntuar los test, la interpretación e información sobre los resultados deberían estar restringidos a los psicólogos	4,39	1,13	4,36	1,13	4,43	1,12	4,38	1,14
10. Los informes generados automáticamente por ordenador no tienen ninguna validez	2,81	1,15	2,68	0,97	2,81	1,14	2,78	1,12
11. Los estándares y directrices que definen las cualidades técnicas mínimas de un test deberían ser obligatorios (por ejemplo los estándares de la Federación Europea de Asociaciones de Psicólogos [EFPA], o los de la Asociación de Psicología)	4,15	0,89	4,18	0,77	4,32	0,81	4,18	0,87
12. Se necesita legislación para controlar los abusos más serios con los test* (C-T, E-T)	3,99	1,01	3,91	1,04	4,36	0,79	4,02	1,01
13. La aplicación de los test por Internet pone en desventaja a algunas personas evaluadas	3,45	1,09	3,28	1,04	3,23	1,02	3,41	1,10
14. Todo aquel que sea capaz de demostrar su competencia en el uso de los test (sea psicólogo o no) debería ser autorizado para utilizarlos	2,09	1,30	2,05	1,18	2,46	1,40	2,08	1,28
15. Si se utiliza adecuadamente, Internet puede mejorar mucho la calidad de la aplicación de los test* (C-T, E-T)	3,33	1,09	3,30	0,95	3,83	0,82	3,36	1,08
16. Los controles sobre los test deberían ser mínimos, pues inhiben el desarrollo de nuevas ideas y nuevos procedimientos de evaluación	1,84	0,97	1,87	0,93	1,74	0,97	1,83	0,96
17. La aplicación de test por Internet no permite proteger la privacidad de los usuarios	2,45	1,19	2,33	1,09	2,16	1,12	2,41	1,18
18. Habría que permitir a los editores que vendiesen cualquier test que consideren adecuado	1,75	1,01	1,80	1,03	1,74	1,08	1,76	1,02
19. El Colegio Oficial de Psicólogos debería ejercer un papel más activo para regular y mejorar el uso que se hace de los test	4,09	0,99	3,99	0,91	4,25	0,85	4,09	0,99
20. La aplicación de test por Internet abre posibilidades de fraude	3,57	1,14	3,36	1,18	3,52	0,99	3,52	1,14
21. En el desempeño de mi profesión utilizo test habitualmente* (C-E, E-T)	3,97	1,19	4,32	0,96	3,72	1,35	4,02	1,18
22. Los test constituyen una excelente fuente de información si se combinan con otros datos psicológicos* (C-E, C-T)	4,41	0,90	4,61	0,72	4,68	0,58	4,47	0,85
23. Utilizados correctamente, los test son de gran ayuda para el psicólogo* (C-E, C-T)	4,40	0,90	4,63	0,71	4,67	0,56	4,46	0,85
24. Teniendo en cuenta todos los aspectos, creo que en la última década el uso de los test ha mejorado en mi país* (C-E)	3,67	0,95	3,86	0,87	3,75	0,88	3,73	0,93
25. La práctica profesional basada en evidencias potencia el uso de los test como fuentes de información*	3,89	0,96	4,07	0,82	4,13	0,78	3,96	0,92

TABLA 4
MEDIA Y DESVIACIÓN TÍPICA DE CADA UNO DE LOS ÍTEMS DE LA ENCUESTA
POR ESPECIALIDADES (CLÍNICA-SALUD, EDUCATIVA, TRABAJO) Y GLOBAL

26. En la toma de decisiones profesionales doy un peso importante a los datos obtenidos mediante test* (C-E)	3,36	1,04	3,86	0,78	3,57	0,81	3,46	1,00
27. Antes de utilizar un test compruebo sus propiedades psicométricas	3,83	1,16	3,87	1,09	4,12	1,04	3,89	1,15
28-1. Hacer fotocopias de materiales sujetos a copyright	3,11	1,43	3,03	1,33	2,78	1,49	3,05	1,42
28-2 Hacer evaluaciones utilizando test inadecuados	2,39	1,29	2,48	1,29	2,68	1,50	2,44	1,31
28-3. No estar al día	3,06	1,26	3,08	1,20	3,16	1,36	3,08	1,26
28-4. No contrastar las interpretaciones con otros	3,06	1,27	2,92	1,27	3,20	1,28	3,04	1,28
28-5. No tener en cuenta los errores de medida de las puntuaciones	2,98	1,26	3,00	1,16	3,04	1,27	3,01	1,25
28-6. No restringir la aplicación de los test a personal cualificado	2,88	1,49	2,95	1,40	3,12	1,59	2,88	1,48
28-7. No tener en cuenta las condiciones locales (país, región) que pueden afectar a la validez	2,86	1,36	2,96	1,23	3,00	1,44	2,89	1,33
28-8. Hacer interpretaciones que van más allá de los límites del test	2,88	1,39	2,82	1,27	2,96	1,38	2,87	1,36
28-9. Utilizar test con baremos inadecuados	2,51	1,30	2,43	1,27	2,59	1,33	2,50	1,30
28-10. Utilizar algunos ítems sin los permisos correspondientes	2,26	1,30	2,14	1,17	2,30	1,22	2,25	1,28
28-11. Tener un precio muy elevado* (C-T, E-T)	4,25	1,06	4,25	1,04	3,65	1,21	4,17	1,11
DT: desviación típica Nota. El asterisco indica que las diferencias de medias entre las especialidades son estadísticamente significativas ($p<0,05$). Los pares de letras entre paréntesis indican entre qué especialidades concretas se dan las diferencias significativas: Clínica-Salud (C), Educativa (E) y Trabajo (T).								

