

# El uso de los Test en España y Latinoamérica: Perspectivas actuales y retos futuros

**Ana Hernández**

*Departamento de Metodología de las Ciencias del Comportamiento e IDOCAL,  
Universidad de Valencia, España*

**Paula Elosua**

*Departamento de Psicología Clínica y de la Salud y Métodos de Investigación  
Universidad del País Vasco, España*

**Solange Wechsler**

*Centro de Ciencias da Vida, Curso Psicologia*

**José Muñiz**

*Universidad Nebrija, España*

Correspondencia: José Muñiz. e-mail: [jmuniz@nebrija.es](mailto:jmuniz@nebrija.es)

## Resumen

El objetivo del presente trabajo es ofrecer una visión general sobre la práctica de los test en España y Latinoamérica, así como apuntar algunos de los retos que tiene planteados la evaluación psicológica cara al futuro. En primer lugar, se revisa cómo se lleva a cabo la formación psicométrica de los profesionales de la psicología en los distintos países, así como las acciones y proyectos desarrollados para lograr que los usuarios de los test tengan las competencias requeridas. A continuación, se analizan las acciones encaminadas a mejorar la calidad de los test y su uso, así como las opiniones de los psicólogos respecto de los test en sus respectivos países, las cuáles resultan clave para poder diseñar acciones de mejora. Finalmente se hacen algunas reflexiones sobre el futuro de la práctica de los test, el cual viene marcado, en gran medida, por los grandes avances de las nuevas tecnologías digitales.

## Abstract

*The goal of this paper is to provide an overview of current testing practices in Spain and Latin America, and discuss some of the challenges that psychological testing will have to face in the next future. We review how professionals acquire their Psychometric knowledge and skills and what actions are taken in order to ensure that test users have the required competences. In addition, we review a number of actions and projects aimed to improve test and testing practices, and also psychologist's opinions about testing practices in their own countries. These opinions are crucial to design and successfully implement actions aimed to test and testing improvement. Finally, we reflect on the future of testing practices, which will be highly influenced by digital technological advances.*

## Palabras clave

Test, uso de test, España, Latinoamérica.

## Keywords

Test, use of test, Spain, Latin america.

## Introducción

En el presente trabajo de revisión se pretende dar una visión general sobre la práctica de los test en España y Latinoamérica, así como apuntar algunos de los retos que tiene planteados la evaluación psicológica cara al futuro. No se trata de una revisión exhaustiva, que excede las posibilidades de un artículo como éste, si bien esperamos que ayude a los lectores a hacerse una idea fidedigna acerca del estado de la cuestión, tal como nos propusieron los editores de la revista.

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. Si bien el proceso de evaluación psicológica es más amplio que la mera aplicación de los test, estos constituyen unas herramientas fundamentales a la hora de obtener evidencias empíricas para la toma de decisiones de los psicólogos (Fernández-Ballesteros, 2014; Fernández-Ballesteros et al., 2001; Groth-Marnat y Wright, 2016; Muñiz et al., 2020; Wright, 2011). Para que los datos provenientes de los test puedan ser utilizados con garantías deben darse tres condiciones básicas: que los test exhiban unas condiciones psicométricas adecuadas, que los profesionales tengan la formación apropiada, y que los test se utilicen de forma correcta (Elosua y Muñiz, 2013; Elosua, 2017). Si se dan estas tres condiciones las evidencias provenientes de las puntuaciones obtenidas mediante los test serán de gran ayuda para los psicólogos. Actualmente disponemos de conocimientos psicométricos y buenas fuentes de información para llevar a cabo una construcción y análisis riguroso de los test (AERA, APA y NCME, 2014; Downing y Haladyna, 2006; Haladyna y Rodríguez, 2013; Irwing et al., 2018; Lane et al., 2016; Muñiz, 2018; Muñiz y Fonseca-Pedrero, 2019; Schmeiser y Welch, 2006). Sin embargo, con frecuencia se observa que muchos test utilizados en la práctica profesional no están a la altura de los conocimientos y técnicas que nos ofrece la psicometría actual. En cuanto a la formación psicométrica de los profesionales, existen bastantes variaciones de unos países a otros, como se comentará más adelante. Una formación psicométrica adecuada de los profesionales que usan los test es esencial. El tercer aspecto, el uso adecuado de los test, es clave, pues de nada vale que el test tenga unas propiedades psicométricas adecuadas y el profesional esté bien formado si luego se lleva a cabo una utilización incorrecta. Conscientes de ello, las organizaciones nacionales e internacionales, tales como los Colegios Oficiales de Psicólogos (COP), la *International Test Commission* (ITC), la Federación Europea de Asociaciones de Psicólogos (EFPA), o las propias universidades y editoriales de test, vienen realizando distintas actividades y proyectos encami-

nados a mejorar los tres aspectos citados: calidad psicométrica de los test, formación de los profesionales y uso adecuado. Estas acciones son muy variadas, si bien pueden encuadrarse dentro de 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, y varían mucho de unos países a otros (Bartram, 1996; Bartram y Coyne, 1998; Evers et al., 2012, 2017; Muñiz et al., 1999). Estas medidas son necesarias, pero no constituyen una garantía total, por lo que hay que complementarlas con otras de carácter *informativo*. Como señalan Muñiz et al. (2020), entre estas acciones informativas 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 et al. (2012), Koocher y Keith-Spiegel (2007), Lindsay et al. (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 et al., 2005; ITC, 2017; Muñiz et al., 2013), o para su uso en investigación, control de calidad y seguridad (Muñiz et al., 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](http://www.intestcom.org)), la EFPA ([www.efpa.eu](http://www.efpa.eu)), o el COP ([www.cop.es](http://www.cop.es)) en el apartado dedicado a la Comisión de Test (Muñiz et al. 2020). Una buena síntesis de las buenas prácticas evaluativas aparecen en la Norma ISO 10667 (AENOR, [www.aenor.es](http://www.aenor.es)), que regula todo lo relativo a la evaluación de personas en entornos laborales y organizacionales. 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 e instituciones que la adoptan (Muñiz et al., 2020). Otra manera de diseminar información sobre los test es llevar a cabo evaluaciones de su calidad psicométrica por parte de expertos y difundirla entre

los profesionales para ayudarles a valorar de forma rigurosa los test que pueden utilizar. Esta actividad varía de unos países a otros, como se verá más adelante. 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 cualitativa 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 (Muñiz et al., 2020). El modelo europeo puede consultarse en la página web de la EFPA ([www.efpa.eu](http://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](http://www.cop.es)) y el proceso de adaptación seguido en Hernández et al. (2016). Estas evaluaciones proporcionan una información de primera mano a los profesionales sobre la calidad de los test, y sirven 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; Muñiz et al., 2020).

