

El Empleo de Test por Psicólogos/as Chilenos/as: Un Inquietante Panorama

The Use of Test by Chilean Psychologists: A Disturbing Panorama

Eugenia V. Vinet, Marcela Rodríguez-Cancino, Ailin Sandoval Domínguez,
Paloma Rojas-Mora y José L. Saiz

Departamento de Psicología, Universidad de La Frontera

Ante la escasez de antecedentes, este estudio indagó, desde un enfoque principalmente cuantitativo, 4 aspectos del empleo de test por parte de los/as psicólogos/as chilenos/as: caracterización del uso, pruebas de uso más frecuente, calidad psicométrica, problemas percibidos y actitudes. Una muestra no probabilística intencionada de 381 psicólogos/as (71,7% mujeres, $M_{edad} = 41,35$ años, $DE = 11,36$) respondió un cuestionario en línea que abordó los aspectos previamente mencionados. Los datos fueron examinados mediante análisis estadísticos descriptivos e inferenciales. Complementariamente, se efectuó una revisión sistemática cualitativa de la literatura sobre fundamentos psicométricos (confiabilidad, validez y/o referentes normativos) de los test empleados habitualmente en Chile. El principal hallazgo reveló que 15 de los 20 test de uso más frecuente no cuentan con evidencia que respalde su calidad psicométrica en poblaciones chilenas. Otros resultados muestran que el empleo de test es una práctica extendida en el ejercicio profesional, orientada preferentemente a la evaluación clínica, de adolescentes y adultos, realizada a través de unos pocos test que se repiten en los distintos campos profesionales y con presencia mayoritaria de técnicas proyectivas. La discusión se centra en la inobservancia de exigencias básicas para una correcta utilización de test y sus eventuales consecuencias. Se plantean posibles explicaciones para este proceder profesional y se ofrecen propuestas para mejorar las buenas prácticas en el empleo de pruebas psicológicas, señalándose, entre otras, el fortalecimiento de la formación de los usuarios de test y la adopción de estrategias restrictivas y formativas para favorecer el uso correcto de los test en Chile.

Palabras clave: test, psicólogos/as, calidad psicométrica, buenas prácticas, Chile

Given the scarcity of antecedents, this study set out to investigate, from a mainly quantitative approach, four aspects of the use of tests by Chilean psychologists: usage characterization, the most frequently used tests, psychometric quality, and perceived problems and attitudes. A purposive sample of 381 psychologists (71.7% female, $M_{age} = 41.35$ years) answered an online survey that addressed the previously mentioned aspects. The data were examined by descriptive and inferential statistical analyzes. In addition, a qualitative systematic review of the literature on psychometric foundations (reliability, validity and/or normative references) of the usual employment tests in Chile was carried out. The main finding revealed that 15 of the 20 most frequently used tests do not have evidence to support their psychometric quality in Chilean populations. Other results show that the use of tests is a widespread practice in the psychology profession, oriented to clinical evaluation, focused on adolescents and adults, and carried out through a few tests that are repeated in the different professional fields, having a majority presence of projective techniques. The discussion focuses on the failure to observe basic requirements for a proper use of tests and its eventual consequences. Possible explanations for this professional proceeding arise and proposals are offered aimed at improving good practices in the use of psychological tests, pointing out, among others, the strengthening of the training of test users and the adoption of restrictive and formative strategies to favor the correct use of tests in Chile.

Keywords: tests, psychologists, psychometric quality, good practices, Chile

Eugenia V. Vinet  <https://orcid.org/0000-0002-2504-4179>

Marcela Rodríguez-Cancino  <https://orcid.org/0000-0003-2012-9701>

Ailin Sandoval Domínguez  <https://orcid.org/0000-0001-8174-560X>

Paloma Rojas-Mora  <https://orcid.org/0000-0002-1690-7489>

José L. Saiz  <https://orcid.org/0000-0002-7137-4646>

Esta investigación fue financiada por la Dirección de Investigación de la Universidad de La Frontera, Chile, según el proyecto DFP 18-0033.

La correspondencia relativa a este artículo debe ser dirigida a Eugenia V. Vinet, Departamento de Psicología, Universidad de La Frontera, Avda. Francisco Salazar 01145, Casilla 54-D, Temuco, Región de la Araucanía, Chile. Email: eugenia.vinet@ufrontera.cl

Los test psicológicos son una de las principales herramientas que utilizan los/as psicólogos/as en su ejercicio profesional (Fonseca-Pedrero & Muñiz, 2017; Muñiz et al., 2020). En general, mediante la presentación de estímulos estandarizados y el registro sistemático de las respuestas de la persona evaluada frente a estos estímulos, los test buscan obtener información que permita valorar conductas o inferir atributos psicológicos (habilidades, conocimientos, rasgos, estilos, intereses, conflictos, entre otros) y tomar subsecuentemente decisiones.

Aunque hay una larga historia de clasificaciones de test (e.g., Yela, 1987/1996), de especial interés para el presente estudio fue la distinción entre pruebas objetivas y pruebas proyectivas (Reynolds et al., 2021, capítulo 10), las cuales proceden de marcos de referencia distintos para explicar y entender los fenómenos psicológicos. Aunque hay otros elementos diferenciadores, las pruebas proyectivas surgen principalmente del modelo psicodinámico y son llamadas así porque buscan exteriorizar contenidos propios del mundo interno de la persona a través del proceso de proyección. Ellas se centran en la pesquisa de aspectos no conscientes de la personalidad y registran las respuestas abiertas y amplias del evaluado ante estímulos generalmente ambiguos o poco estructurados, sin que esas respuestas puedan ser calificadas como correctas o incorrectas. Por su parte, las pruebas objetivas surgen desde el modelo psicométrico y están centradas en la evaluación cognitiva e intelectual o en aspectos más conscientes de la personalidad. En ellas el evaluado debe seleccionar una respuesta desde un conjunto predefinido de opciones. En los test cognitivos o intelectuales la opción elegida es juzgada como correcta o incorrecta o, cuando solicitan una respuesta abierta, esta es graduada según su nivel de corrección. Estas diferencias entre pruebas proyectivas y objetivas según origen, estímulos y respuestas determinan el modo en que son concebidos sus respectivos indicadores de calidad psicométrica. El presente estudio atendió tres de estos indicadores, esto es, confiabilidad, validez y referentes normativos.

Según los Estándares para Pruebas Educativas y Psicológicas (American Education Research Association [AERA], American Psychological Association [APA] & National Council on Measurement in Education [NCME], 2014/2018), en los test objetivos la confiabilidad se refiere a la precisión o coherencia de los puntajes entre replicaciones de un test, sin aludir necesariamente a las variadas técnicas para estimar tal coherencia. Por su parte, la validez es entendida como el grado en que el desarrollo teórico y la evidencia empírica respaldan las interpretaciones de los puntajes de una prueba. La validez es fundamental al momento de elegir una prueba para un determinado propósito. La acumulación de evidencia de validez procedente desde diferentes fuentes (estructura, contenido, relaciones con otras variables, consecuencias y procesos de respuesta) proporciona la base científica para la interpretación de los puntajes de la prueba. Cumplidos estos dos requisitos, confiabilidad y validez, para derivar inferencias útiles sobre el desempeño de una persona, la prueba debe disponer de referentes normativos (generalmente, baremos o puntajes de corte) que posibilitan la comparación del puntaje individual obtenido, con la distribución de puntajes alcanzados en la población a la cual pertenece esa persona.

En las pruebas proyectivas estos indicadores de calidad tienen un tratamiento más laxo. Autores como Lilienfeld et al. (2000), Sabogal (2004) y Sneiderman (2011) coinciden en señalar que la calidad psicométrica de las pruebas proyectivas es cuestionable. Si se juzgan desde el modelo psicométrico, la casi totalidad de ellas sería excluida del arsenal de pruebas psicológicas, porque las respuestas no son cuantificables, no tienen estimaciones de confiabilidad y validez y generalmente no poseen datos normativos adecuados (Sabogal, 2004; Salvia & Ysseldyke, 2001). Desde un punto de vista epistémico, el objetivo principal de estos test es investigar el desarrollo de los procesos que transcurren en la subjetividad y el psiquismo de la persona, presentando una interpretación subjetiva del evaluador que desafía los principios científicos. En estas pruebas, existe consenso que la confiabilidad se refiere a precisión y que puede ser evaluada a través del acuerdo inter-jueces. Por su parte, la validez es concebida como congruencia y se establece comparando la ejecución de la persona en la prueba con otros datos de la propia persona, como su historia clínica, datos de entrevista y la opinión de jueces expertos. Dado que el interés principal de estas pruebas es la individualidad, los referentes normativos generalmente no existen, son pobres o engañosos, llevando a juicios o predicciones erróneas (Lilienfeld et al., 2000). Aunque el estatus científico de las técnicas proyectivas es aún controversial (ver, e.g., Hibbard, 2003), lo cierto es que ellas también deben contar con evidencia sistemática sobre su calidad psicométrica, incluso, definiendo tal calidad en sus propios términos epistémicos.

De acuerdo a Muñiz et al. (2015), además de satisfacer las exigencias de confiabilidad, validez y normas, las pruebas deberían cumplir con otras dos condiciones básicas que involucran a los/as usuarios de los test: hacer un uso apropiado de las pruebas y que sus administradores cuenten con una suficiente formación en dicha tarea. Sin embargo, como se examina a continuación, estas condiciones no siempre se cumplen.

Es frecuente que en naciones en vías de desarrollo se empleen pruebas que no alcanzan una calidad psicométrica aceptable (Oakland, 2009), falencia que, por lo expuesto previamente, suele acentuarse en los test proyectivos. Se ha observado, también, en diversos países que los/as psicólogos/as califican como insuficiente su formación de pregrado en el conocimiento y uso profesional de test (Elosua, 2012). Este déficit formativo, que ocurre particularmente en Latinoamérica (Wechsler et al., 2014) y en España (Muñiz et al., 2020), dificultaría una correcta utilización de los test y revelaría una precaria calificación para el empleo profesional de estos. Los propios profesionales suelen reconocer la ocurrencia de diversos problemas en el uso de test, como, por ejemplo, conocimientos desactualizados y empleo de pruebas sin apoyo psicométrico y/o en grupos carentes de baremos específicos (Elosua, 2017; Muñiz & Fernández-Hermida, 2010). Un apropiado uso de las pruebas psicológicas demanda una formación competente sobre la teoría que respalda la prueba, su proceso de construcción y sus características psicométricas y, además, una suficiente experticia en su administración e interpretación. Una adecuada formación y práctica permitirá elegir el test más propicio, entre distintas alternativas, en coherencia con el propósito de la evaluación y la persona evaluada (Muñiz et al., 2011).