TABLA 4
MEDIA Y DESVIACIÓN TÍPICA DE CADA UNO DE LOS ÍTEMS DE LA ENCUESTA POR ESPECIALIDADES (CLÍNICA-SALUD, EDUCATIVA, TRABAJO) Y GLOBAL (Continuación)

TABLA 5 MEDIA Y DESVIACIÓN TÍPICA DE LAS DIMENSIONES POR ESPECIALIDADES (CLÍNICA-SALUD, EDUCATIVA, TRABAJO)								
Dimensiones	Clínica-Salud N=844		Educativa N=159		Trabajo N=69		Global N=1.248	
	Media	DT	Media	DT	Media	DT	Media	DT
1. Actitudes hacia los test* (C-E)	3,94	0,74	4,17	0,57	4,09	0,56	4,00	0,70
2. Problemas en el uso de los test	2,93	0,94	2,91	0,89	2,95	1,07	2,93	0,94
3. Control y regulación de test y su uso	4,05	0,57	4,08	0,61	4,14	0,56	4,07	0,58
4. Formación y conocimientos sobre los test	2,70	0,78	2,85	0,72	2,76	0,79	2,72	0,77
5. Internet, informatización y automatización* (C-E, C-T)	2,93	0,70	2,79	0,64	2,63	0,59	2,89	0,69
Nota. El asterisco indica que las diferencias de medias entre las especialidades son estadísticamente significativas ($p<0,05$). Los pares de letras entre paréntesis indican entre qué especialidades concretas se dan las diferencias significativas: Clínica-Salud (C), Educativa (E) y Trabajo (T).								

TABLA 5
MEDIA Y DESVIACIÓN TÍPICA DE LAS DIMENSIONES POR ESPECIALIDADES (CLÍNICA-SALUD, EDUCATIVA, TRABAJO)

CONOCIMIENTO DE LA EVALUACIÓN DE TEST

En cuanto a la pregunta de si conocen la evaluación anual de test que lleva a cabo el COP, sorprende comprobar que sólo dicen conocerla el 22,5% de los colegiados, hay, por tanto, un 77,5% que desconocen esta importante evaluación, por lo que está clara la necesidad insistir en su difusión. Entre quienes sí la conocen, les parece importante y necesaria, como se observa en los datos de la tabla 6.

EVOLUCIÓN DE LAS OPINIONES

En la tabla 7 aparecen las puntuaciones medias para cada ítem en las tres evaluaciones llevadas a cabo en los años 2000, 2010 y 2020. Dado que las encuestas utilizadas en las tres evaluaciones no tienen exactamente los mismos ítems, sólo se ofrecen los resultados para los ítems comunes a las tres. Como se puede observar, la característica general más sobresaliente es la estabilidad, pero con algunos matices dignos de mención. Las correlaciones entre los resultados de las tres evaluaciones son muy elevadas: 0,986 entre la 1ª y 2ª, 0,955 entre la 1ª y la 3ª, y 0,982 entre la 2ª y la 3ª.

El primer aspecto a destacar es que los psicólogos españoles aprecian que mejora la información de la que disponen sobre la calidad de los test (ítem 4), pasando de un valor medio en 2010 de 2,73 a una media de

3,04 en 2020. Seguramente no es ajeno a esa percepción la labor de evaluación de los test llevada a cabo por el COP a lo largo de la última década. Otro aspecto a destacar es el aumento del uso de los test en el ejercicio de la profesión (ítem 21), pasando de un valor de 3,76 en 2010 a 4,02 en 2020. Esto constituye un indicador indirecto de la actitud positiva hacia los test, así como una apuesta clara sobre un ejercicio de la profesión basado en evidencias empíricas obtenidas a partir de los test. Cabe reseñar, en tercer lugar, que todos los ítems relativos a los problemas en el uso de los test (ítem 28) bajan sus valores medios, lo que significa que los profesionales perciben una mejora general en el uso de los test. El cambio más significativo se da en el ítem 28-1, relativo a las fotocopias de los test, que pasa de 3,51 en 2010 a 3,05 en 2020. Seguramente el aumento de los test online y las correcciones informatizadas haya favorecido esta percepción. También perciben que los test se ajustan mejor a las condiciones locales (país, región) que hace una década (ítem 28-7), pasando la media de 3,21 a 2,89 al evaluar el problema de la falta de ajuste. Parece surtir efecto el esfuerzo de las organizaciones nacionales como el COP, internacionales, como la ITC, y los propios editores, en traducir y adaptar los test a las condiciones de las poblaciones evaluadas.

TABLA 6 OPINIONES DE LOS PSICÓLOGOS SOBRE LA EVALUACIÓN DE LA CALIDAD DE LOS TEST LLEVADA A CABO ANUALMENTE POR EL COP		
Ítems	Media	DT
Me parece importante para mejorar la calidad de los test	4,46	0,72
Consulta las evaluaciones de los test en la página web del COP	3,81	1,15
Me ayuda en la elección de los test que utilizo	3,65	1,10
Es innecesaria, los psicólogos ya tienen suficiente información para saber qué test utilizar	1,62	1,01
DT: desviación típica		

TABLA 6
OPINIONES DE LOS PSICÓLOGOS SOBRE LA EVALUACIÓN DE LA
CALIDAD DE LOS TEST LLEVADA A CABO ANUALMENTE POR EL COP

TABLA 7 MEDIAS DE LOS ÍTEMS OBTENIDAS EN LAS ENCUESTAS DE LOS AÑOS 2000, 2010 Y 2020			
Ítems	Evaluación-1 2000	Evaluación-2 2010	Evaluación-3 2020
1	2,41	2,43	2,47
4	2,38	2,73	3,04
6	2,57	2,59	2,33
8	4,23	4,12	4,24
9	4,34	4,39	4,38
11	4,33	4,10	4,18
12	4,29	4,01	4,02
14	2,42	2,10	2,08
16	1,85	1,93	1,83
18	1,57	1,77	1,76
19	4,15	4,13	4,09
21	3,56	3,76	4,02
22	4,41	4,46	4,47
23	4,37	4,41	4,46
28-1	3,60	3,51	3,05
28-2	2,63	2,64	2,44
28-5	3,07	3,10	3,01
28-6	2,91	2,92	2,88
28-7	3,28	3,21	2,89
28-8	2,99	2,97	2,87