Dentro de este contexto, vamos a revisar algunos de los aspectos más relevantes de las acciones llevadas a cabo en España y Latinoamérica en torno a la práctica de los test, comentando finalmente algunas tendencias y retos de futuro que tiene planteados la evaluación psicológica. Articularemos la información en torno a tres grandes apartados: la formación psicométrica de los profesionales, la evaluación de la calidad de los test, y las opiniones de los psicólogos sobre la situación de los test en sus respectivos países. Finalmente se harán algunas reflexiones sobre el futuro de la práctica de los test, el cual viene marcado por los grandes avances de las nuevas tecnologías digitales.

## El caso de España

### Formación psicométrica de los profesionales

En España la formación psicométrica se da fundamentalmente en la universidad, dentro de las Facultades de Psicología. Existen cincuenta y ocho Facultades de Psicología, de las cuales en torno a un 80% son públicas. El grado de Psicología son cuatro años, existiendo al menos un cuatrimestre dedicado a la asignatura obligatoria de Psicometría, lo que representa un promedio de unas 180 horas entre clases teóricas, prácticas, laboratorios y tutorías, a las que habría que añadir el trabajo autónomo y grupal del alumnado. Además, es habitual que las facultades ofrezcan otras asignaturas obligatorias y optativas relacionadas con la Psicometría, tales como Evaluación Psicológica, Estadística y Análisis de Datos, Diseño de Experimentos, Elaboración de Encuestas y Técnicas Multiva-

riadas. Si bien existen algunas variaciones entre las distintas facultades, esa es la situación más habitual en los estudios de grado, pudiendo decirse que los alumnos tienen la oportunidad de alcanzar una formación metodológica básica razonable. Una vez terminada esta formación básica hay varias formas de seguir formándose en psicometría, la más habitual es entrar a trabajar con los grupos de investigación especializados en psicometría, de los cuales existen varios en España, algunos de ellos de gran prestigio y proyección internacional. Además, la Universidad Autónoma de Madrid, la Complutense y la Universidad Nacional de Educación a Distancia (UNED), ofrecen un Máster de Metodología de 750 horas (dos años), que incluye Psicometría, con dos opciones: presencial y online. Se trata de un máster prestigioso, con una gran demanda. El máster ofrece cinco líneas de especialización: Diseños de Investigación y Análisis de Datos, Medición y Evaluación, Modelado de Procesos, Investigación Clínica y Big Data. Toda la información puede consultarse en <http://www.metodologiaccs.es/>. Paralelamente las cuatro editoriales de test que operan en España (CEPE, GiuntiEOS Psychometrics, Pearson y TEA) ofrecen su propia formación para el uso de los test específicos que comercializan. Finalmente, cabe señalar que el Consejo General de la Psicología ofrece a sus colegiados un Programa de Formación Continuada a Distancia (FOCAD) que les da la posibilidad de actualizarse en distintos campos, incluyendo la construcción de test, con dos cursos de los que ya se han impartido distintas ediciones (Elosua, 2019; Muñiz y Fonseca-Pedrero, 2017).

Puede decirse, en suma, que existe un conjunto amplio de oportunidades de formación psicométrica, si bien podrían mejorarse en varias direcciones. Una de ellas sería la acreditación de los profesionales en el uso de los test, con varios niveles de especialización, como ya hacen algunos países europeos. También sería deseable desarrollar algún máster específico de Psicometría, que cubriese tanto los aspectos académicos como los profesionales. Se constata una escasez de psicómetras en la mayoría de los países, como demuestran los simposios organizados sobre ese tema en el congreso de 2018 de la *National Council on Measurement in Education* (NCME), celebrado en Toronto, o el de 2021 de la ITC.

## Evaluación de la calidad de los test

Como ya se ha comentado, una de las acciones llevadas a cabo dentro de la estrategia *informativa* para tratar de mejorar todo lo relativo al uso de los test es la evaluación de su calidad psicométrica por parte de expertos. Estas evaluaciones se hacen públicas para que los profesionales puedan valerse de ellas a la hora de elegir y aplicar los test. Utilizando el modelo CET-R (Hernández et al., 2016) ya se han llevado a cabo en España siete evaluaciones anuales de test (Elosua y Gei-

singer, 2016; Fonseca-Pedrero y Muñiz, 2017; Gómez-Sánchez, 2019; Hernández et al., 2015; Hidalgo y Hernández, 2019; Muñiz et al., 2011; Ponsoda y Hontagas, 2013). Actualmente se está completando la octava evaluación, coordinada por la profesora Carme Viladrich, e iniciándose la novena, que coordina el profesor Luis Manuel Lozano. Hasta la fecha se han revisado un total de setenta y dos test, pudiendo consultarse libremente las revisiones en la página web de la Comisión de Test del COP (<https://www.cop.es/index.php?page=evaluacion-tests-editados-en-espana>). Lo ideal sería que la mayoría de los test publicados estuviesen revisados, como ocurre en Holanda, no obstante, en el contexto europeo España es de los países que más avanzado lleva el proyecto de evaluación de test. Por otro lado, hay que destacar que, con cada aplicación anual del modelo se ha realizado una valoración crítica del proceso de revisión y se han ido proponiendo mejoras en el proceso. Este aspecto es fundamental, ya que la implementación de un sistema de garantía de la calidad requiere de la evaluación del propio sistema de calidad para poder mantener y mejorar su buen funcionamiento. También hay que destacar el papel de las editoriales españolas 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. Como señalan Muñiz et al. (2020), 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. De hecho, aunque remarcando algunos aspectos mejorables, las editoriales reconocen que el modelo CET-R resulta una guía útil tanto para el desarrollo y adaptación del test, como para la edición y confección del manual del test, y que el modelo supone un sistema de referencia de mejora continua (Hidalgo y Hernández, 2019). Un tercer aspecto nada desdeñable es su importancia en la docencia y formación de los futuros psicólogos (Hidalgo y Hernández, 2019; Muñiz et al., 2020).