Diversas iniciativas institucionales han sido desarrolladas internacionalmente para consolidar un marco general de buenas prácticas en el uso de test. La International Test Commission (ITC) es una agrupación de asociaciones psicológicas nacionales, comisiones de test, editores y otras organizaciones comprometidas con la promoción de políticas efectivas para el uso de pruebas y también con el desarrollo, la evaluación y el uso adecuado de los instrumentos educativos y psicológicos. La ITC ha formulado detalladas directrices o guías sobre el empleo de test y otros aspectos conexos, como traducción y adaptación, investigación, seguridad y control de calidad, evaluación a gran escala en poblaciones diversas y evaluación informatizada (ITC, 2001). Estas guías están disponibles, además, en su página web (ITC.). Por su parte, la AERA et al., mediante un trabajo conjunto de larga data, han desarrollado sucesivas versiones de los ya citados Estándares para Pruebas Educativas y Psicológicas. La versión actual, que data de 2014, tiene por objetivo proporcionar criterios para el desarrollo, la evaluación y la aplicación de pruebas y evaluar las propiedades psicométricas de las mismas (AERA et al., 2014/2018). Sin embargo, tal como admite Elosua (2012), estas orientaciones internacionales no han sido aún completamente incorporadas en el quehacer profesional cotidiano.

Antecedentes en Chile

Aunque no hay evidencia sistemática y directa sobre un actual incumplimiento en Chile de las exigencias básicas del correcto uso de test, existen indicios de tal incumplimiento. Basándose principalmente en una revisión documental, Vinet y González (2013) identificaron tres sucesos que habrían deteriorado la calidad de la evaluación psicológica en el país: la proliferación de los programas de pregrado, la desregulación del quehacer profesional y un estancamiento en la creación o adaptación de pruebas psicológicas.

La promulgación en 1981 de la ley general de universidades chilenas, aún vigente, generó, según Didier Pino (2014), una institucionalidad educativa desregulada que favoreció la aparición de universidades privadas polivalentes o multifuncionales que buscaban ofrecer la mayor cantidad de carreras, en muchos casos, con escasa formalización. En psicología, la acelerada proliferación de programas de pregrado (carreras) se refleja en la oferta actual de 156 programas (Subsecretaría de Educación Superior, 2020). El crecimiento sostenido de la matrícula en estos programas fue acompañado de un detrimento de la calidad de la formación profesional (Toro & Villegas, 2001), afectando la adquisición de competencias en psicometría y evaluación psicológica. La instauración de mecanismos de acreditación de programas por parte del Estado chileno no mejoró esta situación, ya que estos solo verifican la correspondencia entre el perfil de egreso del programa y la formación ofrecida, sin implicar necesariamente una certificación de calidad de tal formación para el desempeño profesional (Moyano Díaz & Ramos Alvarado, 2013). Pese a todo, dentro del Consorcio de Universidades Estatales de Chile [CUECh], los programas de pregrado en psicología acordaron un marco curricular común basado en competencias genéricas y profesionales necesarias para la formación psicológica de pregrado en el país, incluyendo competencias en psicometría y evaluación psicológica (Juliá Jorquera, 2013). Aunque este marco fue suscrito por las siete universidades del CUECh que, en aquel momento, tenían carreras de psicología acreditadas (Moyano Díaz & Ramos Alvarado, 2013), ninguna universidad privada lo ha suscrito. Así, el impacto de este marco común parece ser bastante limitado, pues el CUECh incluye a solo 10 de los 40 programas de formación de pregrado en psicología registrados por la Comisión Nacional de Acreditación (2020). Además, la implementación de este currículum quedó a la voluntad e iniciativa de cada carrera, no habiendo posteriormente una evaluación de seguimiento en aquellos programas que lo adoptaron.

El desempeño profesional, incluyendo el empleo de test, está normado por el Código de Ética Profesional del Colegio de Psicólogos de Chile (1999). Sin embargo, puesto que en 1981 el régimen militar derogó la obligatoriedad de la colegiatura y eliminó la tuición ética sobre los/las colegiados/as, este colegio profesional no posee atribuciones ni mecanismos legales para el control normativo del quehacer profesional (Didier, 2014), quedando la profesión en un estado de alta precariedad jurídico-legal (Winkler et al., 2007).

Vinet y González (2013) y Didier Pino (2014) plantean algunas consecuencias de la desregulación profesional y de la proliferación de programas de pregrado en el ámbito particular de la evaluación psicológica. Por una parte, la desregulación acarreó un descuido por el cumplimiento de los requisitos teóricos, metodológicos y éticos de las pruebas. Por otra parte, la baja calidad de la formación ha ido acompañada de una disminución de las iniciativas para crear o adaptar pruebas psicológicas, reduciendo, a su vez, la disponibilidad de test psicométricamente respaldados. Esto último ha llevado al empleo profesional de instrumentos, particularmente proyectivos, que no satisfacen los estándares psicométricos. Como contraparte a esta deficitaria situación, el incremento en el país de programas de formación de postgrado en psicología, ocurrido especialmente en la última década, ha parecido estimular una mayor consciencia de la necesidad de crear instrumentos propios o de examinar la pertinencia de test foráneos siguiendo estándares internacionales.

El Presente Estudio

Como se señaló previamente, no hay investigaciones, más allá de revisiones documentales, que hayan abordado el panorama actual de la utilización de test en Chile. Contribuyendo a resolver esta carencia, el presente estudio se propuso indagar diversos aspectos del uso de test en el desempeño profesional de los/as psicólogos/as chilenos/as. Sus objetivos fueron cuatro: (a) caracterizar el empleo de test, (b) identificar los test más frecuentemente usados y sus variaciones según campo profesional, (c) revisar la calidad psicométrica (confiabilidad, validez y referentes normativos) de estos test y (d) examinar tanto los problemas percibidos como las actitudes respecto al empleo de test. Estos objetivos fueron cumplidos con datos primarios de una encuesta, excepto el objetivo 3, que fue alcanzado mediante una revisión sistemática de la literatura derivada de resultados vinculados al segundo objetivo.

Este estudio forma parte del proyecto denominado "Uso de los test psicológicos en los países iberoamericanos", desarrollado por el Grupo de Trabajo en Medición y Evaluación [GT-ME] de la Sociedad Interamericana de Psicología [SIP] (2020). grupo de reciente creación, que incluye a Argentina, Bolivia, Brasil, Chile, Colombia, España, Guatemala, Paraguay, Perú, Portugal, República Dominicana y Venezuela. Este proyecto cuenta con el apoyo de la Federación Iberoamericana de Asociaciones de Psicología, la ITC, el Consejo General de la Psicología de España y los colegios profesionales de los países participantes (Muñiz et al., 2021).

Método

Diseño

Se utilizó un diseño exploratorio, cuantitativo, descriptivo-correlacional y transversal. Este diseño fue complementado con una revisión sistemática de la literatura, de índole cualitativa, orientada a indagar la calidad psicométrica de los test de mayor uso.

Participantes

La población de este estudio estuvo integrada por las personas chilenas que poseen el título profesional de psicólogo/a. Aunque su cuantía es imprecisa, la Superintendencia de Salud (2019) registra aproximadamente a 60.000 profesionales. Focalizada en esta población, el estudio se difundió a través del Colegio de Psicólogos de Chile, organizaciones científicas (Sociedad Chilena de Psicología Clínica, Sociedad Científica de Psicología de Chile) y programas de pregrado en psicología de universidades chilenas. Estas instituciones invitaron a sus asociado/as o titulado/as a participar, proporcionándoles el enlace para acceder al instrumento en línea. Ingresaron a la página Web del cuestionario 707 personas; desde este grupo se extrajo una muestra intencionada de 381 personas que cumplieron cuatro criterios de inclusión: (a) estar ejerciendo profesionalmente la psicología, (b) ser chileno/a, (c) firmar un consentimiento informado y (d) responder al menos las dos primeras secciones, de un total de tres, del cuestionario empleado para recolectar los datos (ver subsección Instrumento). De los 381 participantes, 353 respondieron también la tercera sección

del cuestionario. Entre quienes no cumplieron los criterios de inclusión ($n = 326$), cinco eran extranjeros, 105 solo abrieron el cuestionario, pero no ingresaron datos, 62 omitieron parte de sus datos personales (primera sección) y no continuaron el cuestionario y 154 completaron sus datos personales, pero no abordaron la segunda y tercera parte del cuestionario. Cabe señalar que las 154 personas de este último subgrupo no difirieron significativamente de la muestra de estudio en cuanto a las variables sociodemográficas: sexo: $\chi^2(1, N = 535) = 0,591, p = 0,442$; , edad: $t(533) = -0,378, p = 0,705$ y región de procedencia: $\chi^2(14, N = 513) = 10,200, p = 0,747$. Tampoco hubo diferencias significativas en las siguientes variables profesionales: años de ejercicio: $t(533) = -0,058, p = 0,155$; nivel de formación profesional y/o académica: $\chi^2(3, N = 529) = 4,567, p = 0,206$ y campo de ejercicio profesional principal: $\chi^2(7, N = 535) = 8,765, p = 0,270$.

En la Tabla 1 se proporciona una descripción detallada de la muestra. Los/as 381 participantes reportaron una edad entre 23 y 86 años ($M = 41,35, DE = 11,36$) y fueron mayoritariamente mujeres. Aunque hubo participantes residentes en cada región del país, se concentraron en la Región Metropolitana (50,7%), La Araucanía (17,5%) y Valparaíso (5,5%), estando las dos primeras regiones sobre representadas y la última subrepresentada ya que según los datos censales la Región Metropolitana y Valparaíso son las más pobladas y La Araucanía ocupa el quinto lugar (Instituto Nacional de Estadísticas, 2018). Los/as participantes informaron entre 1 y 60 años de ejercicio profesional de la psicología ($M = 14,17$ años, $DE = 9,98$).

Tabla 1
Caracterización de la Muestra (n = 381)

Variable	f	%
Sexo		
Mujer	273	71,7
Hombre	108	28,3
Edad		
Hasta 29 años	56	14,7
30-39 años	141	37,0
40-49 años	99	26,0
50-59 años	50	13,1
60-69 años	32	8,4
70 años y más	3	0,8
Formación profesional y/o académica		
Título de psicólogo/a	82	21,5
Especialización	111	29,1
Magister	150	39,4
Doctorado	34	8,9
No responde	4	1,1
Años de ejercicio profesional		
Hasta 5 años	73	19,2
6-10 años	93	24,4
11-20 años	137	36,0
21-30 años	54	14,2
31 años y más	24	6,2
Campo de ejercicio profesional principal		
Clínica y de la salud	177	46,5
Enseñanza superior e investigación	54	14,2
Escolar y educación	51	13,4
Trabajo y organizaciones	47	12,3
Jurídico y forense	24	6,3
Social comunitaria	17	4,5
Neuropsicología	8	2,1
Actividad física y deportes	3	0,7

La mayoría de la muestra (68,5%) había cursado alguna especialización y/u obtenido el grado de magister luego de la titulación profesional. El ejercicio profesional principal se concentró en cuatro campos: psicología clínica y de la salud, enseñanza superior e investigación, psicología escolar y educación y psicología del trabajo y las organizaciones. Se destaca que, aunque no estaba incluida como una opción en el cuestionario, la neuropsicología emergió como campo principal mencionado espontáneamente. Un 72,7% de la muestra ($f = 277$) también reportó un campo laboral secundario, predominando la psicología clínica y de la salud (33,2%), psicología escolar y educación (18,1%) y enseñanza superior e investigación (14,8%). Un grupo menor de participantes (27,8%, $f = 106$) reportó un campo laboral terciario, siendo los más mencionados educación superior e investigación (21,7%), psicología clínica y de la salud (17,0%) y psicología del trabajo y las organizaciones (17,0%).