TABLA 7
MEDIAS DE LOS ÍTEMS OBTENIDAS EN LAS ENCUESTAS DE LOS AÑOS 2000, 2010 Y 2020

TEST MÁS UTILIZADOS EN ESPAÑA

En la tabla 8 se muestran los resultados referidos a los test más utilizados en la práctica diaria profesional de los psicólogos. Cabe destacar, en primer lugar, que los 25 test más utilizados son todos psicométricos, quedando los proyectivos relegados a posiciones muy inferiores. Como se puede observar, en primer lugar se encuentra el test de inteligencia para niños WISC, seguido del Inventario Clínico Multiaxial de Millon (MCMI) y de la adaptación española del Personality Assessment Inventory (PAI). También es destacable que todas las pruebas están adaptadas al contexto español, y 6 de las 25 han sido desarrolladas por autores españoles, lo que representa un 24%, lo cual no está nada mal. Comparada esta evaluación con la de hace diez años, se incorporan a la lista diez nuevas pruebas, señaladas en la tabla 8 con un asterisco (PAI, SENA, D2, ENFEN, LSB-50, CARAS, CompeTEA, RIAS, BAI, GDS), algunas de ellas publicadas con posterioridad a la encuesta de 2010 (PAI, SENA, LSB-50). Estos cambios en la lista de los test más utilizados constituyen un buen indicador de la gran actividad que se desarrolla en el campo de la construcción y publicación de test en España.

Cuando se analiza el uso por especialidades (tabla 9), se observa, como no puede ser de otro modo, que los test varían notablemente, así, por ejemplo, en Clínica y Educativa aparece en primer lugar la evaluación de aspectos cognitivos (WISC), mientras que en Trabajo lo hacen los de Personalidad (16PF).

TABLA 8 LOS 25 TEST MÁS UTILIZADOS POR LOS PSICÓLOGOS ESPAÑOLES			
Nombre de la prueba		N	%
WISC	Weschler Intelligence Scale for Children	289	23,2%
MCMI	Millon Clinical Multiaxial Inventory	157	12,6%
PAI*	Inventario de Evaluación de la Personalidad	156	12,5%
16PF	16 Personality Factors	142	11,4%
WAIS	Weschler Adult Intelligence Scale	138	11,1%
BDI	Beck Depression Inventory	122	9,8%
MMPI	Minnesota Multiphasic Personality Inventory	120	9,6%
CUIDA	Evaluación de Adoptantes, Cuidadores, Tutores y Mediadores	110	8,8%
SENA*	Sistema de Evaluación de Niños y Adolescentes	107	8,6%
STAI	State Trait Anxiety Inventory	99	7,9%
PROLEC	Batería de Evaluación de los Procesos Lectores	71	5,7%
MMSE	Mini Mental State Examination	54	4,3%
SCL-90	Symptom Checklist 90	51	4,1%
D2*	Test de Atención	44	3,5%
BASC	Behavior Assessment System for Children	42	3,4%
ENFEN*	Evaluación Neuropsicológica de las Funciones Ejecutivas en Niños	35	2,8%
LSB-50*	Listado de Síntomas Breve	30	2,4%
RAVEN	Matrices Progresivas de Raven	28	2,2%
CARAS*	Test de Percepción de Diferencias	27	2,2%
CompeTEA*	Evaluación de Competencias	26	2,1%
NEO-PI	NEO Personality Inventory	24	1,9%
RIAS*	Reynolds Intellectual Assessment Scales	23	1,8%
BAI*	Beck Anxiety Inventory	20	1,6%
GDS*	Geriatric Depression Scale	20	1,6%
ISRA	Inventario de Situaciones y Respuesta de Ansiedad	20	1,6%

Notas: Se marcan con asteriscos los test que aparecen nuevos respecto a la encuesta de 2010. Se agrupan las distintas versiones de un mismo test bajo el nombre original. Por ejemplo, aparecen agrupadas las formas abreviadas de un test, como es el caso de RIAS y su versión abreviada RIST. También se agrupan las respuestas sobre el MEC, que es la versión adaptada y validada en España del MMSE, o las distintas versiones del WISC y el WAIS, como WISC-R y WISC-V.

TABLA 8
LOS 25 TEST MÁS UTILIZADOS POR LOS PSICÓLOGOS ESPAÑOLES

TABLA 9 LOS DIEZ TEST MÁS UTILIZADOS POR LOS PSICÓLOGOS ESPAÑOLES POR ESPECIALIDADES						
	Clínica-Salud N=844		Educativa N=159		Trabajo N=69	
1	WISC	19,9%	WISC	64,8%	16PF	27,5%
2	MCMI	14,1%	PROLEC	26,4%	CompeTEA	21,7%
3	BDI	14,0%	SENA	12,6%	CTC	8,7%
4	PAI	12,4%	D2	11,3%	PAI	7,2%
5	WAIS	12,4%	WAIS	10,1%	BAT-7	5,8%
6	16PF	10,9%	BADYG	9,4%	BIP	5,8%
7	MMPI	10,6%	ENFEN	7,5%	MMPI	5,8%
8	STAI	10,1%	TAMAI	6,2%	TPT	5,8%
9	SENA	8,8%	CARAS	5,7%	MCMI	4,3%
10	CUIDA	7,2%	16PF	5,0%	PAPI	4,3%

Notas: Se identifican a continuación las siglas de los test de esta tabla que no aparecen descritos en la tabla 8: BADYG (Batería de evaluación de las aptitudes diferenciales y generales); BAT-7 (Batería de Aptitudes de TEA). BIP (Inventario Bochum de Personalidad y Competencias); CTC: Cuestionario TEA Clínico; PAPI (The Personality and Preference Inventory); TAMAI (Test Autoevaluativo Multifactorial de Adaptación Infantil); TPT (Test de Personalidad de TEA)

TABLA 9
LOS DIEZ TEST MÁS UTILIZADOS POR LOS PSICÓLOGOS ESPAÑOLES POR ESPECIALIDADES

ALGUNAS CONCLUSIONES GENERALES

Esta es la tercera encuesta que se hace a los psicólogos españoles acerca de sus opiniones sobre distintas cuestiones relacionadas con la práctica de los test en nuestro país, las otras dos datan de 2000 y 2010.

