



EVALUAR

Revista Evaluar

Laboratorio de Evaluación Psicológica y Educativa
Facultad de Psicología de la Universidad Nacional de Córdoba

2020

VOL 20 - Nº1
ISSN 1667-4545



Validación del Cuestionario de Signos observables asociados a la Inhibición Psicomotriz (SIP)

Validation of the Psychomotor Inhibition Signs Questionnaire (SIP)

Silvia Saal^{1*}, Josefina Reboursin¹, María Fernanda Molina¹

1 - Universidad Nacional de Tres de Febrero (UNTREF), Caseros, Buenos Aires, Argentina

Introducción
Método
Resultados
Discusión
Referencias

Recibido: 06/01/2020 Revisado: 11/03/2020 Aceptado: 03/04/2020

Resumen

El objetivo de este trabajo es presentar la validación de un instrumento de despistaje: el Cuestionario de Signos observables asociados a la Inhibición Psicomotriz (SIP; Saal, 2019). Participaron 675 niños de entre 5 y 10 años que asistían al primer ciclo de escuelas primarias del Área Metropolitana de Buenos Aires. El análisis factorial exploratorio reveló una estructura bidimensional: *inhibición* y *obediencia*. Ambas dimensiones presentan una buena consistencia interna: inhibición, $\alpha = .92$; obediencia, $\alpha = .80$. Si bien ambos factores son conceptualmente relevantes, el puntaje del primer factor se presentó como mejor medida para el rastillaje de la inhibición psicomotriz en niños de primer ciclo. Tuvo un punto de corte de 35.5 que ofrece la mejor combinación de sensibilidad (.78) y especificidad (.72). El área bajo la curva ROC fue de .79 ($p < .001$). En conclusión, el SIP presenta adecuados indicadores de validez y confiabilidad para su uso como instrumento de despistaje.

Palabras clave: instrumento de despistaje, validación, inhibición psicomotriz, trastorno psicomotor, detección temprana, niñez, escuela primaria

Summary

The aim of this article is to introduce the validation of a screening tool: the Psychomotor Inhibition Signs Questionnaire (SIP; Saal, 2019). A total of 675 children aged 5 to 10 years attending the three first years of elementary schools of the metropolitan area of Buenos Aires took part in the study. A bidimensional structure: *inhibition* and *obedience*, was obtained by exploratory factor analysis. Internal consistency was adequate for both dimensions: inhibition, $\alpha = .92$; obedience, $\alpha = .80$. Even if both factors are conceptually relevant, the first factor's score was found to be a better screening measure for psychomotor inhibition. Cut-off point was established at 35.5 where high levels of sensitivity (.78) and specificity (.72) were obtained. The area under the ROC curve was .79 ($p < .001$). In sum, these results suggest that this tool is valid and reliable for screening psychomotor inhibition.

Keywords: screening tool, validation, psychomotor inhibition, psychomotor disorder, early detection, childhood, elementary school

*Correspondencia a: Silvia Saal. Calle 12 de Octubre y Las Camelias, Office Quatro, 1^{er} piso, of. 116, Localidad del Viso, Partido de Pilar, Provincia de Buenos Aires. C.P.1669. E-mail: saalsilvia@gmail.com

Cómo citar este artículo: Saal, S., Reboursin, J., & Molina, M. F. (2020). Validación del Cuestionario de Signos observables asociados a la Inhibición Psicomotriz (SIP). *Revista Evaluar*, 20(1), 01-16. Recuperado de <https://revistas.unc.edu.ar/index.php/revaluar>

Introducción

El Cuestionario de Signos Observables Asociados a la Inhibición Psicomotriz (SIP; Saal, 2019) es un instrumento de despistaje que fue construido con el objetivo de detectar tempranamente signos de inhibición psicomotriz. La *inhibición psicomotriz* se encuadra dentro de los trastornos psicomotores que, si bien no guardan correlación con una lesión neurológica u orgánica (De Ajuriaguerra, 1984) implican, en la mayoría de los casos, una perturbación y un padecimiento que entorpece el cuerpo, el movimiento y la comunicación (Bucher, 1978). En la inhibición psicomotriz el movimiento se encuentra disminuido o prácticamente frenado. Esto deja vacante el valor de la experiencia corporal, esencial para el desarrollo en la infancia.

Diversos autores enfatizan que es mediante el cuerpo en movimiento que el niño puede entrar en relación, ponerse en contacto, sentir y expresar una determinada manera de ser y hacer con el cuerpo en el espacio, en el tiempo, con los objetos y con los otros (De León, 2012; Le Boulch, 1992). Se vuelve relevante considerar su papel en el desarrollo de la inteligencia, al ser los movimientos, junto con las percepciones, la base de la inteligencia senso-motora (Piaget & Inhelder, 1981). El movimiento es el instrumento del que dispone el niño para la acción concreta y la acción es el camino para el acceso al campo de la representación (Wallon, 1965).

Con frecuencia se observa que la merma del movimiento, su disminución o ausencia, no suelen llamar la atención y hasta pueden ser connotadas positivamente. Es habitual que en los distintos ámbitos en los que se desarrolla el niño (en su casa, en la escuela), la reducción del movimiento suela ser asociada por los adultos al *portarse bien*, al *ser buen alumno* y, por lo tanto, sea bien recibida, no sea cuestionada ni problematizada

(Saal, 2011).

Por otra parte, cuando el movimiento se presenta en exceso, cuando se evidencia una dificultad para restringirlo y se manifiesta de modo disruptivo ante el entorno, inevitablemente llama y provoca la mirada del otro. El exceso de movimiento ha sido ampliamente estudiado bajo distintas nominaciones, como “hiperactividad, hiperkinesia, hipermovilidad o inestabilidad psicomotriz” (Untoiglich, 2011, p. 50). La *hiperactividad*, generalmente pensada desde el discurso médico en asociación con la desatención relativa al trastorno por déficit de atención con hiperactividad (American Psychiatric Association, 2014) ha tomado gran protagonismo a partir de fines del siglo pasado, evidenciando en ocasiones un sobrediagnóstico y una creciente medicalización farmacológica (Merten, Cwik, Margraf, & Schneider, 2017; Rodríguez, 2011; Untoiglich, 2011). La *inestabilidad psicomotriz*, por su parte, se caracteriza por impulsividad, inquietud y desatención y suele ser el trastorno psicomotor más estudiado y que mayor casuística clínica presenta (Bergés, 1974, 1985; Boscaini, 2013). El amplio estudio de estas temáticas ha dado lugar a desarrollos teóricos y propuestas de intervención tanto en el campo educativo como clínico.

Si afirmamos que el exceso de movimiento debe ser atendido en tanto puede ser la manifestación de un sufrimiento infantil (Untoiglich, 2011), es imperioso interpelarnos acerca de qué les sucede a aquellos niños que se encuentran en el polo opuesto. Consideramos que es de fundamental importancia abordar la posibilidad de que haya un sufrimiento psíquico-corporal concomitante a las conductas que, en general, se tiende a naturalizar.

La reducción del movimiento y del accionar con el cuerpo es uno de los signos característicos de la inhibición psicomotriz. En la inhibición psicomotriz la reducción o retención del movimiento

y la lentitud en el accionar se presentan sobre un fondo hipertónico. La escasa experiencia corporal trae aparejados signos de torpeza psicomotriz: la posibilidad de ajustarse a los actos cotidianos se desarrolla de manera insuficiente y disarmónica. Se registran también rasgos de timidez, pasividad e hipercontrol (Defontaine, 1980). Cuando esta modalidad de funcionamiento psicomotor se cristaliza y se fija en el tiempo, conlleva malestar y un subyacente padecimiento.

Se considera importante remarcar, de todos modos, que la inhibición como función, es una propiedad operativa del sistema nervioso central que suprime determinadas descargas nerviosas para permitir el flujo de otras (Galimberti, como se citó en Wille, 2007). No se ejerce únicamente para bloquear la actividad, sino que suele presentarse para regular o moderar la impulsividad del gesto (Molina de Costallat, 1987), por lo que cumple un carácter adaptativo en la confrontación con el medio y en el ejercicio de las funciones motrices. Sin embargo, la inhibición se torna patológica cuando se da en exceso, deviniendo expresión de un sufrimiento latente (Benavides, 2003). La inhibición psicomotriz no se define, entonces, por la inhibición de la función exclusivamente, sino que implica la inhibición del funcionamiento y de la realización del cuerpo ligado al movimiento (Bergés, 1974, 1991).