## Opiniones de los psicólogos sobre los test

Otra de las acciones de carácter *informativo* encaminada a mejorar el uso de los test es tratar de conocer de forma objetiva las opiniones de los profesionales sobre los test. Conocer esta opinión es clave para promover acciones y proyectos encaminados a mejorar aquellos aspectos que sean percibidos como deficitarios por los profesionales (Muñiz et al., 2020). Con este fin la EFPA ha impulsado la aplicación de una encuesta para conocer las opiniones de los psicólogos sobre el uso de los test en distintos países europeos. Los primeros datos en seis países europeos se obtuvieron en el año 2000 (Muñiz et al., 2001), y una buena panorá-

mica sobre los resultados obtenidos a partir de entonces en Europa puede verse en Evers et al. (2012). La misma encuesta fue utilizada también en otros países no europeos (Evers 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). El estudio se ha repetido recientemente, con nuevas actualizaciones (Muñiz et al., 2020). En concreto, el cuestionario contaba con un total de 31 ítems referidos a distintas cuestiones como las actitudes hacia los test, la necesidad de controlar y regular los test y su uso, la valoración de la formación y los conocimientos sobre los test, el uso de internet y la informatización y automatización de los test y, finalmente, los problemas en el uso de los test. Se preguntó además por los tres test más utilizados por los profesionales encuestados y, por primera vez, se incluyeron preguntas sobre el modelo CET-R. A continuación, se ofrecen algunos de los resultados más relevantes obtenidos en la última encuesta (Muñiz et al., 2020), tanto globales como diferenciados por área de especialización (Tablas 1-6). También se ofrece la comparativa de resultados a través de los años (Tabla 4), y las conclusiones generales derivadas de los resultados. Los resultados completos y detallados pueden consultarse en Muñiz et al. (2020).

**Tabla 1.- Media y desviación típica de cada uno de los ítems de la encuesta por especialidades (Clínica-Salud, Educativa, Trabajo) y global. (Tomada de Muñiz et al., 2020)**

Ítems	Clínica-Salud N=844		Educativa N=159		Trabajo N=69		Global N=1.248	
	Me- dia	DT	Me- dia	DT	Me- dia	DT	Me- dia	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
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 (C), Educativa (E) y Trabajo (T). La tabla excluye la categoría "otros" que aglutina áreas como deporte, jurídica, tráfico y servicios sociales, entre otras

**Tabla 2.- Media y desviación típica de las dimensiones por especialidades (Clínica-Salud, Educativa, Trabajo) (Tomada de Muñiz et al., 2020)**

Dimensiones	Clínica-Salud N=844		Educativa N=159		Trabajo N=69		Global N=1.248	
	Me- dia	DT	Media	DT	Me- dia	DT	Me- dia	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 (C), Educativa (E) y Trabajo (T). La tabla excluye la categoría "otros" que aglutina áreas como deporte, jurídica, tráfico y servicios sociales, entre otras

**Tabla 3.- Opiniones de los psicólogos sobre la Evaluación de la calidad de los test llevada a cabo anualmente por el COP (Tomada de Muñiz et al., 2020)**

Items	Clínica-Salud N=174		Educativa N=42		Trabajo N=18		Global N=281	
	Me- dia	DT	Media	DT	Me- dia	DT	Me- dia	DT
Me parece importante para mejorar la calidad de los test	4,43	0,72	4,43	,63	4,61	0,98	4,46	0,72
Consulto las evaluaciones de los test en la página web del COP	3,78	1,17	4,05	0,80	4,17	0,71	3,81	1,15
Me ayuda en la elección de los test que utilizo	3,67	1,09	3,83	1,03	3,78	1,00	3,65	1,10
Es innecesaria, los psicólogos ya tienen suficiente información para saber qué test utilizar	1,62	1,01	1,64	1,06	1,44	0,98	1,62	1,01

DT: desviación típica

Nota: A los resultados globales presentados por Muñiz et al. (2020) se añade la comparación por especialidades. La tabla excluye la categoría "otros" que aglutina áreas como Deporte, Jurídica, Tráfico y Servicios Sociales, entre otras

**Tabla 4.- Medias de los ítems obtenidas en las encuestas de los años 2000, 2010 y 2020 (Tomada de Muñiz et al., 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 5.- Los 25 test más utilizados por los psicólogos españoles (Tomada de Muñiz et al., 2020)**

	Nombre de la prueba	N	%
WISC	Weschler Intelligence Scale for Children	289	23,2%
MCFI	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 6.- Los diez test más utilizados por los psicólogos españoles por especialidades (Tomada de Muñiz et al., 2020)**

	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). La tabla excluye la categoría "otros" que aglutina áreas como deporte, jurídica, tráfico y servicios sociales, entre otras

Se incluyen, a continuación, las **conclusiones generales** de Muñiz et al. (2020) sobre los datos obtenidos en la encuesta:

- a. 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.
- b. 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).
- c. 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.

- d. 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.
- e. 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.
- f. 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-Salud, seguido de Educativa y Trabajo.
- g. 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.
- h. 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. La tendencia sugiere que son los psicólogos del trabajo los que valoran más positivamente estas evaluaciones y la información que aportan.
- i. 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.
- j. 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).