Respondieron 1.248 colegiados a las treinta preguntas de la encuesta. A partir de los resultados, ya comentados con detalle en las líneas precedentes, pueden extraerse algunas conclusiones generales.

El análisis factorial de las preguntas de la encuesta indica que estas se articulan en torno a cinco dimensiones principales: Actitudes hacia los test, Problemas en el uso de los test, Control y regulación de los test y su utilización, Formación y conocimientos sobre los test, y finalmente, una quinta dimensión relativa a la influencia de las nuevas tecnologías en la práctica de los test, como informatización, internet y automatización de las evaluaciones.

Se constata una actitud positiva de los psicólogos españoles hacia los test utilizados adecuadamente, con una media de 4 puntos en una escala de 1 a 5 en esta dimensión. Los profesionales más favorables son los del área Educativa (4,17), seguidos de Trabajo (4,09) y Clínica-Salud (3,94).

En lo relativo a los problemas de uso de los test, los psicólogos españoles reconocen que hay aspectos a mejorar, pero consideran los problemas existentes de una gravedad moderada, con una media de 2,93, ligeramente por debajo de la media teórica de la escala que se sitúa en 3.

Los encuestados son claramente partidarios de incrementar el control y la regulación del uso de los test por parte de las instituciones implicadas (media de 4,07), tales como la Federación Europea de Asociaciones de Psicología (EFPA), International Test Commission (ITC), y el COP a nivel nacional.

Se pone de relieve la necesidad de una formación continuada en todo lo relacionado con los test (media de 2,72), pues se reconoce que lo aprendido en los estudios previos, tales como grado, licenciatura o Máster, no es suficiente para utilizar los test actuales con solvencia.

Se comprueba una cierta contención, tal vez prudencia, de los profesionales a la incorporación de las nuevas tecnologías en la práctica de los test, tales como la evaluación online a través de internet o la automatización de los informes (media de 2,89). Este escepticismo es mayor en Clínica, seguido de Educativa y Trabajo.

Los psicólogos consideran que los test tienen un precio demasiado elevado (media de 4,17), especialmente los clínicos y educativos, y en menor grado en el ámbito de trabajo. Los editores deberían de tomar en consideración este aspecto.

La evaluación de la calidad de los test que lleva a cabo la Comisión de Test del COP es poco conocida todavía entre los profesionales, sólo la conocen un 22,5%. Es evidente que hay que seguir haciendo esfuerzos en su difusión. Ahora bien, la buena noticia es que quienes la conocen la consideran muy importante y necesaria.

En cuanto a la evolución de las opiniones de los psicólogos españoles sobre la práctica de los test en las últimas dos décadas, cabe destacar que los profesionales aprecian una mejora de la información de la que disponen sobre la calidad de los test, un incremento del uso de los test en el ejercicio de la profesión, y una mejora en la percepción del uso correcto de los test. Son indicadores en la buena dirección, si bien queda un amplio margen de mejora sobre el que habrá que seguir insistiendo desde todas las instituciones, tanto nacionales como internacionales.

En cuanto a los test más utilizados por los profesionales, destaca que los 25 más citados son todos ellos psicométricos, quedando muy relegados los proyectivos. Los profesionales de las especialidades de Clínica-Salud, Educativa y Trabajo utilizan test muy distintos, como era de esperar, destacando los de carácter cognitivo en Clínica y Educativa (WISC), y los de Personalidad en Trabajo (16PF).

MIRANDO HACIA EL FUTURO

Las opiniones de los profesionales sobre la práctica de los test son una foto fija de sus percepciones actuales, y vistas en perspectiva reflejan una tendencia positiva en relación con las mismas de los años 2000 y 2010, ya veremos qué ocurre en 2030 y siguientes. Todo cambia muy rápidamente, y la evaluación psicológica no es una excepción. Ya Heráclito nos lo advirtió hace veinticinco siglos, todo fluye, nada permanece. Haremos unas breves reflexiones sobre el futuro de la evaluación y de los test siguiendo los trabajos recientes sobre