La semiología que caracteriza a la inhibición psicomotriz ha sido descripta por Bucher (1978), quien destaca la voz baja al hablar, la presencia de vacilaciones, la evitación de la mirada del otro, la respiración bloqueada y la escasa expresividad facial. En el marco del diagnóstico, la autora menciona el rechazo por parte de los niños a varias pruebas. De León (2010) refiere que la tensión que frena los movimientos y la pasividad se presentan de modo persistente y atribuye esto a la intención de los niños de pasar desapercibidos.

La inhibición psicomotriz ha sido clasificada

por distintos autores en función de sus maneras de manifestarse. Bergés (1985) distingue la inhibición por hipercontrol y retención, de la inhibición por suspensión de la iniciativa. La primera se caracteriza por un aumento del tono, movimientos frenados y una censura de la actitud y el gesto que se agravan ante la presencia de otras personas. En el segundo caso, el autor ubica la perturbación antes de la acción, en el momento inicial del acto; contrariamente al caso anterior, la presencia de otros tiende a levantar la inhibición. Wille (2007) propone, dentro de la inhibición psicomotriz, una tipología a la que denomina *conflictiva*, caracterizada por un bloqueo temporal o permanente de la acción, la renuncia al movimiento y un sentimiento de peso de la mirada de los otros, pero conservando el interés hacia el mundo externo. La autora subdivide a esta tipología en dos modos de expresión: acción suspendida y acción tensa. Considera también otra tipología, a la que denomina *inercia* y la subdivide en las formas autística apática y ajetreada. La autora expresa que esta tipología suele estar asociada al retraso mental grave (Wille, 2012). Saal (2013), a partir de la observación de niños en Argentina, señala que la inhibición psicomotriz puede expresarse en modo vigilante, ligada a la retención del movimiento, cuando el cuerpo no encuentra un modo de evacuar las tensiones, y en modo vacilante, ligada a la reducción del movimiento, temor e indecisión que inhabilitan la acción.

Distintos aspectos del desarrollo del niño están comprometidos en este trastorno psicomotor; Saal (2007) los ha clasificado en siete dimensiones: la expresión verbal gestual, la relación tónica emocional, el movimiento del cuerpo en el espacio, la relación con los pares, la participación y el desenvolvimiento en la escuela, el nivel de autonomía e independencia, y la organización del espacio gráfico. En la Tabla 1 se describen estas dimensiones.

Tabla 1

Dimensiones de la inhibición psicomotriz.

Dimensión		Descripción
Expresión gestual	verbal	Los niños que presentan inhibición psicomotriz suelen hablar con tono de voz bajo, usar monosílabos, se expresan solo cuando se les solicita. La gestualidad y la expresión facial se encuentran empobrecidas.
Relación emocional	tónico	Se observa aumento de tono muscular, fatiga y bloqueo frente al movimiento, y escasos recursos para regular estados emocionales: con frecuencia se ruborizan cuando hablan y/o cuando los miran o se dirigen a ellos. Esto se incrementa frente al acercamiento y contacto corporal.
Movimiento del cuerpo en el espacio	del	Hay retención y repliegue del movimiento; el uso del espacio es mayormente reducido (Bucher, 1978), se refugian cerca de las paredes y rincones (Chokler, 1994). Su actitud es observadora y con poca participación; hay escasa exploración de los objetos y del espacio.
Relación con sus pares		Se muestran evitativos, distantes, poco participativos y solitarios. Suelen seguir a los otros, sin proponer iniciativas propias. En la relación grupal, prefieren pasar desapercibidos. En la relación “uno a uno”, esto se atenúa en un vínculo de confianza y seguridad.
Participación y desenvolvimiento en la escuela	y en	Tienden a mostrarse inseguros, temerosos. Se sobreadaptan a las normas y a las exigencias escolares. No se expresan ni se desenvuelven con libertad. No asumen riesgos. Prefieren evitar quedar expuestos a la mirada de los otros.
Nivel de autonomía e independencia	e	Se observa una marcada relación de dependencia respecto de la figura del otro. Poca autoría y autonomía de acción. Se muestran lábiles y vulnerables. Suelen dejar lugar a que los otros hagan y hablen por ellos.
Organización del espacio gráfico	del	Muestran preferencia por esta forma de expresión. Buen dominio óculo manual. La escritura y el dibujo suelen presentar características micrográficas. Se emplazan en espacios reducidos de la hoja. Tienden a incluir muchos detalles en sus dibujos, o por el contrario, ausencia de detalles y marcada pobreza en el dibujo.

De lo presentado surge que los distintos signos que se asocian a la inhibición psicomotriz tienen en común la particularidad de no convocar a la mirada del otro; se presentan de modo silencioso y suelen quedar ocultos, invisibilizados. No son reconocidos como parte de un funcionamiento que al niño le produce malestar y padecimiento. Esta invisibilidad de la inhibición psicomotriz se ve reforzada por dos vías: por un lado, por ser la inhibición funcional al entorno, al evitar todo tipo de provocación y confrontación, y por otro, por la búsqueda de los niños de pasar desapercibidos ante el peso que tiene para ellos la mirada de los otros. Esto nos plantea, entonces, el desafío de detenernos a observar más allá de lo evidente y de ofrecernos como interlocutores para interpretar el mensaje que expresan la quietud y el silencio del

niño inhibido psicomotor como señales de malestar y sufrimiento (Saal, 1995).

Con la intención de dar visibilidad a esta problemática y facilitar la detección temprana de los signos asociados a la inhibición psicomotriz, en el marco de trabajos anteriores (Saal, 2010, 2012, 2014, 2016) elaboramos el instrumento de despistaje SIP. Los instrumentos de despistaje (también denominados de pesquisa o *screening*) son un modo simple y rápido de identificar posibles retrasos o alteraciones que tienen la particularidad de ser inaparentes (Lejarraga, 2004) y pueden ser aplicados a grandes poblaciones sin necesidad de que quienes los administran sean profesionales (Doménech & Gómez de Terreros, 1995). Este instrumento no se propone alcanzar un diagnóstico sino constituirse en un primer

acercamiento, ser una herramienta de orientación que permita identificar indicios, alertar sobre posibles problemas con la intención de prevenir su acentuación.

Al estar diseñado para ser aplicado en las escuelas, el SIP cuenta con dos importantes ventajas: por un lado, es posible aplicarlo a grandes cantidades de niños, y por otro, se cuenta con un actor privilegiado para administrarlo, el docente, que se encuentra a una distancia óptima en relación a ellos (suficientemente cerca para conocerlos en sus modos de ser y hacer, pero no involucrado subjetiva y emocionalmente como los padres). Por estar orientado a la detección temprana, este instrumento se aplica en salas de 4 y 5 años de nivel inicial y en el primer ciclo de la escuela primaria. En este trabajo nos centraremos en los resultados obtenidos en su aplicación en la escuela primaria.

El objetivo de este trabajo es presentar el proceso de validación del SIP para su uso como instrumento de despistaje del trastorno de inhibición psicomotriz en niños de primer a tercer grado (primer ciclo de escolaridad primaria). Se presenta el análisis del funcionamiento de los ítems, el análisis de la estructura interna de la escala y su capacidad de discriminación entre niños con y sin inhibición psicomotriz. Además, se analiza la especificidad y sensibilidad del punto de corte para detectar casos de riesgo.

Método

Participantes

Participaron 675 niños (45.5% de género masculino y 54.5% de género femenino) de entre 5 y 10 años ($M = 7.11$; $DE = 0.91$) que asistían al primer ciclo en cuatro escuelas privadas laicas y religiosas del Área Metropolitana de Buenos Aires (AMBA, Argentina).