Instrumento

Se empleó un cuestionario *online* elaborado por el GT-ME de la SIP para el estudio iberoamericano sobre uso de test psicológicos, antes referido. Este cuestionario, generado a partir de estudios europeos que procuraban el mismo objetivo (ver Muñiz et al., 2021; Muñiz & Fernández Hermida, 2000), incluyó tres secciones. La primera indagó datos personales: sexo, edad, región de residencia, mayor nivel de formación académico/profesional alcanzado (*título profesional o licenciatura, especialización, magister, doctorado*), años de ejercicio profesional y campo laboral principal, secundario y terciario (*clínica y de la salud, escolar y educación, del trabajo y las organizaciones, social y comunitaria, jurídica y forense, de la actividad física y del deporte, enseñanza superior e investigación, otras áreas*).

La segunda sección abordó el uso de test y se inició con la pregunta "En el ejercicio de su profesión ¿hace uso de test y/o cuestionarios psicológicos de forma habitual?" seguida de una respuesta dicotómica (*Sí/No*). Si la respuesta era negativa, se solicitó al/la respondedor/a indicar sus motivos según cuatro opciones cerradas (*falta de formación adecuada, falta de estudios de validación en el país, falta de recursos para adquirir test, falta de confianza en el uso de test*) y dos opciones abiertas. La encuesta continuó solicitando los objetivos del empleo de test mediante la pregunta "Al hacer uso de los test ¿cuál es la frecuencia, en términos porcentuales, de los siguientes objetivos?". Las opciones de respuesta fueron ocho: *screening/cribado/tamizaje, evaluación/diagnóstico, selección o reclutamiento de personal, orientación o asesoramiento, evaluación de programas e intervenciones, investigación, docencia y otro objetivo* (con respuesta abierta). A continuación, se indagaron los grupos etarios receptores de los test, mediante la pregunta "¿Qué porcentaje de los test que utiliza es administrado a los siguientes grupos de edad?". Las alternativas de respuesta fueron: *infantil (menor de 3 años), infantil (3-7 años), infantil (8-12 años), adolescentes y jóvenes (13-25 años), adultos (26-64 años), adultos mayores (65 o más años)*.

Al final de esta segunda sección, se indagó sobre los test más usados en la práctica profesional mediante la pregunta "Cite los test que utiliza con más frecuencia en el ejercicio de su profesión, indicando el nombre completo, el acrónimo o siglas utilizadas para su denominación". La persona participante pudo escribir hasta 10 test, comenzando por aquel más frecuentemente empleado y continuando según un orden decreciente de uso.

La tercera sección incluyó dos partes. En la primera se solicitó la opinión sobre la frecuencia de ocho problemas en el uso de test psicológicos en el país (e.g., Hacer fotocopias de los test). Las opciones de respuesta fueron desde 1 (*problema muy poco frecuente*) hasta 5 (*problema muy frecuente*). La segunda parte de esta sección indagó el grado de acuerdo o desacuerdo con 20 enunciados actitudinales sobre el empleo de los test (e.g., El uso de test psicológicos debe ser limitado solo a psicólogos cualificados). Las opciones de respuesta fueron desde 1 (*total desacuerdo*) hasta 5 (*total acuerdo*).

Procedimiento

Antes de decidir su participación, las personas interesadas debieron leer la invitación y aceptar un consentimiento informado que describían, entre otros, el propósito y relevancia del estudio, la posibilidad de suspender en cualquier momento su participación, garantías sobre un manejo confidencial y seguro de los datos y el correo electrónico de dos investigadoras, en caso de requerir información adicional. La participación fue voluntaria, anónima, no remunerada y sin restricción de tiempo. El cuestionario estuvo disponible para ser contestado en la plataforma Web del Colegio Oficial de Psicólogos [COP] de España entre el 17 de octubre de 2017 y el 5 de junio de 2018 y los/as interesados/as accedieron a ella mediante un enlace específico para Chile. El cuestionario fue respondido en alrededor de 15 minutos y las respuestas se compilaron inicialmente en un archivo de datos Excel.

Análisis de Datos

Datos de la Encuesta

Se efectuaron análisis descriptivos, incluyendo frecuencias, porcentajes, promedios y desviaciones estándar. En las dos preguntas sobre objetivos y grupos etarios receptores de los test, que se solicitaron contestar en términos porcentuales, se calculó el porcentaje medio en cada opción de respuesta. Para identificar los test más empleados, se comenzó computando la frecuencia con que cada test fue reportado por la muestra en los tres primeros lugares de uso profesional, en el listado de hasta 10 menciones. Luego, los 20 test que presentaban los más altos porcentajes de menciones fueron operacionalmente considerados como instrumentos de uso frecuente. Se adoptó esta definición por dos razones. A partir del cuarto lugar del listado, las menciones de test disminuyeron notoriamente y comenzaron a aparecer test con una única mención. Además, este conjunto de 20 test incluyó a aquellos usados frecuentemente por al menos el 2% de la muestra, porcentaje comparable a los obtenidos en estudios europeos similares (Muñiz et al., 2001; Muñiz et al., 2020; Muñiz & Fernández-Hermida, 2010).

A fin de reducir datos y obtener resultados más parsimoniosos, mediante análisis factorial exploratorio (AFE), se indagaron las dimensiones subyacentes, por una parte, a los ocho ítems sobre problemas en el uso de los test y, por otra, a los 20 ítems referidos a actitud hacia el empleo de test. Los AFEs se basaron en la matriz de correlaciones policóricas inter-ítems. Se verificaron los criterios de la adecuación de factorización a través de las pruebas de Kaiser-Meyer-Olkin (KMO) y de esfericidad de Bartlett. Se utilizó el análisis paralelo versión clásica (Horn, 1965) para establecer el número óptimo de factores a retener. Los factores fueron estimados y rotados, respectivamente, por los métodos mínimos cuadrados no ponderados y Promin. Para determinar la pertenencia factorial de los ítems, se estableció una carga mínima de 0,40, junto a su capacidad interpretativa dentro del factor. Luego, según la solución retenida, se computaron los puntajes factoriales mediante el promedio de las respuestas de cada participante a los ítems que integraron cada factor, habiendo invertido previamente las respuestas dadas a los ítems con cargas negativas. Finalmente, se calcularon los promedios y desviaciones estándar de estos puntajes factoriales en la muestra total, promedios que fueron contrastados con el punto central o neutro (3,0) de la escala de respuesta, mediante pruebas *t* de grupo único de Student y el cómputo del coeficiente *d* de Cohen como medida de tamaño del efecto.

En todos los análisis precedentes se empleó el programa SPSS, versión 21, excepto en los AFEs, que fueron efectuados con el software FACTOR, versión 10.10.01 (Ferrando & Lorenzo-Seva, 2017).

Datos de la Revisión Sistemática de la Literatura

La indagación de la calidad psicométrica de los 20 test más frecuentemente usados se inició con una búsqueda de artículos en las bases electrónicas WoS, Scopus y SciELO, complementadas por Google Scholar, incluyendo, además, las revistas chilenas no indexadas pertenecientes a la Asociación Chilena de Revistas Científicas de Psicología (Polanco-Carrasco et al., 2017). Las palabras clave que guiaron la búsqueda fueron el nombre completo de cada test, los nombres alternativos, acrónimo o sigla, cada una junto a la palabra Chile. El periodo de búsqueda de artículos abarcó desde octubre 2019 hasta abril 2020, en idiomas inglés y español. Los criterios de selección de los artículos empíricos fueron tres: (a) publicados durante los últimos 20 años (1999-2019), (b) realizados en muestras chilenas y (c) que informaran sobre confiabilidad, validez y/o referentes normativos (normas o puntos de corte) de alguno de los 20 test en esas muestras.

Resultados

Caracterización del Uso de Test

De los/as 381 participantes, 302 (79,3%) reportaron emplear test en su práctica profesional habitual. Los/as 79 profesionales que señalaron no usarlos habitualmente informaron variados motivos para tal proceder: desconfianza en el uso de test como apoyo profesional (39,2%), carencia de estudios de validación de los test en el país (19,0%), convicción de que los test son innecesarios (16,5%), falta formación adecuada en el empleo de test (13,9%), recursos económicos insuficientes para adquirir test (10,1%), orientación teórica incompatible (7,6%), tiempo excesivo que demanda su uso (2,5%) y/o disgusto inespecífico con los test (2,5%).

La evaluación/diagnóstico fue el objetivo del uso de test que obtuvo la frecuencia media más alta (48,7%). Luego se posicionaron la selección o reclutamiento de personal (12,7%), el *screening*/cribado/tamizaje (10,6%) y la investigación (9,6%). Los objetivos de orientación o asesoramiento, evaluación de programas e intervenciones, y docencia obtuvieron 6,2%, 5,7%, y 4,9%, respectivamente. Otros objetivos presentaron solo un 1,6%. Además, la muestra reportó el uso de test en todas las etapas del ciclo vital: un 1,2% los empleaba en infantes menores de 3 años, 10,6%, en infantes entre 3 y 7 años, 18,2%, en niños y niñas (8 a 12 años), 30,3%, en adolescentes y jóvenes (13 a 25 años), 35,4%, en adultos (26 a 64 años) y 4,3%, en adultos mayores (65 o más años).

Test Más Frecuentemente Usados

La muestra mencionó el uso de 267 test diferentes. Los 20 test más frecuentemente utilizados se presentan en la Tabla 2, detallando la cantidad de participantes que los mencionaron en los tres primeros lugares del listado. De aquí en adelante estos test serán denominados por sus acrónimos, los cuales aparecen definidos en la nota de la Tabla 2. Entre estos test, el Ro fue el más mencionado y el Dibujo del Árbol, el menos mencionado. Solo siete de estos test son pruebas objetivas (WISC, WAIS, 16 PF, MMPI, BDI, Edwards, MACI) y los 13 restantes son pruebas proyectivas, incluyendo estas últimas cinco pruebas gráficas (PBL, HTP, DF, DFH, Árbol).