el tema de Muñiz (2018) y Muñiz y Fonseca-Pedrero (2019). La fuerza motriz más relevante en el cambio de los procesos de evaluación son las nuevas tecnologías de la información y la comunicación, y en especial los avances informáticos, multimedia e Internet. Expertos como Bennet (1999, 2006), Breithaupt, Mills y Melican (2006), Drasgow (2016), Drasgow, Luecht y Bennet (2006), Sireci y Zenisky (2016), o Sireci y Faulkner-Bond (2014), entre otros muchos, consideran que dichas tecnologías están influyendo sobre todos los aspectos de la evaluación psicológica, tales como el diseño de los test, la construcción y presentación de los ítems, la puntuación de los test y la evaluación a distancia. Si bien emergen nuevas formas de evaluación, los test psicométricos seguirán siendo herramientas fundamentales, dada su objetividad y economía de medios y tiempo (Phelps, 2005, 2008). En este contexto de cambio tecnológico aparece la llamada Psicología 2.0 (Armayones et al., 2015), que pretende extender la psicología a través de las facilidades que ofrece Internet y las redes sociales. La evaluación no puede estar ajena a estas nuevas tendencias, apareciendo nuevos enfoques psicométricos conectados con el análisis de las grandes bases de datos (big data) de las que se dispone actualmente (Markovetz, Blaszkiewicz, Montag, Switala y Schlaepfer, 2014). Por ejemplo, las ventajas potenciales de usar los teléfonos móviles y otros portables como terminales para la evaluación abren nuevas posibilidades para la evaluación del futuro (Armayones et al., 2015; Chernyshenko y Stark, 2016; Insel, 2017; Miller, 2012). Además, trabajos como el pionero de Kosinski, Stillwell y Graepel (2013) analizan con éxito la posibilidad de utilizar los “me gusta” de facebook como predictores de distintas características humanas, entre ellas los rasgos de la personalidad, lo que hace preguntarse si nuestros rastros en las redes sociales sustituirán algún día no muy lejano a los cuestionarios y test tal como los conocemos ahora (Muñiz y Fonseca-Pedrero, 2019), nos parece que va a ser que no, pero nunca se sabe. Otro tema que cobra fuerza es el de la evaluación ambulatoria, que si bien no es nuevo en psicología, resurge con fuerza en la actualidad impulsada por las tecnologías de la información y comunicación (Chernyshenko y Stark, 2016; Myin-Germeys et al., 2018; Trull y Ebner-Priemer, 2013; van Os, Delespaul, Wigman, Myin-Germeys y Wichers, 2013). La evaluación ambulatoria abarca diferentes métodos y procedimientos que tratan de estudiar mediante dispositivos móviles las experiencias de las personas (emociones, sentimientos, pensamientos, síntomas psicológicos) en su entorno natural y en la vida diaria. Esta metodología permite evaluar los constructos psicológicos desde una perspectiva más dinámica, personalizada, contextual, micro-longitudinal y ecológica. Ello se consigue realizando evaluaciones varias veces al día para captar suficientemente la variabilidad temporal de los fenómenos. Las preguntas se activan mediante un beep y pueden presentarse de forma aleatoria o en intervalos de tiempo predeterminados. Se trata de una aproximación complementaria a los procedimientos tradicionales de evaluación psicométrica de papel y lápiz (Fonseca-Pedrero y Muñiz, 2017; Muñiz y Fonseca-Pedrero, 2019). Para analizar este tipo de datos se necesitan nuevos modelos psicométricos, entre los que destacan los modelos de redes (Borsboom y Cramer, 2013; Fonseca-Pedrero, 2017, 2018), así como los modelos procedentes de la teoría de los sistemas dinámicos o la teoría del caos (Nelson, McGorry, Wichers, Wigman, y Hartmann, 2017). Estos nuevos enfoques y tecnologías seguramente vienen para quedarse en los reales de la evaluación, y son bienvenidos, pero tendrán que superar la prueba del algodón y convencer a la implacable jueza: la validez. Todas las fantasías sobre los avances tecnológicos pasan por demostrar que aportan mejoras en la medida del constructo evaluado, de lo contrario no dejarán de ser meros fuegos de artificio, pues lo esencial no muta: siempre habrá que aportar evidencias empíricas de la fiabilidad y validez, y así garantizar que los instrumentos de medida evalúan de forma objetiva, equitativa y rigurosa, ese es el norte que no hay que perder de vista (Muñiz y Fonseca-Pedrero, 2019). De hecho, ya asistimos con demasiada frecuencia a propuestas sofisticadas informáticamente para evaluar constructos psicológicos y educativos online que no tienen detrás ninguna evidencia empírica que los avale. Osan algunos de ellos evaluar constructos tan complejos como la Dislexia o el TDAH, y generan gráficos muy aparentes con colores y semáforos varios, se lanzan incluso a diagnosticar, pero detrás no hay nada, no se aporta ninguna evidencia de validez, son mera apariencia, eso sí, seguramente muy lucrativa. Por el bien de los usuarios y de la profesión debemos contribuir entre todos a desenmascarar estos timos, envueltos de marketing y tecnología, pero

carentes de toda validez. Todo instrumento de medida, use la tecnología que use, debe aportar un manual detallado donde consten las evidencias de validez que lo avalan, y si no lo hace no vale nada, así de simple.

Anexo 1

Encuesta utilizada para recoger las opiniones de los psicólogos sobre la utilización de los test

Datos generales

Edad: _____ Sexo: Hombre _____ Mujer _____

Año en el que obtuvo la Licenciatura/Grado en Psicología: _____

Años que lleva colegiado: _____

Especialidad Profesional: Clínica-Salud _____ Educativa _____ Trabajo _____ Otras (diga cuál) _____

Trabaja actualmente como Psicólogo/a: sí _____ no _____

Trabaja en el sector: Público _____ Privado _____ En paro _____

Número de años que lleva en el trabajo actual _____

Instrucciones

Las cuestiones que aparecen a continuación están formuladas para responder en una escala de 1 a 5. Si está *en desacuerdo total* con la frase señale el 1, si está *totalmente de acuerdo* señale el 5, utilice los números 2, 3 y 4 para las situaciones intermedias. La encuesta es totalmente anónima.

Cuestionario

1. La formación recibida en la carrera de Psicología es suficiente para la utilización correcta de la mayoría de los test
2. La formación recibida en cursos y Másteres es suficiente para el uso correcto de la mayoría de los test
3. La *Federación Europea de Asociaciones de Psicólogos* (FEPA) debería establecer un sistema para acreditar la competencia de los usuarios de test
4. Los profesionales disponen de suficiente información (revisiones independientes, investigaciones, documentación...) sobre la calidad de los test editados en nuestro país
5. En mi campo profesional los test informatizados están reemplazando progresivamente a los test de papel y lápiz
6. Mis conocimientos actuales en relación con los test los recibí fundamentalmente durante la carrera de Psicología
7. La aplicación de los test por Internet tiene muchas ventajas en comparación con la aplicación clásica de papel y lápiz
8. El uso de los test psicológicos debería restringirse a psicólogos cualificados
9. Aunque los **no** psicólogos podrían aplicar y puntuar los test, la interpretación e información sobre los resultados deberían estar restringidos a los psicólogos

Anexo 1

Encuesta utilizada para recoger las opiniones de los psicólogos sobre la utilización de los test

10. Los informes generados automáticamente por ordenador **no** tienen ninguna validez

11. Los estándares y directrices que definen las cualidades técnicas mínimas de un test deberían ser obligatorios [por ejemplo los estándares de la Federación Europea de Asociaciones de Psicólogos (EFPA), o los de la Asociación de Psicología Americana (APA)].