El SIP fue completado para cada uno de los niños por su maestra. Se consideró como criterio de exclusión que la maestra fuera suplente y/o hubiera estado al frente del grupo por un tiempo menor a tres meses. También se excluyeron los casos en que los niños participaban de más de un tratamiento ya que los instrumentos de pesquisa se utilizan en niños supuestamente sanos (con este fin se consideraron los tratamientos psicopedagógico, psicológico, fonoaudiológico, psicomotor y neurológico).

Para estudiar la capacidad de discriminación del SIP, del total de la muestra fueron seleccionados 217 niños para arribar a un diagnóstico (ver sección de procedimientos). Para obtener esta submuestra, se tomó a los niños que presentaban mayores puntajes en el SIP y se seleccionaron aleatoriamente niños que no habían obtenido puntajes altos. A partir del diagnóstico, se obtuvo un grupo de 32 niños que presentaban signos de inhibición psicomotriz y un grupo de 185 niños sin signos de inhibición psicomotriz. En la Tabla 2 se presentan las características sociodemográficas de ambos grupos de niños. Como puede observarse, los grupos no presentaban diferencias en las variables sociodemográficas. Si bien se observa una proporción ligeramente mayor de niñas en el grupo que presentaba signos de inhibición psicomotriz, esta diferencia no es estadísticamente significativa ($\chi^2_{(217, 1)} = 3.77, p = .052$).

Instrumentos

El instrumento de despistaje SIP fue elaborado a partir de la profundización teórica sobre el trastorno de inhibición psicomotriz y de información relevada en la casuística clínica. Desde su primera formulación (Saal, 2007), diversos estudios piloto (Saal, 2010, 2012) permitieron ajustarlo progresivamente. Se redujo sensiblemente

Tabla 2

Características sociodemográficas de los niños con y sin signos de inhibición psicomotriz.

	Niños sin signos de inhibición psicomotriz (n = 185)	Niños con signos de inhibición psicomotriz (n = 32)
Edad M (DE)	7.07 (0.94)	7.09 (0.93)
Sexo % (n)		
Femenino	47.0 (87)	65.6 (21)
Masculino	53.0 (98)	34.4 (11)
Grado		
Primero	43.8 (81)	37.5 (12)
Segundo	28.1 (52)	40.6 (13)
Tercero	28.1 (52)	21.9 (7)
Escolaridad		
Jornada simple	61.6 (114)	71.9 (23)
Jornada completa	38.4 (71)	28.1 (9)

la cantidad de ítems que lo conformaban y estos fueron reformulados para ser más cercanos al vocabulario docente con el propósito de facilitar su llenado. Los ítems correspondientes a la dimensión espacio gráfico se omitieron debido a que se evidenció que los docentes no solían contar con información suficiente en relación a ella (Saal & Reboursin, 2017; Saal, Tosto, Moreno, & Reboursin, en prensa).

El SIP se aplica en la escuela y releva diversos signos asociados a la inhibición psicomotriz. Al ser un instrumento de orientación, no pretende tener un valor diagnóstico acabado sino alertar sobre posibles dificultades. Consiste en un cuestionario que debe ser completado por la docente a cargo del grado o sala para cada uno de sus alumnos. Consta de un primer bloque en el que se relevan datos personales del niño. En una segunda parte, incluye 20 ítems presentados como afirmaciones que se responden mediante una escala Likert de cinco opciones: *nunca, poco, a veces, muy frecuente y siempre* (Saal & Reboursin, 2017).

Procedimientos

El estudio se llevó a cabo durante el período 2013-2017 en escuelas de la Ciudad Autónoma de Buenos Aires y el Gran Buenos Aires; estas fueron seleccionadas deliberadamente en función de contactos previos con las instituciones. En una primera etapa, se realizaron encuentros de sensibilización para poner en común con las docentes las manifestaciones corporales a las que el instrumento hace referencia (tono, postura y gestualidad facial, entre otras). Esto permitió arribar a significados compartidos para ciertos términos presentes en el instrumento y aclarar cualquier duda que pudieran tener. Se les entregó un instructivo con las indicaciones necesarias para completarlo correctamente. Se destacó que el SIP no busca rotular a los alumnos ni proporcionar un juicio valorativo de sus conductas sino proporcionar información sobre la actitud y desenvolvimiento del niño en el ámbito escolar. Se solicitó que fuera completado en ausencia de los alumnos.

Luego de la sensibilización a las docentes, se procedió a la aplicación del instrumento a los 675 niños que conformaron la muestra total. Esta fue realizada a partir del mes de junio (en la mayoría de los casos se llevó a cabo entre los meses de octubre y noviembre). La intención fue garantizar que hubiera finalizado el período de adaptación y favorecer un mayor conocimiento de los niños por parte de las docentes, así como también que los niños estuvieran familiarizados entre ellos, con su maestra, y con el espacio y las pautas de la institución.

Al finalizar la aplicación del instrumento a la muestra completa, se seleccionó una submuestra de 217 niños para establecer el diagnóstico que ofició de regla de oro (Pascucci, 2004) para analizar la capacidad de discriminación de los puntajes obtenidos por el instrumento. El diagnóstico, elaborado por psicomotricistas expertas, se basó en

observaciones realizadas en la escuela por estudiantes avanzadas y graduadas de la Licenciatura en Psicomotricidad de la Universidad Nacional de Tres de Febrero (UNTREF) que realizaron un entrenamiento de dos años para tal fin. Dichas observaciones tuvieron un carácter no participante, se extendieron por un período de aproximadamente dos meses y fueron programadas anticipadamente para no entorpecer las actividades escolares. Cada grado fue observado por dos observadoras que se mantuvieron constantes con el fin de minimizar la reactividad. Las observaciones fueron orientadas por una guía de pautas.

Para cubrir el despliegue del niño en diversas instancias de su experiencia escolar se abarcaron cuatro situaciones: clase, recreo, clase de Educación Física, y un taller de psicomotricidad programado intencionalmente para observar y registrar el despliegue del movimiento y accionar del cuerpo en relación al espacio, los objetos y los otros (la participación de los niños fue voluntaria, anónima y confidencial, y con consentimiento informado de los padres).

La información registrada a través de las distintas observaciones conformó el material que, junto con producciones gráficas de los niños realizadas en un taller de psicomotricidad y un dibujo libre elaborado en situación de clase con la maestra, fue evaluado y analizado por tres psicomotricistas expertas con amplia experiencia clínica. Las psicomotricistas debieron determinar para cada niño si presentaba signos asociados a la inhibición psicomotriz o no. La evaluación fue llevada a cabo de manera ciega (las evaluadoras desconocían los puntajes obtenidos en el SIP por los niños). Se consideró que un niño presentaba inhibición psicomotriz cuando se cumplían al menos cuatro de los siguientes criterios: a) tensión e incomodidad frente al acercamiento de otro (reacciones de prestancia, aumento del tono muscular); b) escasez y lentitud de movimientos; c) po-

co despliegue en el espacio y permanencia en la periferia; d) poca expresividad y mímica gestual; e) tono de voz bajo; f) escasez de interacciones espontáneas con pares y adultos; g) torpeza en el desarrollo de las coordinaciones; h) actitud hipoobservadora con respecto al entorno y repliegue de la acción y el uso del cuerpo; i) escasa disposición al juego corporal. Se consideró necesaria la persistencia de estos signos en al menos dos de las cuatro situaciones observadas: aula, recreo, clase de Educación Física y taller de psicomotricidad.

En los casos en que no hubo acuerdo unánime en las decisiones, se implementó una evaluación colectiva, confrontando los argumentos para arribar a un consenso. Solo se consideró válida una decisión diagnóstica cuando las tres psicomotricistas acordaron. La información que arrojó el SIP fue contrastada con las decisiones diagnósticas.

Análisis de datos

En primer lugar, se analizó la capacidad de discriminación de los ítems entre niños con y sin inhibición psicomotriz por medio de las curvas de la característica operativa del receptor (ROC, *receiver operating characteristic*). Se tomó como parámetro el área bajo la curva (ABC). El ABC debe ser mayor a .50 y acercarse a 1. A su vez, este valor no debe estar incluido en el intervalo de confianza (IC 95%; [Cerdeira & Cifuentes, 2012](#)).