Tabla 2

Psicólogos/as que Mencionaron Usar Test en los Tres Primeros Lugares de Frecuencia

Acrónimo del test	Tipo	Lugar 1 (<i>f</i>)	Lugar 2 (<i>f</i>)	Lugar 3 (<i>f</i>)	Suma de <i>f</i> de lugares 1, 2 y 3	%
Ro	P	76	17	18	111	29,13
WISC	O	61	17	9	87	22,84
PBL	P	29	28	21	78	20,47
CAT	P	11	23	20	54	14,17
Luscher	P	19	23	12	54	14,17
WAIS	O	9	33	8	50	13,12
Z	P	15	20	7	42	11,02
TRO	P	6	17	15	38	9,97
HTP	P	13	14	9	36	9,45
DF	P	9	13	8	30	7,87
DFH	P	8	7	7	22	5,77
Bender	P	2	5	14	21	5,51
16 PF	O	5	5	8	18	4,72
Desiderativo	P	2	5	7	14	3,68
MMPI	O	5	1	7	13	3,41
BDI	O	4	2	7	13	3,41
Edwards	O	3	4	4	11	2,89
TAT	P	1	4	5	10	2,63
MACI	O	5	2	3	10	2,63
Árbol	P	0	3	5	8	2,10

Nota. Ro: Psicodiagnóstico de Rorschach; WISC: Escala Wechsler de Inteligencia para Niños; PBL: Dibujo de la Persona Bajo la Lluvia; CAT: Test de Apercepción Infantil; Luscher: Test de Colores de Luscher; WAIS: Escala Wechsler de Inteligencia para Adultos; Z: Test de Zulliger; TRO: Test de Relaciones Objetivas; HTP: Dibujo de Casa, Árbol, Persona; DF: Dibujo de la Familia; DFH: Dibujo de la Figura Humana; Bender: Test Gestáltico Visomotor de Bender; 16PF: Cuestionario de 16 Factores de Personalidad de Cattell; Desiderativo: Test Desiderativo; MMPI: Inventario Multifásico de la Personalidad de Minnesota; BDI: Inventario de Depresión de Beck; Edwards: Inventario de Preferencia Personales; TAT: Test de Apercepción Temática; MACI: Inventario Clínico para Adolescentes de Millon; Árbol: Dibujo del Árbol. La segunda columna indica si el test es proyectivo (P) u objetivo (O). *f* = frecuencia. El porcentaje en la sexta columna indica los participantes que reportan usar un test en los lugares 1, 2, o 3, respecto al total de participantes (*n* = 381).

La Tabla 3 presenta los 20 test más frecuentemente usados, según el campo profesional principal, considerando solo las menciones hechas en primer lugar. Los test con más del 5% de menciones en cada campo aparecen destacados en negrita; por ejemplo, en psicología clínica y de la salud destacan Ro, PBLL y WISC y, en psicología escolar y educación, WISC y PBLL. Globalmente, solo nueve test destacados (seis técnicas proyectivas, esto es, Ro, PBLL, CAT, Luscher, Z y HTP, y tres pruebas objetivas, esto es, WISC, 16PF y MACI), son suficientes para dar cuenta de la evaluación psicológica en los cinco campos profesionales principales de la psicología en Chile.

Tabla 3
Psicólogos/as que Mencionaron Usar Test en el Primer Lugar de Frecuencia, según Campo de Ejercicio Profesional Principal

Acronimo del test	Campo de ejercicio profesional principal									
	Psicología clínica y de la salud (n = 171)		Enseñanza superior e investigación (n = 51)		Psicología escolar y educación (n = 50)		Psicología del trabajo y las organizaciones (n = 45)		Psicología jurídica y forense (n = 23)	
	f	%	f	%	f	%	f	%	f	%
Ro	52	30,4	10	19,6	2	4,0	7	15,6	3	13,0
WISC	17	19,9	8	15,7	31	62,0	1	2,2	1	4,3
PBLL	18	10,5	0	0,0	4	8,0	0	0,0	3	13,0
CAT	8	4,7	0	0,0	1	2,0	0	0,0	2	8,7
Luscher	5	2,9	0	0,0	0	0,0	13	28,9	0	0,0
WAIS	7	4,1	1	2,0	0	0,0	0	0,0	0	0,0
Z	2	1,2	1	2,0	0	0,0	10	22,2	0	0,0
TRO	6	3,5	0	0,0	0	0,0	0	0,0	0	0,0
HTP	6	3,5	0	0,0	1	2,0	0	0,0	2	8,7
DF	5	2,9	2	3,9	2	4,0	0	0,0	0	0,0
DFH	4	2,3	1	2,0	1	2,0	0	0,0	1	4,3
Bender	2	1,2	0	0,0	0	0,0	0	0,0	0	0,0
16PF	1	0,6	1	2,0	0	0,0	3	6,3	0	0,0
Desiderativo	2	1,2	0	0,0	0	0,0	0	0,0	0	0,0
MMPI	1	0,6	2	3,9	0	0,0	2	4,4	0	0,0
BDI	3	1,8	0	0,0	0	0,0	0	0,0	0	0,0
Edwards	0	0,0	0	0,0	0	0,0	2	4,4	1	4,3
TAT	1	0,6	0	0,0	0	0,0	0	0,0	0	0,0
MACI	0	0,0	0	0,0	0	0,0	0	0,0	5	21,7
Árbol	1	0,6	0	0,0	0	0,0	0	0,0	0	0,0

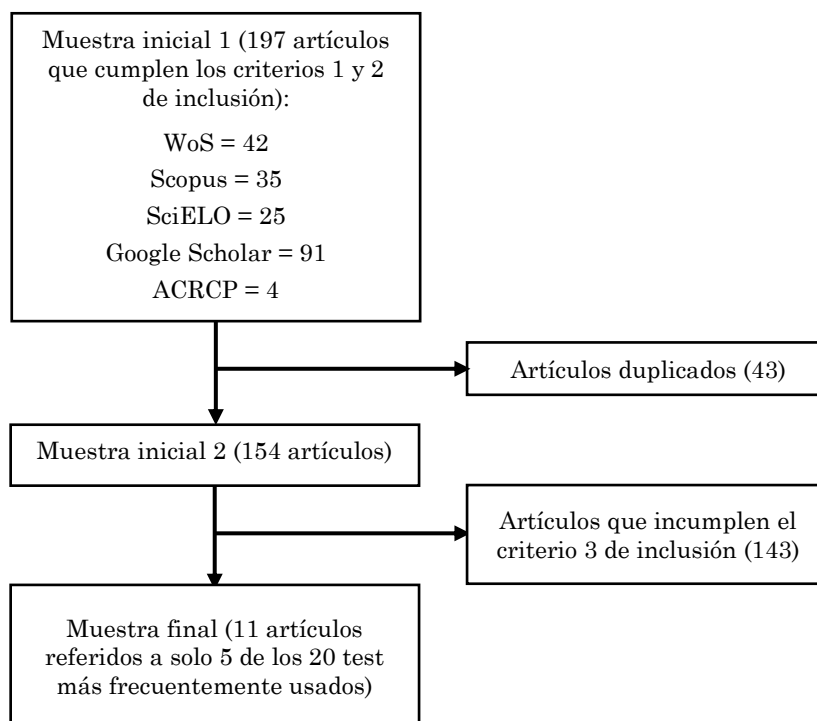
Nota. El nombre extenso de cada test aparece en la Tabla 2. Tres campos fueron excluidos (psicología social comunitaria, neuropsicología, y actividad física y deportes), debido a la escasa cantidad de psicólogos/as que laboraba en ellos (ver Tabla 1). Se destacan en negrita los porcentajes mayores a un 5% en cada campo profesional principal.

Calidad Psicométrica de los Test Más Frecuentemente Usados

La Figura 1 expone los resultados de cada etapa del proceso de la búsqueda de artículos según los criterios de inclusión, en la revisión sistemática de la literatura. Primeramente, se muestra el aporte de las diferentes bases de datos al conjunto inicial de 197 artículos que cumplen los criterios 1 (publicados entre 1999 y 2019) y 2 (realizados en muestras chilenas). De ellos se descartaron 43 artículos duplicados. De los 154 artículos conservados, solo 11 cumplieron con el criterio 3, esto es, reportar algún indicador de calidad psicométrica (confiabilidad, validez y/o referentes normativos) del test examinado.

Figura 1

Flujograma de la Selección de Artículos para la Revisión Sistemática de la Calidad Psicométrica de los 20 Test Más Frecuentemente Usados



Estos 11 artículos seleccionados se refieren solo a cinco de los 20 test más frecuentemente utilizados, a saber, WISC, WAIS, MMPI, BDI y MACI, todos de naturaleza objetiva. La Tabla 4 resume la información psicométrica contenida en esos artículos, respetando la terminología empleada por sus propios autores. Las pruebas de evaluación de la inteligencia WISC-III y WAIS-IV presentan evidencia de adecuada calidad psicométrica, según los tres indicadores analizados (confiabilidad, validez y referentes normativos). En la evaluación de la personalidad adolescente, el MACI también cumple estas tres exigencias. Igual situación se observa en la evaluación de la depresión, a través de dos versiones del BDI (BDI-II y BDI-IA), ofreciendo puntos de corte tanto para población adolescente como adulta. Finalmente, el MMPI-A, aunque ofrece adecuados indicadores de confiabilidad y validez, no reporta referentes normativos. No se encontró evidencia que sustentara la calidad psicométrica en Chile de los restantes 15 test frecuentemente empleados.

Frecuencia Percibida de Problemas en el Uso de Test

La capacidad de factorización de los ocho ítems sobre problemas en el uso de test fue respaldada por la prueba de esfericidad de Bartlett, $\chi^2(28, N = 333) = 1256,3, p = 0,001$ y el test KMO (0,852). El análisis paralelo versión clásica sugirió la presencia de un factor único. La solución unidimensional, con un autovalor de 4,0 y 50,1% de varianza explicada, se expone en la Tabla 5. Todas las cargas factoriales fueron iguales o mayores a 0,40. Mientras mayor es un puntaje en este factor, mayor es la percepción de una ocurrencia muy frecuente de este conjunto de problemas. Aunque los puntajes en este factor ($M = 3,29, DE = 0,92$) fueron significativamente mayores al punto central o neutro (3,0) de la escala, $t(352) = 6,01, p < 0,001$, el valor $d = 0,32$ reveló un tamaño de efecto entre pequeño y mediano (Cohen, 1988). Así, considerando que el promedio apenas sobrepasa el punto central de escala, puede decirse que los/as psicólogos/as perciben que este conjunto de problemas ocurre con una moderada frecuencia. La obtención de un factor único subyacente a la ocurrencia percibida de problemas con los tests ha sido también encontrada en España (Muñiz et al., 2020; Muñiz & Fernández-Hermida, 2010) y Latinoamérica (Muñiz et al., 2021).