## Situación en Latinoamérica y Portugal

### Formación psicométrica de los profesionales

En los últimos años el interés por el análisis de los programas de formación en Psicología ha dado lugar a varios trabajos con enfoques y perspectivas diferentes, si bien todos ellos constatan las notables diferencias entre países (Gallegos, 2016). Los trabajos pioneros de Alarcón (1999, 2004) ofrecen un balance sobre los cincuenta primeros años de los programas de formación en Psicología que se inician en los años 1947 y 1948. Según este autor las prácticas relacionadas con la Psicometría, en lo que se refiere a la adaptación y construcción de test, han sido fundamentales en el desarrollo de la Psicología en América Latina. Los trabajos de revisión constatan marcadas diferencias en la formación del psicólogo entre los países iberoamericanos. Dentro del proyecto Mercosur se creó la asociación de universidades del grupo Montevideo (AUGM) como espacio de referencia para el diseño de políticas de integración (Saenz, 2016). En 1997 se estableció el protocolo de admisión de títulos y grados universitarios para el ejercicio de actividades académicas en los países del Mercosur y el grupo de trabajo Psicólogos del Mercosur reconoció en 1998 que la formación básica en psicología debía de incluir programas sobre evaluación y diagnóstico psicológico a la vez que debía garantizar una formación científica que reconociera a la Psicología como ciencia (Benito, 2009; Saenz, 2016).

En este contexto general encontramos programas y tradiciones bien diferenciadas. Así, en el caso de Perú, el profesor León (2009) reivindica la figura de Reynaldo Alarcón como uno de los impulsores de una tradición psicométrica que hoy sigue viva. La revisión de la situación de la formación en psicología en México que llevan a cabo Zanatta y Yuren (2012) recoge el reconocimiento temprano de la Psicometría como asignatura clave en la formación del psicólogo. Por su parte, Brasil se caracteriza por un elevado desarrollo en el área de evaluación psicológica, sobre todo en las últimas dos décadas. Este desarrollo se relaciona con las exigencias de calidad de las pruebas psicológicas demandadas por parte del Consejo Federal de la Psicología que propició un aumento considerable de cursos específicos de formación (Wechsler, 2013; Wechsler et al., 2019). En Bolivia, Schulmeyer (2016) llevó a cabo una revisión de los planes de estudio de las universidades públicas y privadas y constató que, en las materias del área de la evaluación psicológica, existe poco consenso. La autora cita veinte nombres distintos y contenidos diferentes relacionados con la formación en teoría psicométrica y uso de pruebas psicométricas y proyectivas. En la revisión sobre el estado de la evaluación psicológica y de la psicometría en Chile, las profesoras

Vinet y Gonzalez (2013) mencionan el objetivo del Consorcio de Universidades Estatales de Chile (CUE) de unificar programas e incluir contenidos en materias de Evaluación Psicológica y Psicometría. No obstante, el peso de las universidades públicas es pequeño con relación a las instituciones privadas. En Colombia son varios los trabajos que apuntan a la necesidad de fortalecer la formación universitaria en aspectos relacionados con la psicometría (Puche, 2003; Rodríguez-Jiménez et al., 2011; Rodríguez-Jiménez y Rosero-Burbano, 2014). En Portugal la disciplina de Psicometría es reciente, pues la Psicología se introdujo en las universidades en 1975, en consecuencia, esta área está menos desarrollada en comparación con la mayoría de los países de Europa occidental (Wechsler et al., 2014). Finalmente, en Argentina, la regulación de la formación en Psicología, con una fuerte tradición psicoanalítica, reconoce nueve áreas de contenido, tres de las cuales se relacionan de alguna manera con la Psicometría: Investigación en Psicología, la Evaluación y Diagnóstico Psicológicos, y la ética y Deontología Profesional (Klappenbach, 2015; Klappenbach y Gallegos, 2018).

El papel de las universidades en la formación en Psicología se refuerza con el trabajo de las asociaciones profesionales que a través de cursos, congresos y actividades diversas intentan mejorar y actualizar los perfiles profesionales de sus miembros. En este sentido, una de las sociedades de referencia es la Sociedad Interamericana de Psicología (SIP). Fundada en México el año 1951 (Alarcón, 2004; Ardila, 2004), está compuesta principalmente por psicólogos de América del Norte, Centroamérica, El Caribe y Sur América. La SIP promueve la cooperación en la docencia, investigación y el ejercicio profesional de la Psicología, siendo la sociedad de Psicología más antigua de Latinoamérica (Gallegos et al., 2018). Además de la SIP, en el año 2002 se fundan dos asociaciones profesionales integradas por entidades de Psicología: la Federación Iberoamericana de Asociaciones de Psicología (FIAP) en Colombia, y la Unión Latinoamericana de entidades de Psicología (ULAPSI) en la Ciudad de Puebla en México. Asimismo, a lo largo y ancho de los países latinoamericanos se han ido constituyendo asociaciones profesionales relacionadas con diferentes campos de la Psicología (Ardila, 1998). En relación con la Evaluación Psicológica, en 1995, se constituyó la Asociación Iberoamericana de Diagnóstico y Evaluación Psicológica (AIDEP), que edita la Revista Iberoamericana de Diagnóstico y Evaluación, indexada actualmente en las bases de datos más relevantes. Cabe destacar también la reciente creación dentro de la SIP de la división en Evaluación y Psicometría, lo cual viene a demostrar que la colaboración internacional es fundamental para el desarrollo del conocimiento científico en general y para la mejora en el uso de los test en particular (Elosua y Muñiz, 2013; García et al., 2014; López et al., 2011).

## Evaluación de la calidad de los test

Entre los países latinoamericanos el único que dispone de un proceso de revisión sistemática de test es Brasil. El Consejo Federal de la Psicología viene publicando desde el año 2001 regulaciones sobre la calidad de los test que se apoyan, entre otras, en las directrices de la Comisión Internacional de Test (ITC). Existe un comité de expertos auspiciado por el Consejo que evalúa cada una de las pruebas que se publicarán en el país. Los resultados de la evaluación se ponen a disposición de los interesados (SATEPSI- Sistema de Avaliação de Testes Psicológicos, 2020). En este sentido la aproximación a la mejora en el uso de test encaja en la estrategia restrictiva, considerándose una falta de ética profesional el que un psicólogo utilice pruebas no aprobadas por esta comisión (Primi y Nunes, 2010). Nos consta que en la actualidad el Colegio Colombiano de Psicología está estudiando la implementación de un programa de estudio de la calidad de los test. Parece claro que la evaluación de la calidad de los test por parte de expertos constituye una asignatura pendiente en la psicología Latinoamericana, que sería aconsejable impulsar en los próximos años.