Tomando los ítems que presentaban capacidad de discriminar entre los niños con y sin inhibición psicomotriz, se estudió la estructura interna de la escala por medio de un análisis factorial exploratorio (AFE) a través del análisis paralelo (PA; [Horn, 1965](#)). Se aplicaron los mínimos cuadrados no ponderados (ULS) para determinar el número de factores a retener ([Lorenzo-Seva, 1999](#)). El AFE se basó en matrices de correlacio-

nes policóricas, ya que son las apropiadas para escalas ordinales (Auné & Atorresi, 2017; Lloret-Segura, Ferreres-Traver, Hernández-Baeza, & Tomás-Marco, 2014). Se usó el método de rotación promin (Lorenzo-Seva, 1999). La adecuación de la matriz se evaluó mediante la prueba de Kaiser-Meyer-Olkin (KMO) y la prueba de esfericidad de Bartlett. El criterio de carga elegido para retener cada elemento en cada factor fue que su saturación fuera mayor que .30 (Hair, Anderson, Tatham, & Black, 2001). Adicionalmente, se calculó el porcentaje de varianza explicada.

Para probar la adecuación del modelo se utilizó, en primer lugar, la raíz media cuadrática residual (RMSR). Se siguió el criterio de Kelley (1935, como se citó en Lloret-Segura et al., 2014) para estimar si este índice arrojaba valores adecuados. Según este criterio, el modelo presenta un ajuste aceptable si el RMSR no excede o si se aproxima a $1/\sqrt{n}$ (n es el tamaño de muestra). Para nuestra muestra este valor debe ser próximo a .04. En segundo lugar, se utilizó la raíz cuadrática media del error de aproximación (RMSEA), que debe presentar valores iguales o inferiores a .08 (Hu & Bentler, 1999). En tercer lugar, se estimó la raíz cuadrática media ponderada (WRMR). Se considera que valores de WRMR inferiores a 1.0 representan un buen ajuste (Yu & Muthen, 2002). En cuarto lugar, estimamos el índice de bondad de ajuste (GFI), el índice ajustado de bondad de ajuste (AGFI), el índice de ajuste no normalizado (NNFI) y el índice comparativo de ajuste (CFI). En los últimos cuatro índices se considera que valores superiores a .90 indican un buen ajuste (Hu & Bentler, 1999). Para realizar estos análisis se usó el programa Factor (versión 10.9.02, Ferrando & Lorenzo-Seva, 2017).

Se estudió la confiabilidad por medio del coeficiente alfa de Cronbach y el estimador de fiabilidad ORION (*overall reliability of fully-informative prior oblique N-EAP scores*) propuesto

por Ferrando y Lorenzo-Seva (2016). Se considera que son aceptables valores de consistencia interna superiores a .70 (Hair et al., 2001). A su vez, se indagó la homogeneidad por medio de la correlación ítem-total corregida. Se consideró que un puntaje superior a .30 indicaba una adecuada homogeneidad (Tornimbeni, Pérez, & Olaz, 2008). Para realizar los análisis estadísticos se utilizó el paquete estadístico SPSS (versión 21) para Windows (IBM Corporation, 2012).

Para analizar la capacidad de discriminación de los puntajes del instrumento se calcularon las áreas bajo la curva ROC, la sensibilidad y la especificidad. Se tomó como regla de oro el diagnóstico de *inhibición psicomotriz* (ver Procedimientos). Estos análisis se utilizaron para determinar el punto de corte óptimo para predecir riesgo de presencia del trastorno.

Para realizar el AFE, el estudio de la confiabilidad y de la homogeneidad, se utilizó la muestra completa. La capacidad de discriminación de los ítems y de los puntajes del instrumento se analizó a partir de la submuestra de 217 casos que habían sido evaluados para establecer el diagnóstico de inhibición psicomotriz.

Resultados

Estadísticos descriptivos y capacidad de discriminación de los ítems

En la Tabla 3 se presentan los estadísticos descriptivos de los ítems de la primera versión de la escala. Asimismo, en dicha tabla se presentan los resultados de las curvas ROC para cada ítem. Como se puede ver, los ítems 7, 11 y 20 presentan puntajes de ABC que incluyen en el IC valores inferiores a .50. Esto indica que no discriminan entre los niños con y sin diagnóstico de inhibición psicomotriz. El resto de los ítems presenta puntajes de ABC superiores a .50 (incluyendo el IC).

Tabla 3
Estadísticos descriptivos de los ítems del SIP

Ítem	Mdn	M	DE	Asimetría	Curtosis	Mínimo	Máximo	ABC	IC 95% ABC	p
1	3.00	2.65	1.18	0.12	-0.98	1	5	.73	[.64, .82]	<.001
2	2.00	2.32	1.15	0.40	-0.80	1	5	.80	[.72, .88]	<.001
3	1.00	1.79	0.97	1.05	0.25	1	5	.70	[.60, .81]	<.001
4	2.00	2.21	1.02	0.49	-0.61	1	5	.62	[.51, .73]	.033
5	2.00	2.22	1.19	0.58	-0.71	1	5	.76	[.69, .84]	<.001
6	2.00	2.25	1.05	0.39	-0.65	1	5	.70	[.61, .79]	<.001
7	2.00	2.37	1.22	0.38	-0.77	1	5	.59	[.47, .71]	.107
8	2.00	2.27	1.10	0.47	-0.62	1	5	.68	[.59, .77]	.001
9	4.00	3.68	1.05	0.41	-0.76	1	5	.72	[.63, .80]	<.001
10	3.00	2.91	1.24	0.07	-1.01	1	5	.69	[.59, .79]	.001
11	1.00	1.76	1.04	1.51	0.34	1	5	.57	[.46, .68]	.198
12	2.00	2.24	1.08	0.42	-0.69	1	5	.65	[.54, .75]	.008
13	3.00	3.04	1.22	-0.07	-0.84	1	5	.72	[.62, .81]	<.001
14	2.00	1.97	1.03	0.67	-0.64	1	5	.67	[.58, .76]	.002
15	4.00	3.79	0.99	-0.55	-0.35	1	5	.66	[.55, .77]	.004
16	3.00	3.04	0.93	0.01	-0.10	1	5	.71	[.61, .80]	<.001
17	1.00	1.55	0.79	1.36	1.24	1	5	.61	[.50, .73]	.042
18	3.00	3.17	1.21	-0.02	-1.08	1	5	.63	[.51, .76]	.015
19	2.00	1.96	1.05	0.76	-0.43	1	5	.79	[.70, .87]	<.001
20	1.00	1.61	0.74	1.07	0.80	1	5	.55	[.44, .67]	.341

Nota. ABC = área bajo la curva. IC = intervalo de confianza.

Esto implica que poseen una capacidad, mayor que la que se daría por azar, de diferenciar entre niños con y sin diagnóstico de inhibición psicomotriz (Cerdeira & Cifuentes, 2012; Hajian-Tilaki, 2013).

Validez factorial

El coeficiente KMO y la prueba de esfericidad de Bartlett mostraron la suficiencia de la matriz (ver Tabla 4). Se encontró una estructura de dos factores que explican el 72.2% de la varianza.

El factor *inhibición psicomotriz* (12 ítems) agrupa los ítems que refieren al núcleo de la inhibición psicomotriz y se vinculan al tono y la emoción, la comunicación y relación con los pares, el movimiento y la participación. El factor *obediencia a las normas* (3 ítems) guarda relación con el nivel de dependencia del niño en su vínculo con el adulto. Los factores son independientes entre sí ($r = -.04, p = .36$).

El modelo de dos factores presenta buenos indicadores de ajuste: RMSEA = 0.07; RMSM = 0.04 (valor próximo a 0.04, según el criterio de Kelley, 1935 como se citó en Lloret-Segura et al.,

Tabla 4
Matriz de cargas rotadas SIP.

Ítem	Inhibición Psicomotriz	Obediencia a las normas
1.	.84	
2.	.84	
3.	.67	
4.	.60	
5.	.80	
6.	.81	
8.	.85	
9.		.83
10.	.83	
12.	.74	
14.	.75	
15.		.85
16.	.59	
18.		.71
19.	.68	
% de varianza explicada	53.40	18.82
<i>Eigenvalue</i>	7.26	2.69

Nota. n = 675. KMO = .91, Prueba de Bartlett: $\chi^2_{(675, 105)} = 5786.5$ ($p < .001$).