Tabla 4
Evidencia de Calidad Psicométrica de Test Frecuentemente Usados en Chile

Test y publicación	Muestra	Confiabilidad	Validez	Referentes normativos
WISC-III				
Ramírez & Rosas (2007)	1924 NNA de 6 a 16 años	Consistencia interna	Validez de constructo	Puntaje escala y Puntaje compuesto (Coeficiente intelectual)
Pizarro et al. (2013)*	30 adolescentes de 16 años	No se reporta	Validez convergente con WAIS-IV	No se reporta
WAIS-IV				
Pizarro et al. (2013)*	30 adolescentes de 16 años	No se reporta	Validez convergente con WISC-III	No se reporta
Rosas et al. (2014)	887 adolescentes y adultos de 16 a 90 años	Consistencia interna	Validez de contenido, de constructo y convergente con Examen de Inteligencia Fluida (FIX)	Puntaje escala y Puntaje compuesto (Coeficiente intelectual)
MMPI-A				
Vinet & Alarcón (2003a)	705 adolescentes de 14 a 18 años	Consistencia interna y estabilidad	Validez de puntajes de escalas para discriminar entre grupos	No se reporta
BDI-II				
Melipillán Araneda et al. (2008)	737 adolescentes y jóvenes de 14 a 20 años	Consistencia interna y estabilidad	Capacidad para discriminar entre grupos Validez concurrente y convergente con escalas MACI Evaluación de estructura factorial	Puntaje de corte
BDI-IA				
Valdés et al. (2017)	1105 adultos de 18 a 73 años	Consistencia interna	Evaluación de estructura factorial	Puntaje de corte
MACI				
Vinet & Alarcón (2003b)	436 adolescentes y jóvenes de 13 a 18 años	Consistencia interna	Capacidad de discriminación entre grupos	No se reporta
Vinet & Forns (2006)	617 adolescentes y jóvenes de 13 a 18 años	No se reporta	Validez de puntajes de escalas para discriminar entre grupos	No se reporta
Vinet & Forns i Santacana (2008)	807 adolescentes y jóvenes de 13 a 19 años	No se reporta	No se reporta	Puntajes transformados con punto de corte
Vinet & Forns (2009)	807 adolescentes y jóvenes de 13 a 19 años	No se reporta	Validez concurrente con MMPI-A	No se reporta
Vinet et al. (2014)**	807 adolescentes y jóvenes 13 a 19 años	Consistencia interna	No se reporta	No se reporta

Nota. Se presentan solo los datos globales de las muestras utilizadas sin detallar submuestras. Se incluyen los indicadores de calidad psicométrica (confiabilidad, validez y referentes normativos) tal como son denominados por los autores de los estudios. NNA = Niños, niñas y adolescentes.

* Este artículo es mencionado dos veces, porque aporta evidencia de validez para los test WISC-III y WAIS-IV.

** Este artículo corresponde a un estudio de recapitulación y actualización de los trabajos anteriores y es el único que reporta la confiabilidad de las escalas del MACI correspondientes a la muestra normativa.

Tabla 5*Resultado del AFE que Explora la Frecuencia Percibida de Problemas en el Uso de Test*

Ítem	Carga
2. Hacer evaluaciones utilizando test inadecuados	0,79
8. Hacer interpretaciones que van más allá de los límites del test usado	0,75
7. Hacer caso omiso a las condiciones locales que puedan afectar la validez de los test	0,72
6. Uso de test por personas no calificadas	0,72
5. No tener en cuenta el error típico de medida en las puntuaciones de un test	0,72
3. No estar suficientemente al día con relación a los test usados	0,71
4. No tener oportunidades de contrastar mis evaluaciones con otros psicólogos	0,51
1. Hacer fotocopias de los test	0,40

Nota. Los ítems mantienen su numeración original.

Actitudes Hacia los Test

Dos ítems actitudinales fueron omitidos, ya que no diferenciaban adecuadamente: casi la totalidad de los/as participantes (94,3% en ítem 5: La interpretación de los resultados de los test y conclusiones debe ser realizada solo por psicólogos, y 89,2% en ítem 9: Los test psicológicos son útiles solo si se complementan con otras fuentes de información) reportó estar totalmente de acuerdo con dichas afirmaciones. Una serie de AFEs con los ítems restantes sugirió eliminar tres de ellos [ítems 12: Los profesionales disponen de suficiente información (revisiones independientes, investigaciones, documentación) sobre la calidad de los test editados en Chile, 19: La formación de grado o carrera debería de incluir cursos específicos sobre construcción y evaluación de test y cuestionarios, y 20: Si se utiliza adecuadamente Internet puede mejorar la calidad de la aplicación de los test], por presentar cargas factoriales inferiores a 0,40 en las soluciones obtenidas.

El AFE de los 15 ítems restantes informó que los datos eran aptos para ser explorados en su factorización, $\chi^2(105, N = 333) = 2186,4, p = 0,001$, según la prueba de esfericidad de Bartlett y $KMO = 0,78$. El análisis paralelo versión clásica sugirió una solución de cinco factores. Esta solución penta factorial oblicua, que explicó un 71,3% de la varianza, es plenamente interpretable y se expone en la Tabla 6.

Un revisor anónimo de este artículo indicó que la presencia de dos factores (F2 y F4) integrados por dos ítems podría responder a una sobrefactorización. AFEs complementarios de los datos mostraron que soluciones con una menor cantidad de factores generan una subfactorización y que, por tanto, la solución penta factorial retenida es adecuada. En la solución tetra factorial los originales F1 y F2 aparecieron fusionados, pese a que estos dos factores correlacionaron bajo (0,28) en el modelo penta factorial, el ítem 3 (originalmente en F1) no cargó en ningún factor y los ítems 1 y 2 continuaron conformando un factor con dos ítems. La solución trifactorial careció de una estructura simple puesto que los ítems 1 y 2 presentaron cargas cruzadas en dos factores, siendo dos de esas cuatro cargas mayores a 1 (casos Heywood). Estos resultados sugieren que los factores 1 y 2 reflejan dimensiones sustantivas y que la cantidad de ítems que contienen no es producto de un artefacto estadístico, sino más bien de la ausencia en el cuestionario de otros ítems que cubran estas dimensiones.

En la solución penta factorial, el primer factor (F1), denominado *Exclusividad*, incluye ítems referidos a la opinión de que el uso de test (ítem 3), su administración y corrección (ítem 4) e informe de resultados (ítem 6) deben ser competencias exclusivas de los/as profesionales psicólogos/as. El segundo factor (F2) fue rotulado *Restricción*, pues sus ítems (invertidos) reflejan la certeza de que es necesario controlar el desarrollo (ítem 14) y comercialización (ítem 15) de los test. El tercer factor (F3), denominado *Confianza*, contiene ítems que aluden a la opinión de que el empleo correcto de test (ítem 10) proporciona información confiable (ítem 17), útil (ítem 10) y de calidad (ítem 18). El cuarto factor (F4) fue llamado *Formación Inicial Limitada*, puesto que sus ítems (invertidos) señalan la convicción de que los conocimientos sobre test recibidos durante la formación de pregrado (ítem 1) han sido insuficientes para la práctica profesional de la psicología (ítem 2).

Tabla 6
Estructura Factorial de los 15 Ítems que Indagan Actitudes Hacia el Uso de Test

Ítem	F1	F2	F3	F4	F5
3. El uso de test psicológicos debe ser limitado solo a psicólogos cualificados	0,65	-0,15	0,05	0,10	0,22
4. Los test pueden administrarse y corregirse por personas que no son psicólogos	-0,50	-0,24	0,05	-0,00	0,20
6. Los resultados de los test solo pueden ser transmitidos por psicólogos	0,66	0,01	0,09	-0,03	-0,03
14. Los controles sobre los test deberían ser mínimos, pues inhiben el desarrollo de nuevas ideas y nuevos procedimientos de evaluación	-0,02	-0,61	0,02	-0,10	-0,13
15. Habría que permitir a los editores u otras entidades que comercialicen cualquier test que ellos consideren adecuado	-0,02	-0,61	0,01	0,08	-0,03
10. Utilizados correctamente los test son de gran ayuda para el psicólogo	-0,04	-0,05	0,71	-0,05	0,22
17. En el desempeño de mi profesión, el uso de test me ofrece confianza	0,04	-0,05	0,94	0,02	-0,06
18. Estoy satisfecho con la calidad de las pruebas e instrumentos que utilizo	0,09	0,09	0,74	0,06	-0,16
1. La formación recibida durante el grado o la carrera de psicología es suficiente para la práctica profesional de la psicología	-0,06	-0,02	0,16	-0,59	-0,09
2. Mis conocimientos actuales en relación con los test los recibí fundamentalmente durante la carrera de psicología	0,02	-0,03	-0,02	-0,86	0,01
7. Es necesaria una legislación oficial específica para el control del uso de test en mi país	0,17	-0,03	-0,14	0,02	0,81
8. El control del uso de los test por parte de asociaciones y órganos profesionales de psicólogos en mi país debe ser mayor	0,08	0,01	-0,15	0,04	0,89
11. Se debería establecer un sistema para acreditar la competencia de los usuarios de los test	-0,03	-0,07	0,06	0,04	0,81
13. Los estándares y directrices que definen las cualidades técnicas mínimas de un test deberían ser obligatorios (por ejemplo, los estándares de la ITC o los de la APA)	-0,04	0,10	0,09	-0,11	0,60
16. Las asociaciones profesionales deberían ejercer un papel más activo para regular y mejorar el uso que se hace de los test	-0,19	0,15	0,18	0,00	0,75
Autovalores	4,6	2,1	1,5	1,4	1,1
% de varianza explicada	30,8	13,9	10,0	9,2	7,4
<i>M</i>	4,36	3,96	3,97	3,15	4,02
<i>DE</i>	0,74	0,91	0,86	1,04	0,89
<i>t</i>	34,49	19,76	21,21	2,82	21,55
<i>p</i> ≤	0,001	0,001	0,001	0,005	0,001
<i>d</i>	1,84	1,05	1,13	0,15	1,15

Nota. F = factor. Los ítems mantienen su numeración original. Las cargas factoriales $\geq 0,40$ están destacadas en negrita. Las últimas cinco filas de la tabla presentan los resultados de la prueba *t* de una muestra, comparando el puntaje promedio de cada factor con el punto central o neutro (3,0) de la escala de respuesta.

Finalmente, los ítems del quinto factor (F5) se refieren a la certeza de que las organizaciones profesionales (ítem 8), respaldadas por una legislación oficial (ítem 7), deben ejercer una mayor regulación (ítem 16) y control en Chile (ítem 7) sobre el empleo de test y las competencias exigibles a sus usuarios (ítem 11), según estándares y directrices internacionales (ítem 13). Este factor fue denominado *Control Normativo Institucional*.

AFE realizados en España (Muñiz et al., 2020; Muñiz & Fernández-Hermida, 2010) identificaron dimensiones conceptualmente muy similares a los factores obtenidos en esta solución penta factorial: percepción de los test como una herramienta útil (similar a F3, *Confianza*), formación y conocimiento sobre

test (similar a F4, *Formación Inicial Limitada*) y necesidad de control y regulación de los test (similar a F5, *Control Normativo Institucional*, F1, *Exclusividad*, y F2, *Restricción*). Es importante destacar que esta similitud se obtiene a pesar de que esos estudios analizaron un conjunto de ítems que solo se superpone parcialmente con aquellos del presente estudio y emplearon otros métodos de factorización y rotación.

Como se exhibe en la sección inferior de la Tabla 6, los promedios de estos cinco factores fueron significativamente superiores al punto central (3,0) de la escala, con efectos de tamaño grande, excepto en F4, donde el efecto fue pequeño ($d = 0,15$). Así, se aprecia un alto acuerdo con el uso exclusivo de test por parte de la psicólogos/as (F1), la necesidad de restringir su desarrollo y comercialización (F2), el provecho profesional de un empleo correcto de los test (F3) y la urgencia de un control normativo formal del uso de test y de las aptitudes de sus usuarios/as (F5). En contraste, los y las participantes adoptaron una actitud cercana a neutra respecto a considerar suficiente su formación de pregrado en el uso de test (F4), probablemente reflejando la heterogénea calidad de la formación inicial en psicología en Chile (ver Vinet et al., 2015).