## Opiniones de los psicólogos sobre los test

Actualmente en Latinoamérica se está llevando a cabo un estudio de carácter internacional con el objetivo de obtener una descripción detallada de las opiniones de los psicólogos sobre el uso de test en la práctica profesional. En el estudio han participado 3.638 psicólogos y psicólogas que ejercen su labor en Argentina, Bolivia, Brasil, Chile, Colombia, Guatemala, México, Perú, Portugal y Venezuela (Tabla 7). De ellos el 34,8% se dedican profesionalmente al área de Psicología Clínica y de la Salud; el 16,9% trabajan en el área de Psicología del Trabajo y de las Organizaciones; el 13,9% son especialistas en el área de Psicología Educativa, y el resto trabajan en campos como la Psicología del Deporte, Neuropsicología o Psicología Forense, o la enseñanza universitaria.

El cuestionario se construyó utilizando como referencia el estudio de la EFPA, y adaptando algunos de los ítems a la situación latinoamericana. Hay que destacar este proyecto como un claro exponente de la importancia de la colaboración internacional (Elosua, 2017), ya que cuenta con el apoyo de la Federación Iberoamericana de Asociaciones de Psicología (FIAP), la Sociedad Interamericana de Psicología (SIP), el Consejo General de la Psicología de España, la Comisión Internacional de Test y los colegios profesionales de los países participantes. Si bien los resultados del proyecto serán publicados en breve con todo detalle, avanzamos aquí algunas de las conclusiones más relevantes derivadas del análisis preliminar de datos:

**Tabla 7.- Participantes en el estudio latinoamericano sobre actitudes hacia los test**

Países participantes	Respuestas completas
Argentina	63
Bolivia	138
Brasil	867
Chile	350
Colombia	760
Guatemala	68
México	89
Perú	161
Portugal	1021
Venezuela	121

- a. Los ítems de la encuesta utilizada se articulan en torno a cuatro áreas 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*. Esta estructura es muy similar a la encontrada en España, las diferencias se explican porque el cuestionario latinoamericano únicamente contiene un ítem referido al uso de internet en la evaluación, con lo cual la quinta dimensión española carece de contenido.
- b. Se constata una actitud positiva de los psicólogos latinoamericanos hacia los test. Sobre una escala de 5 puntos la media aritmética del ítem "Utilizados correctamente los test son de gran ayuda para el psicólogo" ha sido 4,63 (DT=0,83). Este valor se mantiene constante con pequeñas fluctuaciones para la mayoría de los países participantes en el proyecto ( $h=0,02$ ).
- c. En lo relativo a los problemas de uso de los test, la consideración de la gravedad de prácticas no adecuadas es mayor que la valoración realizada por los psicólogos españoles: la media de los ítems que componen esta escala es de 3,35, frente a 2,93 en el caso de España.
- d. Los psicólogos latinoamericanos ven necesaria la implementación de sistemas de regulación sobre el uso de los test. La media del ítem que indaga sobre la necesidad de legislación oficial específica es muy elevada: 4,29. Este valor se repite cuando se pregunta a los psicólogos sobre la necesidad de control por parte de las asociaciones profesionales ( $M=4,27$ ).
- e. Coinciden los psicólogos latinoamericanos con los españoles al valorar la importancia de la mejora y adecuación de la formación en el uso de test.

Consideran que: a) la formación recibida proviene en su mayor parte del grado (Media=3); b) el grado debiera incluir formación psicométrica específica (Media=4,42); y c) el profesional no dispone de información suficiente sobre la calidad de los test (Media=2,99). Este dato pone de manifiesto la necesidad de potenciar la evaluación de la calidad de los test en Latinoamérica.

- f. En cuanto a los test más utilizados por los profesionales latinoamericanos observamos una preferencia mayor hacia las pruebas proyectivas (Dibujo de la Figura Humana, Casa Árbol y Persona – HTP, Rorschach, Wartegg, CAT, TAT, Persona Bajo la Lluvia) en comparación con la observada entre los psicólogos europeos. Los test de inteligencia más mencionados son las escalas Wechsler (niños y adultos), y el test de Matrices Progresivas de Raven. Entre las escalas de personalidad las que muestran una mayor frecuencia de uso son el Inventario de Personalidad de Millon y el 16 PF de Cattell.

**Tabla 8- Test más utilizados en Latinoamérica**

Nombre de la prueba	
DFH	Test de la figura humana
HTP	House-Tree-Person
WISC	Wechsler Intelligence Scale for Children
MCFI	Millon Clinical Multiaxial Inventory
16PF	16 Personality Factors
	Test de relaciones objetales de Phillipson
	Rorschach
	Test de Wartegg
	Test Persona bajo la lluvia
TAT	Test de apercepción temática
	Test de la familia kinética
WAIS	Wechsler Adult Intelligence Scale
RAVEN	Matrices Progresivas de Raven
CAT	Test de apercepción temática para niños
EPQ	Cuestionario de personalidad de Eysenck
	Fábulas de Düss
FROSTIG	Test de Desarrollo de la Percepción Visual
BENDER	Test gestáltico visomotor niños
MMPI	Minnesota Multiphasic Personality Inventory
PFISTER	Test de las pirámides coloridas
ROTTER	Escala de Locus de Control Interno-Externo
	Test de Zulliger
REY	Test de copia de una figura compleja
BENDER	Test gestáltico visomotor adultos

Este avance de resultados describe un perfil diferenciado de los psicólogos latinoamericanos con respecto a los psicólogos europeos, particularmente los españoles, en lo que respecta a los test más utilizados, con una mayor preferencia hacia el uso de técnicas proyectivas. Los datos requieren de un análisis detallado posterior que analice los usos y costumbres por país y permita profundizar en la heterogeneidad y riqueza latinoamericana.