2014); WRMR = 0.05; GFI = 0.99; AGFI = .99; NNFI = 0.98; CFI = 0.99.

Ambos factores presentan una buena consistencia interna: *inhibición*, $\alpha = .92$; *obediencia*, $\alpha = .80$. asimismo, el estimador de confiabilidad ORION indica una buena confiabilidad para ambas subescalas (ver Tabla 5).

Tabla 5
Confiabilidad de las subescalas.

	Inhibición psicomotriz	Obediencia a las normas
Alfa de Cronbach	.92	.80
ORION	.95	.88

Los ítems presentan una buena homogeneidad: coeficientes de correlación ítem/total corregida superiores a .30. Ninguno de los ítems aumenta el alfa al ser eliminado (ver Tabla 6).

Tabla 6
Homogeneidad de los ítems y alpha de Cronbach si se elimina un elemento.

Ítem	Correlación ítem/total corregida	Alpha si se elimina el ítem
Inhibición psicomotriz		
1.	.76	.91
2.	.77	.91
3.	.57	.92
4.	.54	.92
5.	.73	.91
6.	.73	.91
8.	.77	.91
10.	.75	.91
12.	.64	.92
14.	.66	.92
16.	.54	.92
19.	.60	.92
Obediencia a las normas		
9.	.65	.69
15.	.66	.69
18.	.60	.76

Tabla 7

Puntos de corte, sensibilidad, especificidad y área bajo la curva de las dimensiones de la inhibición psicomotriz.

Dimensión	Punto de corte	Sensibilidad	Especificidad	ABC	IC 95% ABC	<i>p</i>
Inhibición Psicomotriz	35.5	.78	.72	.79	[.72, .87]	.001
Obediencia a las normas	9.5	.82	.48	.70	[.59, .80]	.001

Sensibilidad, especificación y capacidad de discriminación

El ABC del primer factor fue de .79 ($p < .001$; ver Tabla 7). Se estableció un punto de corte de 35.5, que ofrece la mejor combinación de sensibilidad (.78) y especificidad (.72). El ABC para el segundo factor fue de .70 ($p < .001$). Sin embargo, el punto de corte (9.5) ofrece una buena sensibilidad (.82), pero una baja especificidad (.48).

Discusión

El presente trabajo tuvo como objetivo presentar el proceso de validación del SIP para su uso como instrumento de despistaje del trastorno de inhibición psicomotriz en niños de primer a tercer grado (primer ciclo de escolaridad primaria). El desarrollo y la validación de este instrumento obedecen a la necesidad de contar con un instrumento útil para detectar tempranamente el trastorno de inhibición psicomotriz.

Al tratarse de niños que no generan disturbios en su relación con el entorno, obedientes, muchas veces sobreadaptados a los requerimientos que se les imponen; con escasa y limitada participación; con poco uso del cuerpo como modo de expresar, a través del movimiento y de la acción, una modalidad de conquista activa y prota-

gónica; los niños con inhibición psicomotriz suelen pasar desapercibidos e inadvertidos por el entorno. Así, es bastante extraño que se identifique esta sintomatología como una posible alteración que impacta en su estructuración psíquica-corporal. Por estos motivos se ha considerado necesario plantear estos signos como un problema que trasciende la investigación científica y académica, ya que su lectura por las personas en el entorno y los modos en que se manifiesta, suelen ser en líneas generales valorados como señales positivas de conducta obediente (muchas veces por desconocimiento, otras por resultar poco disruptivos al entorno). Al tratarse de un aspecto de la conducta de los niños que no está problematizado en el contexto social y escolar, la “buena conducta” asociada a la inhibición psicomotriz hace que se dificulte visibilizar y reconocer el malestar y padecimiento que a ella se asocian, como un funcionamiento perturbador para el niño, su niñez y su infancia (Saal, 2011, 2013; Saal & Reboursin, 2017).

A partir de los resultados presentados se puede inferir que el SIP es un instrumento confiable y válido para identificar y visibilizar niños en riesgo de presentar inhibición psicomotriz. La evidencia encontrada señala que los ítems que componen la escala presentan una adecuada capacidad de discriminación y una buena homogeneidad.

En cuanto a la validez factorial, se ha logrado identificar una estructura interna de dos dimen-

siones empíricas con sentido teórico. La primera dimensión, *inhibición psicomotriz* (12 ítems) abarca los ítems que se desprenden de cinco de las siete dimensiones teóricas consideradas: *relación tónico-emocional, expresión verbal-gestual, movimiento del cuerpo en el espacio, participación y desenvolvimiento en la escuela y relación con los pares*. Se constituye así en una dimensión empírica que cubre el núcleo de la inhibición psicomotriz, considerando sus signos centrales.

Resulta relevante destacar que los ítems que hacen referencia al movimiento no se han incluido en la versión final del instrumento, habida cuenta de que las primeras formulaciones del instrumento (en las que se consideraban explícitamente los desplazamientos en el espacio) mostraron una tendencia de las docentes a marcar para todos los niños la ausencia de movimiento, que podría estar en relación con lo que se considera esperable y deseable en el marco de la institución escolar.

Los ítems que estudian el desenvolvimiento del niño en los recreos no se incluyen en ninguno de los factores, debido a que los resultados mostraron que estos ítems no permiten discriminar entre los niños que presentan inhibición psicomotriz y los que no. Consideramos que esto podría deberse a distintos motivos. Por un lado, podría pensarse que las docentes utilizan el momento del recreo como un momento de descanso, en el que su atención está focalizada en el cuidado de los niños (evitando que se lastimen, o que se peleen) pero no lo valoran como un espacio para observar y conocer mejor a los niños. Otra posibilidad es que se deslice en las respuestas de las docentes lo que consideran como esperable para la instancia del recreo, esto es, que los niños participen y disfruten de él.

La segunda dimensión, *obediencia a las normas* (3 ítems) se relaciona con la dimensión teórica de nivel de autonomía e independencia en relación al otro, en particular a la figura del do-

cente. Las dimensiones empíricas mencionadas no presentan una asociación significativa entre sí. Esto quiere decir que, aunque ambas son aspectos relacionados con este trastorno, se presentan de manera independiente. Por este motivo, no se considera indicado calcular un puntaje total para la escala.

En cuanto a la confiabilidad de la escala, se encontraron buenos índices de consistencia interna y del estimador de confiabilidad ORION. Asimismo, la consistencia interna no aumenta al eliminar ninguno de los ítems.

Por otro lado, el análisis de las curvas ROC presenta evidencia estadísticamente significativa de que los puntajes obtenidos por el instrumento permiten diferenciar entre niños con y sin inhibición psicomotriz. Asimismo, se pudo identificar un puntaje de corte para la primera dimensión con adecuados niveles de sensibilidad y especificidad. Esto indica que dicha dimensión es una buena medida para el rastillaje de la inhibición psicomotriz en niños de primer ciclo. Sin embargo, no se pudo identificar un punto de corte para la segunda dimensión que combine adecuados niveles de sensibilidad y especificidad. De todos modos, sería relevante que futuras investigaciones analicen las posibilidades de combinar estas subescalas para potenciar la capacidad del instrumento de identificar casos de riesgo de inhibición psicomotriz.

Cuando los puntajes altos en la dimensión de obediencia se presentan junto con valores altos en la dimensión de inhibición psicomotriz, la primera de ellas indicaría un alto nivel de dependencia del otro, de sumisión. Por otra parte, cuando los puntajes en la dimensión de obediencia no se acompañan de altos valores en la dimensión de inhibición psicomotriz, sino que se asocian a altos niveles de participación en la escuela (niños que participan activamente, que se ofrecen para mostrar sus trabajos y presentar sus ideas), suelen ser

indicios de los tipos de conducta que tradicionalmente se asocian con los “buenos alumnos” que respetan las normas.

De este modo, el SIP se presenta como un instrumento válido y confiable que permite discriminar estas dos situaciones. Por un lado, conductas propias de lo que supone ser un buen alumno, obediente y ordenado, sin que ello implique un padecimiento, sino solo un modo de ser y estar. Por otro lado, una situación bien distinta, estaría vinculada a una alteración en la utilización y en el despliegue del funcionamiento y accionar del cuerpo, que coarta y bloquea el desarrollo creativo, autónomo y enriquecido.