Discusión

Este estudio reveló que, entre los/as participantes, el empleo de test es una práctica frecuente y extendida en el ejercicio profesional. Este uso habitual está principalmente orientado a la evaluación clínica o diagnóstico, actividad que parece ser nuclear tanto en el rol percibido como desempeñado por psicólogos/as. El empleo de test se concentra en dos grupos etarios (adolescentes o jóvenes y adultos), con escasa presencia en los grupos extremos (infantes menores de 3 años y adultos mayores) reflejando, quizás, una escasa disponibilidad de test para estos grupos extremos y/o una exigua dedicación profesional en la atención de estas personas.

La muestra mencionó el uso habitual de algunos test que registran una larga historia en la literatura internacional sobre evaluación psicológica. Entre éstos aparecieron instrumentos multidimensionales como WISC y WAIS, en la evaluación intelectual, y Ro y MMPI, en la evaluación de personalidad y psicopatología. Sin embargo, surgieron también instrumentos, como el Luscher y el PBLL, que no aparecen mencionados, al menos, en revisiones europeas (e.g., Muñiz et al., 2001) ni iberoamericanas (e.g., Wechsler et al., 2014).

Los test de uso frecuente son mayoritariamente de naturaleza proyectiva. La inspección según los campos de ejercicio principal reveló un empleo concentrado en unos pocos test y repetido en distintos campos, con escasa inclusión de instrumentos especializados que aborden las particularidades de cada ámbito. Se constató, en cada campo, la presencia de técnicas proyectivas. En el campo clínico los test más utilizados fueron Ro, WISC y PBLL. En educación superior e investigación las menciones se centraron en Ro y WISC, instrumentos cuya administración es generalmente enseñada en los programas chilenos de pregrado en psicología, junto a otras técnicas proyectivas (Vinet et al., 2015). Llama la atención la enseñanza de Ro en la formación de pregrado, por cuanto se trata de una técnica de alta complejidad que requiere un entrenamiento individualizado y una madurez profesional que el estudiante de pregrado difícilmente posee. Ilustrando, al menos en Estados Unidos, el Ro es impartido mayoritariamente en programas doctorales de psicología clínica (Meyer et al., 2014); incluso, algunos autores (e.g., Evans & Finn, 2017) plantean que la evaluación psicológica, tanto proyectiva como objetiva, debería ser reconocida como una especialidad postdoctoral.

En el campo de la psicología escolar y educación, el test marcadamente más utilizado fue el WISC. Es probable que esta difundida práctica obedezca a la exigencia de uso del WISC hecha por el Ministerio de Educación (Fija normas para determinar los alumnos con necesidades educativas especiales, 2009) para las evaluaciones intelectuales que permiten el acceso de niños/as a los programas de apoyo escolar. En el campo de la psicología del trabajo y las organizaciones, la muestra informó el empleo preferente de pruebas proyectivas (Luscher, Z y Ro, en ese orden), concordando plenamente con Didier Pino (2014), quien, además, reporta que estas tres pruebas son las más demandadas a los profesionales de este campo. Cabe señalar que estos test son descritos por la Dirección Nacional del Servicio Civil (2008) como pruebas especializadas para evaluar la personalidad. El uso del Ro, una técnica eminentemente clínica, en este campo profesional es un ejemplo de administración en ámbitos para los cuales el test no fue originalmente concebido, práctica que carece, hasta ahora, de suficiente justificación. Finalmente, en psicología jurídica y forense se observó mayor diversidad en el uso de test, coexistiendo varias técnicas proyectivas (Ro, PBLL, CAT y HTP) con el MACI, un instrumento objetivo bastante estudiado y difundido en Chile que, a diferencia de los cuatro test anteriores, está específicamente orientado a la evaluación de diferentes grupos de adolescentes, especialmente de aquellos en situaciones de vulnerabilidad social (Vinet et al., 2014).

Quizás el hallazgo más relevante del presente estudio es la constatación de falta de evidencia sobre la calidad psicométrica de una gran porción de los 20 test más frecuentemente empleados en Chile. La revisión sistemática de la literatura de 154 artículos no redundantes identificó únicamente 11 trabajos que contenían información psicométrica favorable sobre solo cinco de estos 20 test, en particular sobre WISC, WAIS, MMPI, BDI y MACI. Los otros 15 test, mayoritariamente proyectivos, no contaron con respaldo psicométrico en Chile según las diversas fuentes consultadas. Este escenario es extremadamente preocupante, por cuanto revela el incumplimiento de exigencias básicas para un correcto uso de test (e.g., ITC, 2001, Muñiz et al., 2015), situación que, a su vez, genera legítimas dudas sobre la exactitud de las inferencias y decisiones que se hacen desde la información recogida por estos instrumentos (Elosua, 2017; Muñiz & Fernández-Hermida, 2010). Puesto que los test son comúnmente empleados para tomar decisiones sobre personas, este cuestionable proceder entraña reproches éticos esenciales. Decisiones carentes de respaldo psicométrico pueden perjudicar injusta y gravemente aspectos relevantes no solo de quienes son evaluados, sino también de las organizaciones y la comunidad en su conjunto (Carretero-Dios & Pérez, 2005; Fonseca-Pedrero & Muñiz, 2017).

Ilustrando, piénsese simplemente en el impacto negativo que inferencias infundadas podrían tener en la vida de personas que postulan a un trabajo o a un ascenso laboral. Este impacto es altamente probable, ya que, como reveló el presente estudio, en el ámbito de la psicología del trabajo y las organizaciones destaca el empleo habitual de test sin información psicométrica y, además, un segmento importante de los/as participantes del presente estudio reportó usar test para selección y reclutamiento de personal. La inexistencia de fundamento psicométrico en los test proyectivos empleados en este campo profesional ha sido también observada fuera de Chile. Específicamente, Didier Pino (2014), en una revisión de artículos tanto en inglés como español en reconocidas bases de datos, constató la ausencia de evidencia internacional sobre la calidad psicométrica de las pruebas proyectivas Luscher, Z y PBLI en el ámbito particular de la psicología laboral y de las organizaciones.

En la literatura internacional, la generalizada falta de evidencia sobre la fundamentación psicométrica de los test proyectivos ha llevado a discutir la pertinencia de su uso (Lilienfeld et al., 2000) o, al menos, su denominación como pruebas psicológicas (Sabogal, 2004). Pese a las dificultades para cuantificar las respuestas obtenidas en test proyectivos (Sneiderman, 2011), ha habido esfuerzos exitosos en la generación de respaldo psicométrico para este tipo de instrumentos; tal es el caso, por ejemplo, de algunos índices del Sistema Comprensivo del Ro (Exner Jr., 2003) examinados por Sahly et al. (2011) y Giromini et al. (2015). Otros ejemplos de técnicas proyectivas con evidencia de rigor psicométrico pueden ser encontrados en Brasil (Moraes Cardoso & da Silva-Filho, 2018).

¿Por qué los/as psicólogo/as emplean test carentes de respaldo psicométrico en Chile? Se discutirán cinco posibles razones, no necesariamente excluyentes. Primero, habría una limitada disponibilidad de pruebas psicométricamente adecuadas para poblaciones chilenas. Por una parte, cerca de un quinto de las personas encuestadas no usa frecuentemente test, debido a la falta de estudios de validación en el país. Por otra parte, existiría un muy bajo involucramiento de estos/as profesionales en investigación orientada al análisis de test pertinentes a la realidad nacional. Los autores de los 11 artículos identificados en la revisión de la literatura constituyen un número muy reducido y se concentran en solo cuatro núcleos investigativos: el Centro de Desarrollo de Tecnologías de Inclusión de la Escuela de Psicología de la Pontificia Universidad Católica de Chile, responsable actualmente de la estandarización nacional de las Escalas Wechsler de Inteligencia (WISC y WAIS), un equipo de la Universidad de Concepción y otro del Instituto Milenio para la Investigación en Depresión y Personalidad que examinaron dos versiones del BDI, y un equipo de la Universidad de La Frontera que estudió instrumentos de evaluación de la personalidad para adolescentes (MACI y MMPI-A). Segundo, la escasa disponibilidad de test estaría agravada por un aparente desconocimiento de los/as psicólogos/as sobre otros test que sí poseen respaldo psicométrico en Chile. Tal sería el caso, por ejemplo, del Inventario de Síntomas de Derogatis-Revisado (SCL-90-R), examinado en Chile por Gempp Fuentealba y Avendaño Bravo (2008), inventario multiescalas mencionado por solo tres encuestados/as en alguno de los 10 lugares provistos. Tercero, las instituciones que efectúan o contratan servicios de evaluación psicológica pueden sugerir o exigir, incluso sin argumentación explícita, el empleo de instrumentos que no cuentan con evidencia psicométrica nacional. Este podría ser el caso, por ejemplo, de la Dirección Nacional del Servicio Civil (2008), al incluir algunas pruebas proyectivas como instrumentos de selección de personal. Cuarto, la injustificable enseñanza de test sin respaldo psicométrico durante la formación de pregrado, particularmente de pruebas proyectivas (Vinet et al., 2015), podría contribuir a la mantención de este proceder profesional. Finalmente, la continuidad de esta práctica entre los/as psicólogos/as nacionales parece ser favorecida por la

inexistencia de un organismo regulador nacional y/o de fuentes de información, institucionalmente respaldadas, sobre las propiedades de los test usados en el país.

¿Están los/as psicólogos/as plenamente conscientes del incorrecto proceder al usar test desprovistos de soporte psicométrico? Los resultados sobre actitudes y percepción de problemas con los test no proporcionan una respuesta completamente afirmativa. En un plano actitudinal, junto con mostrarse muy favorables a que el desarrollo y uso de test sean actividades profesionales exclusivas, restringidas y controladas formalmente, los/as participantes reconocieron que el correcto empleo de los test proporciona valiosa información. Sin embargo, en un plano perceptivo, ellos/as no son tan unánimes respecto a la existencia frecuente de problemas ligados al empleo de test en el país, incluyendo aquellos referidos específicamente al uso de instrumentos inadecuados y otras deficiencias conexas, como desatender las condiciones locales, no considerar el error de medida, interpretar más allá de las posibilidades del test o no contrastar las interpretaciones con colegas o expertos. Habría una cierta paradoja: pese a valorar el uso correcto de test, los/as participantes utilizan ampliamente algunas pruebas sin fundamento psicométrico, al tiempo que perciben que esta errónea práctica no es muy habitual. Quizás, al sostener esto último, algunos/as profesionales atenúan la relevancia que tendría el surgimiento de reproches éticos.

Mirando Hacia Adelante

Se discuten a continuación cuatro propuestas destinadas a resguardar y promover las buenas prácticas en el empleo de test en Chile. Siguiendo a Prieto y Muñiz (2000), la primera propuesta corresponde al fortalecimiento de la formación inicial y avanzada de los usuarios de los test. En la formación de pregrado cabría impartir contenidos básicos de psicometría y desarrollar competencias generales de evaluación psicológica, dejando contenidos y competencias de nivel avanzado para una formación de especialización, postítulo o postgrado. En estos tres niveles de formación, la enseñanza debería incluir solo pruebas que cuenten con evidencia psicométrica en Chile, según estándares internacionales de calidad (e.g., Muñiz et al., 2015). Otro aspecto formativo crucial es la enseñanza de directrices internacionales sobre test, particularmente aquellas referidas a su correcto empleo (e.g., AERA et al., 2014/2018; ITC, 2001).