## Mirando hacia el futuro

El futuro ya está aquí, toda la evaluación psicológica y la práctica profesional están incorporando gradualmente todo un conjunto de nuevos modelos psicométricos y avances tecnológicos, tales como la Teoría de la Respuesta a los Ítems (TRI), los modelos de diagnóstico clínico, los modelos de redes, los Test Adaptativos Informatizados (TAI), la obtención de puntuaciones automatizadas, la generación de informes automatizados, la innovación en nuevos formatos de ítems/test, por citar sólo algunos ejemplos. Todas estas innovaciones vienen propiciadas por las nuevas tecnologías de la información y la comunicación, y en especial los avances informáticos, multimedia e Internet (Muñiz, 2018; Muñiz y Fonseca-Pedrero, 2019). De hecho, expertos como Bennett (1999, 2006), Breithaupt et al., (2006), Drasgow (2016), Drasgow et al. (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. Esta tendencia se enmarca en 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. Aparecen así nuevos enfoques psicométricos conectados con el análisis de las grandes bases de datos (*big data*) de las que se dispone en la actualidad (Markovetz et al., 2014). Además, el uso de 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), con especial relevancia para la evaluación ambulatoria. La evaluación ambulatoria abarca diferentes métodos y procedimientos que tratan de estudiar mediante dispositivos móviles las experiencias de las personas (emociones, pensamientos, síntomas psicológicos, conductas) en su entorno natural y en la vida diaria. Al realizar mediciones varias veces al día para captar suficientemente la variabilidad temporal de los fenómenos, esta metodología permite evaluar los constructos psicológicos desde una perspectiva más dinámica, personalizada, contextual, micro-longitudinal y ecológica. Esta aproximación complementa los procedimientos tradicionales de evaluación psicométrica de papel y lápiz (Fon-

seca-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 et al., 2017). También están comenzando a emplearse las redes sociales para realizar evaluaciones. Así, por ejemplo, Kosinski et al. (2013) analizaron con éxito la posibilidad de utilizar los “me gusta” de Facebook como predictores de distintas características humanas, entre ellas los rasgos de la personalidad. Queda pendiente realizar estudios comparativos sistemáticos que permitan determinar el valor añadido de utilizar la información de las redes sociales, y si pueden considerarse sustitutos o, probablemente, complementos de los cuestionarios y test tal como los conocemos ahora (Muñiz y Fonseca-Pedrero, 2019). Sin negar la utilidad y potencial de las nuevas formas de evaluación, creemos que los test psicométricos, informatizados o no, seguirán siendo herramientas fundamentales, dada su objetividad y economía (Phelps, 2005, 2008). De hecho, estos nuevos enfoques y tecnologías seguramente serán una práctica común en un futuro no muy lejano, siempre y cuando muestren suficiente rigor psicométrico en términos de fiabilidad y validez y, con ello, se garantice que los instrumentos de medida evalúan de forma objetiva, equitativa y rigurosa. Esta es la prueba del algodón (Muñiz y Fonseca-Pedrero, 2019; Muñiz et al., 2020). Si no se supera la prueba, los avances tecnológicos no dejarán de ser meros fuegos de artificio, vistosos envoltorios, pero sin ningún valor psicométrico demostrado. En este sentido, ya hay ejemplos en el mercado de algunas pruebas sofisticadas informáticamente, con buenas campañas de marketing, que prometen evaluar constructos psicológicos y educativos tan complejos como el TDAH o la dislexia de forma simplista. Queda por demostrar que realmente midan los constructos pretendidos, ya que no tienen detrás ninguna evidencia empírica publicada que los avale. Eso sí, generan gráficos muy aparentes con colores y semáforos varios, y se lanzan incluso a diagnosticar y a ofrecer programas de formación que deben resultarles muy lucrativos. Por el bien de los usuarios y de la profesión debemos contribuir entre todos a desenmascarar los posibles 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 no vale con decir que los estudios realizados avalan su validez, es necesario presentar una información detallada de los estudios y sus resultados, para que el profesional pueda tomar una decisión informada sobre la calidad de los instrumentos. Usar un test, por sofisticado que sea, sin contar con garantías de fiabilidad y validez, es un disparate (Muñiz et al., 2020). A modo de conclusión general puede decirse que, tanto en España

como en Latinoamérica, los profesionales de la psicología son conscientes de la importancia de los datos provenientes de los test para aportar evidencias empíricas que les ayuden a la hora de tomar decisiones. Estos datos deben ser objetivos, para lo cual es fundamental que las puntuaciones de los test muestren las propiedades psicométricas adecuadas, que los profesionales tengan una formación rigurosa en Psicometría y Evaluación Psicológica, y que los instrumentos de medida se usen adecuadamente. En la mayoría de los países, unos más que otros, se están llevando a cabo acciones y proyectos encaminados a mejorar la práctica de los test. Queda mucho por hacer, pero como reza el proverbio hindú, las grandes caminatas empiezan por un pequeño paso. Estamos en ello.

## Referencias

- Alarcón, R. (1999). La psicología en América Latina en el siglo XX: Un análisis de sus características. *Revista de Psicología de la PUCP*, 17 (2), 131-146.
- Alarcón, R. (2004). Medio siglo de Psicología Latinoamericana. Una visión de conjunto. *Revista Interamericana de Psicología*, 38(2), 307-316.
- American Educational Research Association, American Psychological Association y National Council on Measurement in Education [AERA, APA & NCME] (2014). *Standards for educational and psychological testing*. American Psychological Association.
- Ardila, R. (1998). *La psicología en América latina: pasado, presente y futuro*. (2ª. Edición). Siglo XXI Editores.
- Ardila, R. (2004). La Psicología Latinoamericana: El Primer Medio Siglo. *Revista Interamericana de Psicología*, 38(2), 317-322.
- 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, 153-160.
- Bartram, D. (1996). Test qualifications and test use in the UK: The competence approach. *European Journal of Psychological Assessment*, 12, 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, 249-260.
- Benito, E. (2009). La formación en psicología: revisión y perspectivas. *Revista Psiencia*, 1, 3-10.
- Bennett, R. E. (1999). Using new technology to improve assessment. *Educational Measurement: Issues and Practice*, 18, 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*. 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*. John Wiley and Sons (pp. 219-251).