Al ser los signos asociados a la inhibición psicomotriz poco molestos, silenciosos, poco disruptivos para la institución educativa y para el entorno, el problema a plantear está emparentado a su escasa visibilización; lo cual podría explicar que no haya antecedentes sobre instrumentos que permitan evaluar esta problemática en la que se ven afectados los niños en el devenir de su infancia como sujetos plenos de derechos para expresarse con plenitud. Esta situación de invisibilidad de la inhibición psicomotriz es opuesta a la de notable visibilización que alcanzan afecciones como la inestabilidad psicomotriz (Bergés, 1974) o el trastorno por déficit de atención con hiperactividad (American Psychiatric Association, 2014) que ya hemos esbozado en un apartado anterior.

El contraste en cuanto a la profundización conceptual y al desarrollo de investigaciones en relación a estos dos trastornos radica en que los niños que presentan inestabilidad psicomotriz provocan y llaman a la mirada del otro (Bergés, 1985), generando malestar en el entorno, mientras que los niños más retraídos, inhibidos, suelen pasar desapercibidos.

Por otra parte, el presente trabajo no está exento de limitaciones. En primer lugar, se trabajó con una muestra no probabilística y por conve-

nencia. Por esta razón los resultados no son directamente generalizables a todos los niños de primer ciclo de Argentina. Sin embargo, se contó con una muestra heterogénea de niños pertenecientes a escuelas laicas y religiosas, y de zonas del AMBA con distinto nivel socioeconómico. De todos modos, sería importante que futuras investigaciones repliquen este estudio con niños de otras regiones del país y de áreas rurales. En segundo lugar, ante la imposibilidad de observar a todos los niños para los que se administró el SIP, se debió tomar una submuestra para realizar las decisiones diagnósticas. Al ser los niños con puntajes altos en el SIP una minoría (lo que guardaría relación con la baja incidencia de la inhibición psicomotriz en la población), tomar esta submuestra aleatoriamente podría haber traído aparejado el inconveniente de que los casos de puntajes altos en el SIP no formaran parte de ella. Se tomó, por ello, a los niños que presentaban puntajes altos en el instrumento y se completó aleatoriamente el resto de cada grado para garantizar que la variedad de puntajes en el SIP estuvieran representados en la submuestra. Futuros estudios deberán replicar este análisis tomando una muestra aleatoria, lo suficientemente grande como para incluir la cantidad necesaria de niños con altos puntajes en el SIP.

Habiendo ya aclarado que el SIP no es una herramienta diagnóstica, sino que, en tanto despistaje, es un primer acercamiento que alerta acerca de la presencia de signos asociados a la inhibición psicomotriz, podemos afirmar que este instrumento brinda la información necesaria para actuar oportunamente con la intención de prevenir su acentuación. Esto se puede implementar, según el caso, ofreciendo herramientas a los docentes o haciendo posible la solicitud de una interconsulta y/o derivación a un especialista para que los niños reciban un diagnóstico y tratamiento adecuado. Arribar a un diagnóstico implica una tarea del profesional psicomotricista implementando técni-

cas y métodos específicos del quehacer psicomotor que exceden el propósito de este instrumento.

Sobre la base de estos análisis, los resultados obtenidos demuestran la importancia de validar este instrumento ya que quienes lo aplican, los docentes, cuentan con la oportunidad de observar a los niños en diferentes instancias del tránsito escolar, lo que brinda a su vez al psicomotricista la posibilidad de conocer estas facetas, y aportar una interpretación más completa de lo observado. La cantidad de casos relevados que presentaban signos asociados a la inhibición psicomotriz confirman lo beneficioso y adecuado de su implementación en el primer ciclo. Se evidencia así el lugar protagónico que adquiere el cuerpo, el accionar concreto en la realidad y el despliegue activo e interactivo en esta etapa vital del desarrollo del niño en el proceso de su infancia. Todos estos son procesos fundantes para la construcción y apropiación de los aprendizajes.

Al mismo tiempo es fundamental destacar la importancia de contar con esta herramienta de aplicación que complementa al campo de la educación con la promoción en el campo de la salud. Se abre un espacio innovador con nuevas oportunidades de acercamiento, de intercambio e interacción temprana, oportuna y directa entre profesionales psicomotricistas y docentes a la hora de evaluar diferentes modos de intervención.

Consideramos que problematizar esta modalidad corporal es el camino para intervenir oportuna y eficazmente. Rescatamos que al implementarse el instrumento de despistaje SIP, este, junto con la sensibilización que lo antecede, tendrá consecuencias también en la mirada de las maestras, que como pudimos registrar a lo largo de la investigación, se solidarizan con la mirada psicomotriz hacia el funcionamiento integral y autónomo del cuerpo del niño en sus modos de ser y estar.

Referencias

- De Ajuriaguerra, J. (1984). *Manual de psiquiatría infantil* (4ª ed.). Barcelona, España: Masson.
- American Psychiatric Association. (2014). *Manual diagnóstico y estadístico de trastornos mentales* (5ª ed). Madrid, España: Médica Panamericana.
- Auné, S. E., & Attorresi, H. F. (2017). Dimensionalidad de un Test de Conducta Prosocial. *Revista Evaluar*, 17(1), 29-37. doi: [10.35670/1667-4545.v17.n1.17072](https://doi.org/10.35670/1667-4545.v17.n1.17072)
- Benavides, T. (2003). La clinique psychomotrice. En T. Benavides, *Sémiologie psychomotrice de l'enfant* (pp. 11-22). París, Francia: Université Paris - VI. Faculté de Médecine Pierre et Marie Curie.
- Bergés, J. (1974). Algunos temas de investigación en psicomotricidad. *Cuadernos de Terapia Psicomotriz*, 1, 5-14.
- Bergés, J. (1985). Los trastornos psicomotrices en la infancia. En S. Lebovici, R. Diatkine & M. Soulé (Coords.), *Traité de psychiatrie de l'enfant et de l'adolescent*. París, Francia: Universitaires de France.
- Bergés, J. (1991). El cuerpo: De la neurofisiología al psicoanálisis. *Cuadernos de Psicomotricidad y Educación Especial*, 1(2), 5-15.
- Boscaini, F. (2013). Por una contribución al debate sobre la inestabilidad psicomotora. En F. Boscaini (Comp.), *Clínica Psicomotriz* (pp. 183-194). Verona, Italia: CISERPP.
- Bucher, H. (1978). El esquema de una investigación. En H. Bucher, *Estudio de la personalidad del niño a través de la exploración psicomotriz* (pp. 39-48). Barcelona, España: Toray Masson.
- Cerda, J., & Cifuentes, L. (2012). Uso de curvas ROC en investigación clínica. Aspectos teórico-prácticos. *Revista Chilena de Infectología*, 29(2), 138-141. doi: [10.4067/s0716-10182012000200003](https://doi.org/10.4067/s0716-10182012000200003)
- Chokler, M. (1994). *Los organizadores del desarrollo psicomotor*. Buenos Aires, Argentina: Cinco.
- De León, C. (2010). Una propuesta de clasificación. En C. De León, *Las alteraciones psicomotrices*. Diagnós-