La segunda propuesta plantea la necesidad de estimular la creación y continuidad de nuevos equipos u organismos investigadores dedicados a generar fundamento psicométrico al empleo profesional de test. Para mejorar tal uso, es indispensable, previamente, "desarrollar pruebas eficientes" (Prieto & Muñiz, 2000, p. 65), incluyendo la construcción de nuevas pruebas, la correcta traducción y/o adaptación de otras desarrolladas en otros países/culturas y la revisión periódica de la calidad psicométrica de las ya disponibles. Además, sería conveniente promover el examen de las propiedades psicométricas y el desarrollo de referentes normativos en segmentos relevantes de la población nacional como, por ejemplo, aquellos definidos por la procedencia rural o urbana, el origen étnico, el nivel socioeconómico o la discapacidad.

La tercera propuesta se aboca a la falta de regulación formal del uso de test en Chile, debilidad claramente advertida por la muestra estudiada. Esto implicaría la adopción de una estrategia restrictiva destinada a limitar el uso de test solo a psicólogos/as cualificados/as (Muñiz et al., 2015) y, complementariamente, certificar oficialmente la calidad psicométrica de los test (Primi, 2010). Aunque parece remota la posibilidad de que alguna organización recupere la tuición legal de la profesión psicológica en Chile, la experiencia brasileña (Wechsler et al., 2019) puede ser inspiradora. El Consejo Federal de Psicología (CFP) de Brasil, organismo que orienta, regula y fiscaliza la psicología profesional en ese país, creó en 2001 el Sistema de Evaluación de Pruebas Psicológicas (SATEPSI; Moraes Cardoso & da Silva-Filho, 2018). Este sistema evalúa y certifica la calidad técnico-científica de los test según el cumplimiento de exigencias mínimas (fundamentación teórica, precisión, validez y normalización) para ser aprobados como instrumentos de empleo profesional por psicólogos/as. En términos prácticos, el SATEPSI genera dos categorías de test: aquellos que cumplen esas exigencias mínimas y, por tanto, se autoriza su uso profesional, y aquellos que reciben una evaluación desfavorable o aún no han sido evaluados y, por tanto, se impide su empleo. El CFP considera que el uso de pruebas incluidas en esta última categoría constituye un proceder reñido con la ética.

Finalmente, la cuarta propuesta consiste en producir información objetiva, experta, actualizada y periódica, de fácil acceso, que permita a los/as psicólogo/as seleccionar fundamentadamente los test a usar. Esta propuesta, a diferencia de la anterior, no pretende certificar sino más bien informar y, como tal, corresponde a la estrategia formativa (Muñiz & Fernández-Hermida, 2000). Un ejemplo de esta propuesta es la Comisión de Test, creada por el COP de España, destinada a potenciar el empleo adecuado de los test en ese país. Coordinada con organizaciones internacionales, esta comisión inició en 2010 la revisión sistemática de las

pruebas editadas en España, empleando el actualmente denominado Cuestionario de Evaluación de Test Revisado (CET-R; Hernández et al., 2016) para establecer la calidad de cada prueba y las limitaciones de su uso. Las evaluaciones, hechas mediante el CET-R por expertos independientes, son compendiadas en un informe final, con acceso público, que incluye las fortalezas y debilidades del test (Fonseca-Pedrero & Muñiz, 2017).

Comentarios Finales

Este estudio es el primero que aborda ampliamente las prácticas de los psicólogos/as chileno/as en el uso profesional de test. Sus resultados revelan, en lo principal, el empleo habitual de pruebas carentes de fundamentación empírica conocida, situación que cuestiona la calidad de la evaluación psicológica en Chile.

Pese a la relevancia de sus hallazgos, este estudio presenta al menos dos limitaciones. Por una parte, debido a la estrategia muestral no probabilística adoptada, resulta imposible determinar si la muestra es representativa de la población de psicólogos/as chileno/as. Se optó por esta estrategia, ya que no se dispone de registros fiables sobre esta población (Colegio de Psicólogos de Chile, comunicación personal vía email, 26 de agosto de 2020). Por otra parte, como el cuestionario de este estudio fue previamente consensuado entre los países participantes del proyecto iberoamericano, es posible que eventuales singularidades del empleo de test en Chile hayan quedado sin cubrir; por ejemplo, no se pudo contar con información sobre la organización específica a través de las cuales los/as profesionales accedieron a la invitación a participar.

Finalmente, es relevante destacar que este estudio, al menos en sus fases de formulación general y recolección de datos, se basó esencialmente en la acción coordinada de distintos profesionales iberoamericanos interesados en la evaluación psicológica. Concordando con Elosua (2017), los/as autores de este artículo resaltan la acción colaborativa internacional como estrategia para cautelar y fomentar las buenas prácticas en el uso de test.

Referencias

- American Educational Research Association, American Psychological Association & National Council on Measurement in Education. (2018). *Estándares para pruebas educativas y psicológicas* (M. Lieve, Trad.). American Educational Research Association. (Obra original publicada en 2014)
- Carretero-Dios, H. & Pérez, C. (2005). Normas para el desarrollo y revisión de estudios instrumentales. *Revista Internacional de Psicología Clínica y de la Salud*, 5 (3), 521-551. <https://www.redalyc.org/articulo.oa?id=33705307>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences: Second edition*. Lawrence Erlbaum.
- Colegio de Psicólogos de Chile. (1999). *Código de ética profesional*. http://colegiopsicologos.cl/web_cpc/wp-content/uploads/2014/10/CODIGO-DE-ETICA-PROFESIONAL-VIGENTE.pdf
- Comisión Nacional de Acreditación. (2020). *Búsqueda avanzada de acreditaciones*. <https://www.cnachile.cl/Paginas/buscador-avanzado.aspx>
- Didier Pino, N. O. (2014). Selección de personal en Chile. *Revista Iberoamericana de Psicología*, 7(1), 103-113. <https://doi.org/10.33881/2027-1786.rip.7109>
- Dirección Nacional del Servicio Civil. (2008). *Manual selección de personas en servicios públicos (2ª edición)*. Gobierno de Chile. MAVAL. <https://www.empleospublicos.cl/documentos/aspectosGenerales.pdf>
- Elosua, P. (2012). Tests publicados en España: Usos, costumbres y asignaturas pendientes. *Papeles del Psicólogo*, 33(1), 12-21. <http://www.papelesdelpsicologo.es/pdf/2031.pdf>
- Elosua, P. (2017). Avances, proyectos y retos internacionales ligados al uso de tests en psicología. *Estudios de Psicología (Campinas)*, 34(2), 201-210. <https://doi.org/10.1590/1982-02752017000200002>
- Evans, F. B. & Finn, S. E. (2017). Training and consultation in psychological assessment with professional psychologists: Suggestions for enhancing the profession and individual practices. *Journal of Personality Assessment*, 99(2), 175-185. <https://doi.org/10.1080/00223891.2016.1187156>
- Exner Jr., J. E. (2003). *The Rorschach: A comprehensive system. Volume 1: Basic foundations and principles of interpretation* (4th ed.). John Wiley & Sons.
- Ferrando, P. J. & Lorenzo-Seva, U. (2017). Program FACTOR at 10: Origins, development and future directions. *Psicothema*, 29(2), 236-240. <https://doi.org/10.7334/psicothema2016.304>
- Decreto 170 de 2009 [Ministerio de Educación de Chile]. Por el cual se establecen normas para determinar los alumnos con necesidades educativas especiales que serán beneficiarios de las subvenciones para educación especial. 21 de abril de 2010. https://especial.mineduc.cl/wp-content/uploads/sites/31/2018/06/DTO-170_21-ABR-2010.pdf
- Fonseca-Pedrero, E. & Muñiz, J. (2017). Quinta evaluación de tests editados en España: mirando hacia atrás, construyendo el futuro. *Papeles del Psicólogo*, 38(3), 161-168. <https://doi.org/10.23923/pap.psicol2017.2844>
- Gempp Fuentealba, R. & Avendaño Bravo, C. (2008). Datos normativos y propiedades psicométricas del SCL-90-R en estudiantes universitarios chilenos. *Terapia Psicológica*, 26(1), 39-58. <https://doi.org/10.4067/S0718-48082008000100004>
- Giromini, L., Viglione, D. J., Brusadelli, E., Lang, M., Reese, J. B. & Zennaro, A. (2015). Cross-cultural validation of the Rorschach Developmental Index. *Journal of Personality Assessment*, 97(4), 348-353. <https://doi.org/10.1080/00223891.2014.960927>
- Hernández, A., Ponsoda, V., Muñiz, J., Prieto, G. & Elosua, P. (2016). Revisión del modelo para evaluar la calidad de los tests utilizados en España. *Papeles del Psicólogo*, 37(3), 192-197. <http://www.papelesdelpsicologo.es/pdf/2775.pdf>