12. Se necesita legislación para controlar los abusos más serios con los test

13. La aplicación de los test por Internet pone en desventaja a algunas personas evaluadas

14. Todo aquel que sea capaz de demostrar su competencia en el uso de los test (sea psicólogo o no) debería ser autorizado para utilizarlos

15. Si se utiliza adecuadamente, Internet puede mejorar mucho la calidad de la aplicación de los test

16. Los controles sobre los test deberían ser mínimos, pues inhiben el desarrollo de nuevas ideas y nuevos procedimientos de evaluación

17. La aplicación de test por Internet **no** permite proteger la privacidad de los usuarios

18. Habría que permitir a los editores que vendiesen cualquier test que consideren adecuado

19. El Colegio Oficial de Psicólogos debería ejercer un papel más activo para regular y mejorar el uso que se hace de los test

20. La aplicación de test por Internet abre posibilidades de fraude

21. En el desempeño de mi profesión utilizo test habitualmente

22. Los test constituyen una excelente fuente de información si se combinan con otros datos psicológicos

23. Utilizados correctamente, los test son de gran ayuda para el psicólogo

24. Teniendo en cuenta todos los aspectos, creo que en la última década el uso de los test ha mejorado en mi país

25. La práctica profesional basada en evidencias potencia el uso de los test como fuentes de información

26. En la toma de decisiones profesionales doy un peso importante a los datos obtenidos mediante test

27. Antes de utilizar un test compruebo sus propiedades psicométricas

28. Estime la frecuencia de los siguientes problemas de uso de los test en su entorno profesional (1: muy poco frecuente; 5: muy habitual)

1. Hacer fotocopias de materiales sujetos a copyright
2. Hacer evaluaciones utilizando test inadecuados
3. No estar al día
4. No contrastar las interpretaciones con otros
5. No tener en cuenta los errores de medida de las puntuaciones
6. No restringir la aplicación de los test a personal cualificado
7. No tener en cuenta las condiciones locales (país, región) que pueden afectar a la validez
8. Hacer interpretaciones que van más allá de los límites del test
9. Utilizar test con baremos inadecuados
10. Utilizar algunos ítems sin los permisos correspondientes
11. Tener un precio muy elevado

29. En relación con la evaluación anual de test que realiza la Comisión de Test del COP:

1. La conozco ☐ sí ☐ no (si responde no pase a la pregunta 30, en realidad en la encuesta informatizada ya lo hace automáticamente)
2. Me parece importante para mejorar la calidad de los test
3. Consulto las evaluaciones de los test en la página web del COP
4. Me ayuda en la elección de los test que utilizo
5. Es innecesaria, los psicólogos ya tienen suficiente información para saber qué test utilizar

30. Cite los tres test que utiliza con más frecuencia en el ejercicio de su profesión:

1.
2.
3.

Observaciones: Comente cualquier otro aspecto que considere oportuno

Anexo 1

Encuesta utilizada para recoger las opiniones de los psicólogos sobre la utilización de los test (continuación)

AGRADECIMIENTOS

Los autores desean expresar su agradecimiento a todos los investigadores y profesionales que participaron en el desarrollo del cuestionario utilizado, que fueron muchos. Especial mención merecen, el presidente del COP, Francisco Santolaya, que apoyó el proyecto de forma clara y decidida, los miembros de la Comisión de Test del COP (Milagros Antón, José Luis Galve, Rosario Martínez Arias, Vicente Ponsoda y Frédérique Vallar), así como quienes coordinaron las evaluaciones anuales de test: Paula Elosua Oliden, Laura E. Gómez Sánchez, María Dolores Hidalgo, Eduardo Fonseca Pedrero y Carme Viladrich. La implementación informática de la encuesta corrió a cargo de José Ramón Correas González. Muchas gracias a todos, sin su ayuda este trabajo no hubiese sido posible.

CONFLICTO DE INTERESES

No existe conflicto de intereses

REFERENCIAS

- American Educational Research Association, American Psychological Association y National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.
- Armayones, M., Boixadós, M., Gómez, B., Guillamón, N., Hernández, E., Nieto, R., Pousada, M. y Sara, B. (2015). Psicología 2.0: oportunidades y retos para el profesional de la psicología en el ámbito de la e-salud. *Papeles del Psicólogo*, 36(2), 153-160.
- Bartram, D. (1996). Test qualifications and test use in the UK: The competence approach. *European Journal of Psychological Assessment*, 12(1), 62-71.
- Bartram, D. y Coyne, I. (1998). Variations in national patterns of testing and test use: The ITC/EFPPA international survey. *European Journal of Psychological Assessment*, 14(3), 249-260.
- Bennett, R. E. (1999). Using new technology to improve assessment. *Educational Measurement: Issues and Practice*, 18(3), 5-12.
- Bennett, R. E. (2006). Inexorable and inevitable: The continuing story of technology and assessment. En D. Bartram and R. K. Hambleton (Eds.), *Computer-based testing and the Internet*. Chichester: John Wiley and Sons. (pp. 201-217).
- Borsboom D. y Cramer, A.O.J. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, 9, 91-121.
- Breithaupt, K. J., Mills, C. N. y Melican, G. J. (2006). Facing the opportunities of the future. En D. Bartram y R. K. Hambleton (Eds.), *Computer-based testing and the Internet* (pp. 219-251). Chichester: John Wiley and Sons.
- Calderón, C., Navarro, D., Lorenzo-Seva, U. y Ferrando, P. J. (2019). Multidimensional or essentially unidimensional? A multi-faceted factor-analytic approach for assessing the dimensionality of tests and items. *Psicothema*, 31(4), 450-457.
- Carlson, J. F. y Geisinger, K. F. (2012). Test reviewing at the Buros Center for Testing. *International Journal of Testing*, 12(2), 122-135.
- Chernyshenko, O. S. y Stark, S. (2016). Mobile psychological assessment. En F. Drasgow (Ed.) (2016). *Technology and testing*. Nueva York : Routledge.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.
- Downing, S. M. y Haladyna, T. M. (Eds.) (2006). *Handbook of test development*. Hillsdale, NJ.: LEA.
- Drasgow, F. (Ed.) (2016). *Technology and testing*. Nueva York : Routledge.
- Drasgow, F., Luecht, R. M. y Bennett, R. E. (2006). Technology and testing. En R. L. Brennan (Ed.), *Educational measurement*. Westport, CT: ACE/Praeger. (págs. 471-515).
- Elosua, P. y Geisinger, K. F. (2016). Cuarta evaluación de test editados en España: forma y fondo. *Papeles del Psicólogo*, 37(2), 82-88.
- European Federation of Professional Psychologists' Associations (2005). *Meta-Code of ethics*. Brussels: Author (www.efpa.eu).
- Evers, A., McCormick, C., Hawley, L., Muñoz, J. et al. (2017). Testing practices and attitudes toward tests and testing: An international survey. *International Journal of Testing*, 17(2), 158-190.
- Evers, A., Muñoz, J., Bartram, D., Boben, D., Egeland, J., Fernández-Hermida, J. R., et al. (2012). Testing practices in the 21st Century: Developments and European Psychologists' opinions. *European Psychologist*, 17(4), 300-319.
- Evers, A., Muñoz, J., Hagemester, C., Hostmaelingen, A., Lindley, P., Sjöberg, A., y Bartram, D. (2013). Assessing the quality of tests: Revision of the EFPA review model. *Psicothema*, 25(3), 293-201.
- Fernández-Ballesteros, R. (2014). *Evaluación psicológica*. Madrid: Pirámide.
- Fernández-Ballesteros, R., De Bruyn, E., Godoy, A., Hornke, L., Ter Laak, J., y Vizcarro, C. et al. (2001). Guidelines for the assessment process (GAP): A proposal for discussion. *European Journal of Psychological Assessment*, 17(3), 187-200.