- Carlson, J. F. y Geisinger, K. F. (2012). Test reviewing at the Buros Center for Testing. *International Journal of Testing*, 12, 122-135.
- Chernyshenko, O. S. y Stark, S. (2016). Mobile psychological assessment. En F. Drasgow (Ed.) (2016). *Technology and testing*. Routledge. (pp. 206-216).
- Downing, S. M. y Haladyna, T. M. (Eds.) (2006). *Handbook of test development*. LEA.
- Drasgow, F. (Ed.) (2016). *Technology and testing*. Routledge.
- Drasgow, F., Luecht, R. M. y Bennett, R. E. (2006). Technology and testing. En R. L. Brennan (Ed.), *Educational measurement*. ACE/Praeger. (pp. 471-515).
- Elosua, P. (2017). Avances, proyectos y retos internacionales ligados al uso de tests en psicología. *Estudios de Psicología*, 34, 201-220.
- Elosua, P. (2019). *Los test psicológicos en la práctica profesional*. FOCAD. Consejo General de Psicología de España.
- Elosua, P. y Geisinger, K. F. (2016). Cuarta evaluación de test editados en España: forma y fondo. *Papeles del Psicólogo*, 37, 82-88.
- Elosua, P. y Muñiz, J. (2013). Proyectos españoles para una mejora en el uso de los tests. *Psiciencia. Latin American Journal of Psychological Science*, 6, 5-9.
- European Federation of Professional Psychologists' Associations [EFPA] (2005). *Meta-Code of ethics*. (www.efpa.eu).
- Evers, A., McCormick, C., Hawley, L., Muñiz, J. et al. (2017). Testing practices and attitudes toward tests and testing: An international survey. *International Journal of Testing*, 17, 158-190.
- Evers, A., Muñiz, J., Bartram, D., Boben, D., Egeland, J., Fernández-Hermida, J.R., y col. (2012). Testing practices in the 21st Century: Developments and European psychologists' opinions. *European Psychologist*, 17, 300-319.
- Evers, A., Muñiz, J., Hagemester, C., Høstmælingen, A., Lindley, P., Sjöberg, A., & Bartram, D. (2013). Assessing the quality of tests: Revision of the EFPA review model. *Psicothema*, 25, 283-291.
- Fernández-Ballesteros, R. (2014). *Evaluación psicológica*. 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, 187-200.
- Fonseca-Pedrero, E. (2017). Análisis de redes: ¿una nueva forma de comprender la psicopatología? *Revista de Psiquiatría y Salud Mental*, 10, 183-224.
- Fonseca-Pedrero, E. (2018). Análisis de redes en psicología. *Papeles del Psicólogo*, 39, 1-12.
- Fonseca-Pedrero, E. y Muñiz, J. (2017). Quinta evaluación de test editados en España: mirando hacia atrás, construyendo el futuro. *Papeles del Psicólogo*, 38, 161-166.
- Gallegos, M. (2016). Historia de la psicología y formación en psicología en América latina: convergencias temáticas. *Enseñanza e Investigación en Psicología*, 21, 319-335.
- Gallegos, M., Salas, G., Ardila, R., Caycho-Rodriguez, T., y Burgos, C. (2018). Sociedad Interamericana de Psicología: historia y proyección estratégica. *Revista Mexicana de Psicología*, 35, 72-87.
- García, A., Acevedo-Triana, C. A., y López-López, W. (2014). Cooperación en las ciencias del comportamiento latinoamericanas: Una investigación documental. *Terapia psicológica*, 32, 165-174.
- Geisinger, K. F. (2012). Worldwide test reviewing at the beginning of the twenty-first century. *International Journal of Testing*, 12, 103-107.
- Geisinger, K. F. (2016). Test evaluation. En S. Lane, M. R. Raymond y T. M. Haladyna (eds.), *Handbook of test development*. Routledge.

- Gómez-Sánchez, L. E. (2019). Séptima evaluación de test editados en España. *Papeles del Psicólogo*, 40, 205-210.
- Groth-Marnat, G. y Wright, A. J. (2016). *Handbook of psychological assessment*. Wiley.
- Haladyna, T. M. y Rodríguez, M. C. (2013). *Developing and validating test items*. Routledge.
- Hambleton, R. K., Merenda, P. F. y Spielberger, C. D. (Eds.) (2005). *Adapting educational and psychological tests for cross-cultural assessment*. LEA.
- Hernández, A., Ponsoda, V., Muñiz, J., Prieto, G. y Elosua, P. (2016). Revisión del modelo para evaluar la calidad de los test utilizados en España. *Papeles del Psicólogo*, 37, 192-197.
- Hernández, A., Tomás, I., Ferreres, A. y Lloret, S. (2015). Tercera evaluación de test editados en España. *Papeles del Psicólogo*, 36, 1-8.
- 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, 21-30.
- Insel, T. R. (2017). Digital phenotyping: Technology for a new science of behavior. *JAMA*, 318, 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*. John Wiley & Sons Ltd.
- ITC Guidelines for Translating and Adapting Tests (Second Edition) (2017). *International Journal of Testing*, 17, 1-34, ([www.InTestCom.org](http://www.InTestCom.org))
- Klappenbach, H. A. (2015). La formación universitaria en psicología en Argentina: Perspectivas actuales y desafíos a la luz de la historia. *Universitas Psychologica*, 14, 937-960.
- Knapp, S. J., Gottlieb, M. C., Handelsman, M. M., & VandeCreek, L. D. (Eds.). (2012). *APA handbooks in psychology. APA handbook of ethics in psychology, Vol. 1. Moral foundations and common themes*. American Psychological Association
- Klappenbach, H. A., y Gallegos, M. (2018). The long road to the profession of psychologists: Current status, challenges and Perspectives. In R. Ardila (Ed.), *Psychology in Latin America. Current Status, Challenges and Perspectives* (pp. 27-51). Springer
- Koocher, G. y Kith-Spiegel, P. (2007). *Ethics in psychology*. 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 (2<sup>nd</sup> edition)*. 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, 71-88.
- León, R. (2009). La contribución de Reynaldo Alarcón al desarrollo y fortalecimiento de la psicometría en el Perú. *Revista de Investigación en Psicología*, 12, 239-250.
- Lindsay, G., Koene, C., Ovreeide, H., y Lang, F. (2008). *Ethics for European psychologists*. Hogrefe.
- López, W. L., Silva, L. M., García-Cepero, M. C., Bustamante, M.C.A., y Aguado, E. (2011). Retos para la colaboración nacional e internacional en la psicología latinoamericana: un análisis del sistema RedALyC, 2005-2007. *Estudios de Psicología*, 16, 17-22.
- Markovetz, A., Blaszkiewicz, K., Montag, C., Switala, C. y Schlaepfer, T. E. (2014). Psycho-Informatics: Big Data shaping modern psychometrics. *Medical Hypotheses*, 82, 405-411.
- Miller, G. (2012). The smartphone psychology manifesto. *Perspectives on Psychological Science*, 7, 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). TEA Ediciones.