- Hibbard, S. (2003). A critique of Lilienfeld et al.'s (2000) "The scientific status of projective techniques". *Journal of Personality Assessment*, 80(3), 260-271. https://doi.org/10.1207/S15327752JPA8003_05
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179-185. <https://doi.org/10.1007/BF02289447>
- Instituto Nacional de Estadísticas (2018, Junio). *Síntesis resultados Censo 2017*. <http://www.censo2017.cl/descargas/home/sintesis-de-resultados-censo2017.pdf>
- International Test Commission. (2001). International guidelines for test use. *International Journal of Testing*, 1(2), 93-114. https://doi.org/10.1207/S15327574IJT0102_1
- International Test Commission. (9 de Agosto, 2020). *Guidelines*. <https://www.intestcom.org/page/28>
- Juliá Jorquera, M. T. (Ed.). (2013). *Competencias del psicólogo en Chile: propuestas desde las universidades estatales*. Editorial Universidad de La Serena.
- Lilienfeld, S. O., Word, J. M. & Garb, H. N. (2000). The scientific status of projective techniques. *Psychological Science in the Public Interest*, 1(2), 27-66. <https://doi.org/10.1111/1529-1006.002>
- Melipillán Aranedo, R., Cova Solar, F., Rincón González, P. & Valdivia Peralta, M. (2008). Propiedades psicométricas del Inventario de Depresión de Beck-II en adolescentes chilenos. *Terapia Psicológica*, 26(1), 56-59. <https://doi.org/10.4067/S0718-48082008000100005>
- Meyer, G. J., Viglione, D. J. & Giromini, L. (2014). An introduction to Rorschach-based performance assessment. En R. P. Archer & S. R. Smith (Eds.), *Personality assessment: Second edition* (pp. 301-369). Routledge.
- Moraes Cardoso, L. & da Silva-Filho, J. H. (2018). Satepsi e a qualidade técnica dos testes psicológicos no Brasil [Satepsi y la calidad técnica de las pruebas psicológicas en Brasil]. *Psicologia: Ciência e Profissão*, 38(Número spe), 40-49. <https://doi.org/10.1590/1982-3703000209112>
- Moyano Díaz, E. & Ramos Alvarado, N. (2013). Transformaciones en la formación de psicólogos en universidades del Estado de Chile. *Integración Académica en Psicología*, 1(2), 29-37. <https://integracion-academica.org/vol1numero2-2013/21-transformaciones-en-la-formacion-de-psicologos-en-universidades-del-estado-de-chile>
- Muñiz, J., Bartram, D., Evers, A., Boben, D., Matesic, K., Glabeke, K., Fernández-Hermida, J. R. & Zaal, J. N. (2001). Testing practices in European countries. *European Journal of Psychological Assessment*, 17(3), 201-211. <https://doi.org/10.1027/1015-5759.17.3.201>
- Muñiz, J. & Fernández Hermida, J. R. (2000). La utilización de los tests en España. *Papeles del Psicólogo*, 76, 41-49. <http://www.papelesdelpsicologo.es/resumen?pii=838>
- Muñiz, J. & Fernández-Hermida, J. R. (2010). La opinión de los psicólogos españoles sobre el uso de los tests. *Papeles del Psicólogo*, 31(1), 108-121. <http://www.papelesdelpsicologo.es/pdf/1801.pdf>
- Muñiz, J., Fernández-Hermida, J. R., Fonseca-Pedrero, E., Campillo-Álvarez, Á. & Peña-Suárez, E. (2011). Evaluación de tests editados en España. *Papeles del Psicólogo*, 32(2), 113-128. <http://www.papelesdelpsicologo.es/pdf/1947.pdf>
- Muñiz, J., Hernández, A. & Fernández-Hermida, J. R. (2020). Utilización de los test en España: el punto de vista de los psicólogos. *Papeles del Psicólogo*, 41(1), 1-15. <https://doi.org/10.23923/pap.psicol2020.2921>
- Muñiz, J., Hernández, A. & Ponsoda, V. (2015). Nuevas directrices sobre el uso de los tests: investigación, control de calidad y seguridad. *Papeles del Psicólogo*, 36(3), 161-173. <http://www.papelesdelpsicologo.es/pdf/2609.pdf>
- Muñiz, J., Hernández, A., Elosua, P., & Wechsler, S. (2021). El uso de los test en España y Latinoamérica: Perspectivas actuales y retos futuros. *Anuario Internacional de Revisiones en Psicología*, (1), 67-94. <https://doi.org/10.14635/REVPSY.0.4>
- Oakland, T. (2009). How universal are test development and use? En E. L. Grigorenko (Ed.), *Multicultural psychoeducational assessment* (pp. 1-40). Springer.
- Pizarro, M., Arroyo, R., Aparicio, A., Tenorio, M. & Rosas, R. (2013). *Evidencia de validez convergente entre las versiones chilenas de WAIS-IV y WISC-III* (Papeles de Investigación 1). Pontificia Universidad Católica de Chile, Escuela de Psicología, Centro de Desarrollo de Tecnologías de Inclusión. <http://descargas.cedeti.cl/2015/06/Papeles-No-1-%E2%80%9315-11-2013-Validez-convergente-WAIS-IV-y-WISC-IIIvch.pdf>
- Polanco-Carrasco, R., Gallegos, M., Salas, G. & López-López, W. (2017). Las revistas de psicología en Chile: historia y situación actual. *Terapia Psicológica*, 35(1), 81-93. <https://doi.org/10.4067/S0718-48082017000100008>
- Prieto, G. & Muñiz, J. (2000). Un modelo para evaluar la calidad de los test utilizados en España. *Papeles del Psicólogo*, 77, 65-71. <http://www.papelesdelpsicologo.es/resumen?pii=1102>
- Primi, R. (2010). Avaliação psicológica no Brasil: Fundamentos, situação atual e direções para o futuro [Evaluación psicológica en Brasil: Fundamentos, situación actual y directrices para el futuro]. *Psicologia: Teoria e Pesquisa*, 26(Número spe), 25-35. <https://doi.org/10.1590/S0102-37722010000500003>
- Ramírez, V. & Rosas, R. (2007). Estandarización del WISC-III en Chile: descripción del test, estructura factorial y consistencia interna de las escalas. *Psykhe*, 16(1), 91-109. <https://doi.org/10.4067/S0718-22282007000100008>
- Reynolds, C. R., Altmann, R. A. & Allen D. N. (2021). *Mastering modern psychological testing: Theory & methods*. Second edition. Springer.
- Rosas, R., Tenorio, M., Pizarro, M., Cumsille, P., Bosch, A., Arancibia, S., Carmona-Halty, M., Pérez-Salas, C. P., Pino, E., Vizcarra, B. & Zapata-Sepúlveda, P. (2014). Estandarización de la Escala Wechsler de Inteligencia para Adultos-Cuarta Edición en Chile. *Psykhe*, 23(1), Artículo 1. <https://doi.org/10.7764/psykhe.23.1.529>
- Sabogal, L. F. (2004). Pruebas proyectivas: acerca de su validez y confiabilidad. *Duazary: Revista de la Facultad de Ciencias de la Salud*, 1(2), 134-137. <https://doi.org/10.21676/2389783X.595>
- Sahly, J., Shaffer, T. W., Erdberg, P. & O'Toole, S. (2011). Rorschach intercoder reliability for protocol-level comprehensive system variables in an international sample. *Journal of Personality Assessment*, 93(6), 592-596. <https://doi.org/10.1080/00223891.2011.608761>
- Salvia, J. & Ysseldyke, J. E. (2001). *Assessment in special and remedial education* (8ª ed.). Houghton Mifflin.
- Sneiderman, S. (2011). Consideraciones acerca de la confiabilidad y validez en las técnicas proyectivas. *Subjetividad y Procesos Cognitivos*, 15(2), 93-110. <http://www.redalyc.org/articulo.oa?id=339630258005>
- Sociedad Interamericana de Psicología. (13 de agosto de 2020). *Grupos de trabajo*. <https://sipsych.org/gt/?lang=es>
- Subsecretaría de Educación Superior. (2020, 13 de agosto). *Dónde y qué estudiar. Buscador de carreras*. Gobierno de Chile, Ministerio de Educación. <https://www.mifuturo.cl/buscador-de-carreras/?tipo=carrera>
- Superintendencia de Salud. (12 de abril de 2019). *Estadísticas de prestadores individuales de salud. Periodo: 1 de abril de 2009 al 31 de marzo de 2019*. Gobierno de Chile, Ministerio de Salud. http://www.supersalud.gob.cl/documentacion/666/articulos-18219_recurso_1.pdf

- Toro, J. P. & Villegas, J. F. (2001). Chile. En J. P. Toro & J. F. Villegas (Eds.), *Problemas centrales para la formación académica y el entrenamiento profesional del psicólogo en Las Américas* (Vol. I, pp. 171-228). Sociedad Interamericana de Psicología. JVE Ediciones.
- Valdés, C., Morales-Reyes, I., Pérez, J. C., Medellín, A., Rojas, G. & Krause, M. (2017). Propiedades psicométricas del Inventario de Depresión de Beck IA para la población chilena. *Revista Médica de Chile*, 148(8), 1005-1012. <https://doi.org/10.4067/s0034-98872017000801005>
- Vinet, E. V. & Alarcón, P. (2003a). Evaluación psicométrica del Inventario Multifásico de Personalidad de Minnesota para Adolescentes (MMPI-A) en muestras chilenas. *Terapia Psicológica*, 21(2), 87-103. o <https://psycnet.apa.org/record/2004-12498-002>
- Vinet, E. V. & Alarcón, P. (2003b). El Inventario Clínico para Adolescentes de Millon (MACI) en la evaluación de adolescentes chilenos. *Psykhē*, 12(1), 39-55. <http://ojs.uc.cl/index.php/psykhe/article/view/20341>
- Vinet, E. V., Barrera-Herrera, A. & Salinas-Oñate, N. (2014). El MACI en Chile: desarrollo culturalmente pertinente de un test para adolescentes. *Revista Iberoamericana de Diagnóstico y Evaluación Psicológica*, 37(1), 123-145. <https://www.aidep.org/sites/default/files/articles/R37/Art6.pdf>
- Vinet, E. V. & Fornis i Santacana, M. (2006). El Inventario Clínico para Adolescentes de Millon (MACI) y su capacidad para discriminar entre población general y clínica. *Psykhē*, 15(2), 69-80. <https://doi.org/10.4067/S0718-22282006000200007>
- Vinet, E. V. & Fornis i Santacana, M. (2008). Normas chilenas para el MACI: una integración de criterios categoriales y dimensionales. *Terapia Psicológica*, 26(2), 151-163. <https://doi.org/10.4067/S0718-48082008000200001>
- Vinet, E. V. & Fornis i Santacana, M. (2009). Validación de los puntajes de corte del MACI a través de las escalas clínicas del MMPI-A. *Psykhē*, 18(1), 11-25. https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-22282009000100002
- Vinet, E. V. & González, N. (2013). Desarrollos actuales y desafíos futuros en la evaluación psicológica en Chile. *Psiencia: Revista Latinoamericana de Ciencia Psicológica*, 5(2), 134-138. <https://doi.org/10.5872/psiencia.5.2.74>
- Vinet, E. V., Hueche, C., Antivilo, A., González, I., Castro, P., González, M. & Ramos, N. (12-16 de julio de 2015). Diversidad en la formación en evaluación psicológica: el caso de las universidades chilenas. En S. Wechsler (Coord.), *La formación en evaluación psicológica en Latinoamérica: una tensión no resuelta entre teoría y práctica* [Simposio]. XXXV Congreso Interamericano de Psicología, Lima, Perú.
- Wechsler, S. M., Hutz, C. S. & Primi, R. (2019). O desenvolvimento da avaliação psicológica no Brasil: avanços históricos e desafios. [El desarrollo de la evaluación psicológica en Brasil: Avances históricos y retos] *Avaliação Psicológica*, 18(2), 121-128. <https://doi.org/10.15689/ap.2019.1802.15466.02>
- Wechsler, S. M., Oakland, T., León, C., Vivas, E., de Almeida, L., Franco, A., Pérez-Solís, M. & Contini, N. (2014). Test development and use in five Iberian Latin American countries. *International Journal of Psychology*, 49(4), 233-239. <https://doi.org/10.1002/ijop.12068>
- Winkler, M. I., Pasmanik, D., Alvear, K. & Reyes, M. I. (2007). Cuando el bienestar psicológico está en juego: la dimensión ética en la formación profesional de psicólogos y psicólogas en Chile. *Terapia Psicológica*, 25(1), 5-24. <https://doi.org/10.4067/S0718-48082007000100001>
- Yela, M. (1996). Los tests. Reproducido de Introducción a la teoría de los tests (pp. 3-29). Universidad Complutense de Madrid. *Psicothema*, 8(Suplemento 1), 249-263. <https://www.psicothema.com/pdf/660.pdf>

Fecha de recepción: Noviembre 2020.

Fecha de aceptación: Diciembre 2021.