- tico en psicomotricidad* (pp. 57-90). Montevideo, Uruguay: Tradinco.
- De León, C. (Coord.). (2012). *Diálogos con Bernard Aucouturier*. Montevideo, Uruguay: Espacio Lunes.
- Defontaine, J. (1980). *Manual de reeducación psicomotriz*. Barcelona, España: Médica y Técnica.
- Domènech, E., & Gómez de Terreros, M. (1995). Evaluación psicológica en la temprana infancia. En F. Silva Moreno (Ed.), *Evaluación psicológica en niños y adolescentes* (pp. 173-219). Madrid, España: Síntesis.
- Ferrando, P. J., & Lorenzo-Seva, U. (2016). A note on improving EAP trait estimation in oblique factor-analytic and item response theory models. *Psicológica*, 37(2), 235-247. Recuperado de <https://www.uv.es/psicologica>
- Ferrando, P. J., & Lorenzo-Seva, U. (2017). Program FACTOR at 10: Origins, development and future directions. *Psicothema*, 29(2), 236-240. Recuperado de <http://www.psicothema.com>
- IBM Corporation. (2012). IBM SPSS Statistics 21.0. [software de cómputo]. Armonk, NY: IBM.
- Hajian-Tilaki, K. (2013). Receiver Operating Characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Caspian Journal of Internal Medicine*, 4(2), 627-635. Recuperado de <http://caspijim.com>
- Hair, J., Anderson, R., Tatham, R., & Black, W. (2001). *Análisis multivariante* (5ª ed.). Madrid, España: Prentice Hall Iberia.
- Horn, J. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179-185. doi: [10.1007/bf02289447](https://doi.org/10.1007/bf02289447)
- Hu, L., & Bentler, P. M. (1999). Cut-off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55. doi: [10.1080/10705519909540118](https://doi.org/10.1080/10705519909540118)
- Le Boulch, J. (1992). *Hacia una ciencia del movimiento humano. Introducción a la psicokinética*. Barcelona, España: Paidós.
- Lejarraga, H. (2004). La pesquisa de problemas de desarrollo. La Prueba Nacional de Pesquisa. En H. Lejarraga (Ed.), *Desarrollo del niño en contexto* (pp. 463-475). Buenos Aires, Argentina: Paidós.
- Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., & Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: Una guía práctica, revisada y actualizada. *Anales de Psicología*, 30(3), 1151-1169. doi: [10.6018/analesps.30.3.199361](https://doi.org/10.6018/analesps.30.3.199361)
- Lorenzo-Seva, U. (1999). Promin: A method for oblique factor rotation. *Multivariate Behavioral Research*, 34(3), 347-365. doi: [10.1207/s15327906mbr3403_3](https://doi.org/10.1207/s15327906mbr3403_3)
- Merten, E. C., Cwik, J. C., Margraf, J., & Schneider, S. (2017). Overdiagnosis of mental disorders in children and adolescents (in developed countries). *Child and Adolescents Psychiatry and Mental Health*, 11(1), 1-11. doi: [10.1186/s13034-016-0140-5](https://doi.org/10.1186/s13034-016-0140-5)
- Molina de Costallat, D. (1987). *Psicomotricidad III. Educación gestual*. Buenos Aires, Argentina: Losada.
- Pascucci, M. C. (2004). Validación de instrumentos de pesquisa de trastornos de desarrollo. En H. Lejarraga (Ed.), *Desarrollo del niño en contexto* (pp. 477-515). Buenos Aires, Argentina: Paidós.
- Piaget, J., & Inhelder, B. (1981). *Psicología del niño* (10ª ed.). Madrid, España: Morata.
- Rodríguez, M. B. (2011). Déficit atencional. Perspectiva psicomotriz en la comprensión de una temática actual. En L. González (Comp.), *Temas de investigación en psicomotricidad* (pp. 35-56). Sáenz Peña, Argentina: Eduntref.
- Saal, S. (1995). Psicomotricidad y el niño inhibido. *Crónicas clínicas en relajación terapéutica y psicomotricidad*, 3, 38-49.
- Saal, S. (2007). *Caracterización y estudio del trastorno de inhibición psicomotriz* (Tesis de licenciatura inédita). Universidad Nacional de Tres de Febrero, Tres de Febrero, Buenos Aires, Argentina.
- Saal, S. (2010). *Desarrollo de un instrumento de despistaje del Trastorno de inhibición Psicomotriz en la infancia (TIP) a través del reconocimiento de sus signos asociados*. Proyecto de Investigación correspondiente a la programación científica 2010-2011. Universidad Nacional de Tres de Febrero, Caseros, Buenos Aires, Argentina.

- Aires, Argentina.
- Saal, S. (2011). Caracterización y estudio del trastorno de inhibición psicomotriz. En L. González (Comp.), *Temas de investigación en psicomotricidad* (pp. 7-34). Sáenz Peña, Argentina: Eduntref.
- Saal, S. (2012). *Ajustes a un instrumento de despistaje del Trastorno de Inhibición Psicomotriz en la infancia (TIP) y primeras aproximaciones a su validación*. Proyecto de Investigación correspondiente a la programación científica 2012-2013. Universidad Nacional de Tres de Febrero, Caseros, Buenos Aires, Argentina.
- Saal, S. (2013). Sui “buoni comportamenti” e i “bravi scolari”. Uno sguardo sull’inibizione psicomotoria nell’ambito scolastico. *Res. Recherche e studi in psicomotricità*, XXI(2), 5-10.
- Saal, S. (2014). *Avances y mejoras al instrumento de despistaje del trastorno de inhibición psicomotriz en la infancia TIP e información relevada de los primeros cruces con el criterio externo (observación e interpretación diagnóstica)*. Proyecto de Investigación correspondiente a la programación científica 2014-2015. Universidad Nacional de Tres de Febrero, Caseros, Buenos Aires, Argentina.
- Saal, S. (2016). *Validación del instrumento de despistaje del trastorno de inhibición psicomotriz en la infancia TIP*. Proyecto de Investigación correspondiente a la programación científica 2016-2017. Universidad Nacional de Tres de Febrero, Caseros, Buenos Aires, Argentina.
- Saal, S., & Reboursin, J. (2017). ¿Cuándo el planteo del problema se vuelve un problema a plantear? Hacia la validación de un instrumento de despistaje de la inhibición psicomotriz. En N. Fernández-Lamarra (Org.), *Políticas, actores y prácticas. Estudios de política y administración de la educación III* (pp. 385-402). Sáenz Peña, Argentina: Eduntref.
- Saal, S. (2019). *Cuestionario de Signos Observables Asociados a la Inhibición Psicomotriz (SIP)*. Dirección Nacional del Derecho de Autor. Propiedad intelectual número: RE-2019-06247311-APN-DNDA#MJ. Referencia: RL-2019-06247248-APN-DNDA#MJ.
- Saal, S., Tosto, P., Moreno, J., & Reboursin, J. (en prensa). ¿Cómo visibilizar los signos asociados a la inhibición psicomotriz? Cartografía de una investigación. En L. González (Comp.), *La construcción de un oficio. Investigaciones en Psicomotricidad 2009-2018*. Buenos Aires, Argentina: Eduntref.
- Tornimbeni, S., Pérez, E., & Olaz, F. (2008). *Introducción a la Psicometría*. Buenos Aires, Argentina: Paidós.
- Untoiglich, G. (2011). *Versiones actuales del sufrimiento infantil. Una investigación psicoanalítica acerca de la desatención y la hiperactividad*. Buenos Aires, Argentina: Noveduc.
- Wallon, H. (1965). *Los orígenes del carácter en el niño*. Buenos Aires, Argentina: Lautaro.
- Wille, A. M. (2007). *L’inibizione psicomotoria. L’azione tesa e l’azione sospesa*. Napoli, Italia: Cuzzolin.
- Wille, A. M. (2012). *L’inerzia decisionale. La terapia psicomotoria con persone senza iniziativa motoria e con grave ritardo mentale*. Napoli, Italia: Cuzzolin.
- Yu, C., & Muthen, B. (2002). *Evaluation of model fit indices for latent variable models with categorical and continuous outcomes*. Trabajo presentado en el Annual Meeting of the American Educational Research Association, New Orleans, LA.

Test de Raven, baremos argentinos para el rango 19-30 años y efecto Flynn

Raven's Progressive Matrices, Argentinean Norms for ages 19 to 30 and Flynn Effect

Lilia Rossi-Casé¹, Stella Maris Doná¹, Ramiro Garzaniti^{1*}, Bruno Biganzoli¹

1 - Facultad de Psicología de la Universidad Nacional de La Plata.