- Muñiz, J. (2018). *Introducción a la psicometría*. Pirámide.
- Muñiz, J. y Bartram, D. (2007). Improving international test and testing. *European Psychologist*, 12, 206-219.
- 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, 108-121.
- Muñiz, J., y Fonseca, E. (2017). *Construcción de instrumentos de medida en psicología*. Madrid: FO-CAD. Consejo General de Psicología de España
- Muñiz, J. y Fonseca, E. (2019). Diez pasos para la construcción de un test. *Psicothema*, 31, 7-16.
- Muñiz, J., Elosua, P. y Hambleton, R.K. (2013). Directrices para la traducción y adaptación de los tests: segunda edición. *Psicothema*, 25, 151-157.
- Muñiz, J., Hernández, A., y Fernández-Hermida, J. R. (2020). La utilización de los test en España: el punto de vista de los psicólogos. *Papeles del Psicólogo*, 41, 1-15.
- 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, 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, 151-157.
- Muñiz, J., Fernández-Hermida, J. R., Fonseca, E., Campillo, A. y Peña, E. (2011). Evaluación de test editados en España. *Papeles del Psicólogo*, 32, 113-128.
- 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, 201-211.
- 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, 528-534.
- Papeles del Psicólogo (2009). *Número monográfico sobre Ética Profesional y Deontología*. Vol. 30, 182-254.
- Phelps, R. (Ed.) (2005). *Defending standardized testing*. LEA.
- Phelps, R. (Ed.) (2008). *Correcting fallacies about educational and psychological testing*. American Psychological Association.
- Ponsoda, V. y Hontagas, P. (2013). Segunda evaluación de test editados en España. *Papeles del Psicólogo*, 34, 82-90.
- Primi, R., y Nunes, C. H. S. (2010). Psychological Testing & Assessment in Brazil: Advances in Research and Federal Regulations. *Testing International*, 24, 4-5.
- Puche, R. (2003). *Elementos relevantes para pensar un "estado del arte de la psicología académica" en Colombia*. [http://www.ascofapsi.org.co/observatorio/documentos/Elem\\_relevantes\\_Arte\\_psyco\\_Academica.pdf](http://www.ascofapsi.org.co/observatorio/documentos/Elem_relevantes_Arte_psyco_Academica.pdf).
- Rodríguez-Jiménez, O. R. y Rosero-Burbano, R. (2014). La formación en psicometría en los programas de pregrado de psicología. *Diversitas: Perspectivas en Psicología*, 10, 165-174.
- Rodríguez-Jiménez, O. R., Rosero-Burbano, R. F., Botia Sanabria, M. L., y Duarte Mateus, L. H. (2011). Producción de Conocimiento en Psicometría en Instituciones de Educación Superior de Bogotá y Chía. *Revista Colombiana de Psicología*, 20, 9-25
- Saenz, I. (2016). Procesos integrativos. Los psicólogos y el Mercosur. *Revista de Educación y Desarrollo*, 38, 14-24.
- SATEPSI, Sistema de Avaliação de Testes Psicologicos (2020). Testes aprovados. <http://satepsi.cfp.org.br/>

- Schmeiser, C. B. y Welch, C. (2006). Test development. En R. L. Brennan (Ed.), *Educational measurement* (4th edition) (pp. 307-353). American Council on Education/Praeger.
- Schulmeyer, M. K. (2016). La formación en evaluación psicológica en carreras de psicología de Bolivia. *Revista Interamericana de Psicología*, 50, 288-300.
- Sireci, S. y Faulkner-Bond, M. (2014). Validity evidence based on test content. *Psicothema*, 26, 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*. Routledge.
- Vinet, E. V. y González, N. (2013). Desarrollos actuales y desafíos futuros en la evaluación psicológica en Chile. *Psiciencia. Revista Latinoamericana de Ciencia Psicológica*, 5, 134-138.
- Wechsler, S. (2013). El movimiento para el desarrollo de los test psicológicos en el Brasil. Una experiencia pionera. *Psiciencia. Revista Latinoamericana de Ciencia Psicológica*, 5, 130-133.
- Wechsler, S., Hutz, C. S., y Primi, R. (2019). O desenvolvimento da avaliação psicológica no Brasil: avanços históricos e desafios. *Avaliação Psicológica*, 18 (2), 121-128. Doi:
- Wechsler, S. M., Oakland, T., Leon, C., Vivas, E., Almeida, L., Franco, A., Pérez-Solís, M., y Contini, N. (2014). Test development and use in five Iberian Latin American Countries. *International Journal of Psychology*, 49 (4), 233-239.
- Wright, A. J. (2011). *Conducting psychological assessment: A guide for practitioners*. Wiley.
- Zanatta, E. y Yurén, E. (2012). La formación profesional del psicólogo en México: trayecto de la construcción de su identidad disciplinar. *Enseñanza e Investigación en Psicología*, 17, 151-170.

## Agradecimientos

Trabajo parcialmente financiado por el Ministerio de Ciencia, Innovación y Universidades de España. Referencias: PSI2017-85724-P; PID2019-103859RB-I00.