- Ferrando, P.J. y Anguiano, C. (2010). El análisis factorial como técnica de investigación en Psicología. *Papeles del Psicólogo*, 31(1), 18-33.
- Ferrando, P. J. y Lorenzo-Seva, U. (2017). Program factor at 10: Origins, development and future directions. *Psicothema* 29(2), 236–240. doi: 10.7334/psicothema2016.304.
- Fonseca-Pedrero, E. (2017). Análisis de redes: ¿una nueva forma de comprender la psicopatología? *Revista de Psiquiatría y Salud Mental*, 10(4), 183-224.
- Fonseca-Pedrero, E. (2018). Análisis de redes en psicología. *Papeles del Psicólogo*, 39(1), 1-12.
- Fonseca-Pedrero, E. y Muñoz, J. (2017). Quinta evaluación de test editados en España: mirando hacia atrás, construyendo el futuro. *Papeles del Psicólogo*, 38(3), 161-16.
- Geisinger, K. F. (2012). Worldwide test reviewing at the beginning of the twenty-first century. *International Journal of Testing*, 12(2), 103-107.
- Geisinger, K. F. (2016). Test evaluation. En S. Lane, M. R. Raymond y T. M. Haladyna (eds.), *Handbook of test development*. Nueva York: Routledge.
- Goodman, D.P. y Hambleton, R.K. (2004). Student test score reports and interpretive guides: Review of current practices and suggestions for future research. *Applied Measurement in Education*, 17(2), 145-220.
- Gómez-Sánchez, L. E. (2019). Séptima evaluación de test editados en España. *Papeles del Psicólogo*, 40(3), 205-210.
- Groth-Marnat, G. y Wright, A. J. (2016). *Handbook of psychological assessment*. Hoboken, NJ: Wiley.
- Haladyna, T. M. y Rodríguez, M. C. (2013). *Developing and validating test items*. Londres: Routledge.
- Hambleton, R. K., Merenda, P. F. y Spielberger, C. D. (Eds.) (2005). *Adapting educational and psychological tests for cross-cultural assessment*. Londres: LEA.
- Hernández, A., Tomás, I., Ferreres, A. y Lloret, S. (2015). Tercera evaluación de test editados en España. *Papeles del Psicólogo*, 36(1), pp. 1-8
- Hernández, A., Ponsoda, V., Muñoz, J., Prieto, G. y Elosua, P. (2016). Revisión del modelo para evaluar la calidad de los test utilizados en España. *Papeles del Psicólogo*, 37(3), 192-197.
- Hidalgo, M. D. y Hernández, A. (2019). Sexta evaluación de test editados en España: resultado e impacto del modelo en docentes y editoriales. *Papeles del Psicólogo*, 40(1), 21-30.
- IBM Corp. (2016). *IBM SPSS Statistics for Windows, Version 24.0*. Armonk, NY: IBM Corp.
- Insel, T. R. (2017). Digital phenotyping: Technology for a new science of behavior. *JAMA*, 318(2), 1215-1216.
- Irwing, P., Booth, T. y Hughes, D. J. (Eds.) (2018). *The Wiley handbook of psychometric testing: A multidisciplinary reference on survey, scale and test development*. UK: John Wiley y Sons Ltd.
- ITC Guidelines for Translating and Adapting Tests (Second Edition) (2017). *International Journal of Testing*, 17, 1-34, (www.InTestCom.org)
- Izquierdo, I., Olea, J. y Abad, F. J. (2014). Exploratory factor analysis in validation studies: Uses and recommendations. *Psicothema*, 26(3), 395-400.
- Joint Committee on Testing Practices. (2002). *Ethical principles of psychologists and code of conduct*. Washington DC: Joint Committee on Testing Practices.
- Knapp, S.J. (Ed.) (2012). *APA handbook of ethics in psychology*. Washington, DC: American Psychological Association.
- Koocher, G. y Kith-Spiegel, P. (2007). *Ethics in psychology*. Nueva York: Oxford University Press.
- Kosinski, M., Stillwell, D., y Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behaviour. *Proceedings of the National Academy of Sciences (PNAS)*, 110, 5802-5805.
- Lane, S., Raymond, M.R. y Haladyna, T. M. (2016). *Handbook of test development (2nd edition)*. Nueva York, NY: Routledge.
- Leach, M. y Oakland, T. (2007). Ethics standards impacting test development and use: A review of 31 ethics codes impacting practices in 35 countries. *International Journal of Testing*, 7(1), 71-88.
- Lindsay, G., Koene, C., Ovreeide, H., y Lang, F. (2008). *Ethics for European psychologists*. Gottingen and Cambridge, MA: Hogrefe.

- Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., y Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: una guía práctica, revisada y actualizada. *Anales de Psicología*, 30(3), 1151-1169.
- Lorenzo-Seva, U. y Ferrando, P. J. (2013). FACTOR 9.2 a comprehensive program for fitting exploratory and semiconfirmatory factor analysis and IRT models. *Applied Psychological Measurement*, 37(6), 497-498. doi: 10.1177/0146621613487794
- Markovetz, A., Blaszkiewicz, K., Montag, C., Switala, C. y Schlaepfer, T. E. (2014). Psycho-Informatics: Big Data shaping modern psychometrics. *Medical Hypotheses*, 82(4), 405-411.
- Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. Minneapolis: University of Minnesota Press.
- Miller, G. (2012). The smartphone psychology manifesto. *Perspectives on Psychological Science*, 7(3), 221-237.
- Muñiz, J. (1997). Aspectos éticos y deontológicos de la evaluación psicológica. En A. Cordero (Ed.), *Evaluación psicológica en el año 2000* (pp. 307-345). Madrid: TEA Ediciones.
- Muñiz, J. (2018). *Introducción a la psicometría*. Madrid: Pirámide.
- Muñiz, J. y Bartram, D. (2007). Improving international tests and testing. *European Psychologist*, 12(3), 206-219.
- Muñiz, J., Bartram, D., Evers, A., Boben, D., Matesic, K., Glabeke, K., Fernández-Hermida, J.R. y Zaal, J. (2001). Testing practices in European countries. *European Journal of Psychological Assessment*, 17(3), 201-211.
- Muñiz, J., Elosua, P. y Hambleton, R.K. (2013). Directrices para la traducción y adaptación de los tests: segunda edición. *Psicothema*, 25(2), 151-157.
- Muñiz, J. y Fernández-Hermida, J.R. (2000). La utilización de los test en España. *Papeles del Psicólogo*, 76, 41-49.
- Muñiz, J. y Fernández-Hermida, J. R. (2010). La opinión de los psicólogos españoles sobre el uso de los test. *Papeles del Psicólogo*, 31(1), 108-121.
- Muñiz, J., Fernández-Hermida, J. R., Fonseca, E., Campillo, A. y Peña, E. (2011). Evaluación de test editados en España. *Papeles del Psicólogo*, 32(2), 113-128.
- Muñiz, J. y Fonseca-Pedrero, E. (2019). Diez pasos para la construcción de un test. *Psicothema*, 31(1), 7-16. doi: 10.7334/psicothema2018.291.
- Muñiz, J., Hernández, A. y Ponsoda, V. (2015). Nuevas directrices sobre el uso de los test: investigación, control de calidad y seguridad. *Papeles del Psicólogo*, 36(3), 161-173.
- Muñiz, J., Prieto, G., Almeida, L., y Bartram, D. (1999). Test use in Spain, Portugal and Latin American countries. *European Journal of Psychological Assessment*, 15(2), 151-157.
- Myin-Germeys, I., Kasanova, Z., Vaessen, T., Vachon, H., Kirtley, O., Viechtbauer, W., y Reininghaus, U. (2018). Experience sampling methodology in mental health research: New insights and technical developments. *World Psychiatry*, 17(2), 123-132.
- Nelson, B., McGorry, P. D., Wichers, M., Wigman, J. T. y Hartmann, J. A. (2017). Moving from static to dynamic models of the onset of mental disorder. *JAMA Psychiatry*, 74(5), 528-534.
- Papeles del Psicólogo (2009). *Número monográfico sobre Ética Profesional y Deontología*. Vol. 30, 182-254.
- Phelps, R. (Ed.) (2005). *Defending standardized testing*. Londres: LEA.
- Phelps, R. (Ed.) (2008). *Correcting fallacies about educational and psychological testing*. Washington: APA.
- Ponsoda, V. y Hontagas, P. (2013). Segunda evaluación de test editados en España. *Papeles del Psicólogo*, 34(2), 82-90.
- Ryan, J. M. (2006). Practices, issues, and trends in student test score reporting. En S. M. Downing y T. M. Haladyna (Eds.) (2006). *Handbook of test development*. Hillsdale, NJ.: LEA. (págs. 677-710).
- Schmeiser, C. B. y Welch, C. (2006). Test development. En R. L. Brennan (Ed.), *Educational measurement (4th edition)* (pp. 307-353). Westport, CT: American Council on Education/Praeger.
- Sireci, S. y Faulkner-Bond, M. (2014). Validity evidence based on test content. *Psicothema*, 26(2), 100-107.
- Sireci, S. y Zenisky, A. L. (2016). Computerized innovative item formats: Achievement and credentialing. En S. Lane, M. R. Raymond, y T. M. Haladyna (Eds.), *Handbook of test development*. Nueva York: Routledge.
- Trull, T. J. y Ebner-Priemer, U. W. (2013). Ambulatory assessment. *Annual Review of Clinical Psychology*, 9, 151-176.

- van Os, J., Delespaul, P., Wigman, J., Mying-Germays, I. y Wichers, M. (2013). Beyond DSM and ICD: Introducing precision diagnosis for psychiatry using momentary assessment technology. *World Psychiatry*, 12(1), 113-117.
- Wright, A. J. (2011). *Conducting psychological assessment: A guide for practitioners*. Hoboken, NJ: Wiley.
- Zenisky, A. L. y Hambleton, R. K. (2016). A model and good practices for score reporting. En S. Lane, M. R. Raymond y T. M. Haladyna (eds.), *Handbook of test development*. Nueva York: Routledge.