Introducción
Método
Resultados
Discusión
Referencias

Recibido: 19/12/2019 **Revisado:** 06/02/2020 **Aceptado:** 25/02/2020

Resumen

El objetivo de este artículo es presentar los resultados definitivos del proyecto de investigación dedicado a la construcción de baremos para el Test de Raven, Escala General (Raven, Raven, & Court, 2003), para las edades de 19 a 30 años en la ciudad de La Plata, Argentina. La muestra estuvo conformada por 979 participantes y se dividió en cuatro intervalos de edad. Los resultados mostraron que a medida que aumenta la edad de los participantes, el rendimiento promedio de los grupos mejora; y que la heterogeneidad mostrada en sus respuestas resulta similar entre grupos. No se encontraron diferencias estadísticamente significativas entre los rendimientos de ambos sexos, ni entre los resultados actuales y los baremos del año 2000. Se constata una detención del incremento de puntajes esperable según el efecto Flynn. Se ensayan discusiones en torno a las diferencias cognitivas intergeneracionales que podrían originar este fenómeno, y se compara estos resultados con los obtenidos en otros países.

Palabras clave: *Test de Raven, baremos, Argentina, 19-30 años, efecto Flynn*

Abstract

This article aims to show the final results of the research project devoted to build norms for Raven's Standard Progressive Matrices (Raven, Raven, & Court, 2003), on subjects aged 19 to 30 years old in the city of La Plata, Argentina. The sample consisted of 979 subjects of both sexes, and was divided into four age groups. The results show that as the participant's age increases, the average group's performance improves, and variability remains similar between groups. There were no statistically significant differences found between sexes, or between the current results and those from the 2000 norms. It was found that the scores' increase, expected according to the Flynn effect, has stopped, and the possible intergenerational cognitive differences which could explain this phenomenon are discussed. These results are also compared with those from recent research made in other countries.

Key words: *Raven's Progressive Matrices, norms, Argentina, ages 19-30, Flynn effect*

* **Correspondencia a:** Ramiro Garzaniti. E-mail: ramiro.garzaniti.unlp@gmail.com

Cómo citar este artículo: Rossi-Casé, L., Doná, S. M., Garzaniti, R., & Biganzoli, B. (2020). Test de Raven, baremos argentinos para el rango 19-30 años y efecto Flynn. *Revista Evaluar*, 20(1), 17-31. Recuperado de <https://revistas.unc.edu.ar/index.php/revaluar>

Introducción

El Test de Matrices Progresivas de Raven (en adelante Test de Raven) fue publicado por primera vez en el año 1938 por John C. Raven (Rossi-Casé, et al. 2016). Se trata de un test no verbal, de capacidad intelectual, de habilidad mental general. Examina el factor *g*, que surge de la Teoría Ecléctica de los dos Factores, propuesta por Charles Spearman en 1904 (Raven, Raven, & Court, 2003). Con este fin, pone en juego procesos de educación de relaciones y correlaciones sobre un material en el que las variables no son obvias, de modo que se deben extraer nuevas comprensiones a partir de la información dada (Raven et al., 2003). Si bien hay investigaciones que afirman que los puntajes del test no solo reflejan el factor *g*, sino también otros factores como la capacidad visoespacial, la motivación y las estrategias de resolución de problemas (Gignac, 2015; Hayes, Petrov, & Sederberg, 2015), Raven es considerado el test con mayor saturación de factor *g* (Gignac, 2015).

Se trata de un test que mide la inteligencia fluida, por lo que coincidimos con Fernández-Liporace, Ongarato, Saavedra y Casullo (2004, p. 52) al afirmar que “esto explica la utilización tan difundida de las Matrices en los ámbitos de investigación”, ya que los resultados que aporta permiten comparar poblaciones y/o sujetos que “han sido expuestos a situaciones educativas formales e informales de naturaleza bien disímil”.

A partir del análisis de baremos de diversos tests de inteligencia y de los datos obtenidos de casi 50 mil participantes por internet, Hartshorne y Germine (2015) afirman que no hay una única edad en la que las personas logran un desempeño máximo en todas las tareas, o en la mayoría de ellas. Parten de la ya mencionada separación entre inteligencia fluida e inteligencia cristalizada, de las cuales suele entenderse que la primera

alcanza su punto máximo en la adultez temprana, mientras que la segunda lo hace en la mediana edad.

Los autores refieren que hay una gran variabilidad en estas edades de desempeño máximo entre las diversas tareas que involucra cada uno de estos dos tipos de inteligencia (por ejemplo, memoria a corto plazo para nombres o rostros en el caso de la inteligencia fluida, vocabulario e información en el de la cristalizada), y por ese motivo la distinción entre inteligencia fluida y cristalizada en cuanto a un momento de desempeño máximo ya no sería relevante. No obstante, el estudio del cambio de los puntajes en los tests de inteligencia a través de las generaciones ha mantenido vivo el interés en estos estudios en la actualidad.

El uso generalizado de los tests de inteligencia demostró que los puntajes medios obtenidos aumentaban regularmente y de manera notable con el correr del tiempo en todo el mundo (Flynn, 1984; Rossi-Casé, Neer, & Lopetegui, 2001). Esta constatación indica que, para la misma población, las normas para el cociente intelectual se vuelven obsoletas conforme pasan los años. Es decir que en la actualidad, para obtener el mismo puntaje transformado que hace algunas décadas, las puntuaciones directas deberían ser más altas, debiendo resolverse un mayor número de problemas. Estos aumentos son mayores en aquellos tests que miden la inteligencia fluida (Sundet, Barlaug, & Torjussen, 2004). Esto llevó a pensar que el aumento está relacionado al factor *g* de la inteligencia y no a los factores específicos (i.e., factores *e*).

En su estudio inicial, Flynn (1984) estableció que la magnitud de aumento del cociente intelectual, medido inicialmente con las escalas Wechsler y Stanford-Binet, era de 0.3 puntos por año o 3 puntos por década. Estos estudios se realizaron comparando muestras estadounidenses

entre los años 1932 y 1978. Más adelante, la misma tendencia se corroboró con datos de otros 20 países (Flynn & Rossi-Casé, 2012; Gignac, 2015; Sundet et al., 2004; Teasdale & Owen, 2007).

Con respecto a la Argentina, este equipo realizó una primera constatación del efecto Flynn en la ciudad de La Plata y alrededores al actualizar los baremos del Test de Raven, Escala General, en el año 2000, y compararlos con la estandarización del año 1964 (Rossi-Casé et al., 2001). Esta comparación mostró un considerable aumento en los puntajes directos del test para todas las edades (Flynn & Rossi-Casé, 2011, 2012; Rossi-Casé, Neer, & Lopetegui, 2002, 2017; Rossi-Casé et al., 2014, 2016). No obstante, durante la década siguiente, empezó a constatarse que dichos puntajes aumentaban a un ritmo mucho menor.

La mencionada desaceleración en el aumento de los puntajes fue observada por vez primera en los países escandinavos. Tal es el caso de investigaciones longitudinales hechas en Noruega (Sundet et al., 2004), en las cuales se muestra una disminución en los puntajes medios de una batería de tests administrados a los jóvenes de 18 años, antes de entrar al servicio militar, desde los años 50. Las pruebas utilizadas son tests de matemática y lenguaje, similares a los subtests del WAIS IV (Wechsler, 2008), y un test no verbal que fue construido de forma similar al Test de Raven. Los dos primeros miden la inteligencia cristalizada, mientras que el último mide la inteligencia fluida.

Teasdale y Owen (2007) estudiaron los datos aportados por una batería similar en Dinamarca, la cual desde 1957 se administra a todos los jóvenes de 18 años que ingresan al servicio militar. De los cuatro tests que la componen, los autores afirman que el de Matrices de Letras es el más similar a las Matrices Progresivas de Raven. Aquí encontraron que, si bien hubo un pequeño incremento de los puntajes entre 1988 y 1998, los mismos disminuyeron en 2003-2004, incluso por

debajo de los obtenidos en 1988. Esto se observó en jóvenes de todos los niveles educativos. Con estos datos, los autores apoyan la afirmación de que en lo que va del siglo XXI ha habido muy poca evidencia sobre la continuidad del efecto Flynn, tal como se lo define hasta este momento, es decir, como el aumento continuo de las medias en los tests de inteligencia. Este mismo *efecto meseta* fue constatado en los puntajes del Test de Raven en la ciudad de La Plata, Argentina, para las edades 13-18 años por este equipo de investigación (Rossi-Casé et al., 2016).

Cuando se realizaron las primeras observaciones del efecto Flynn surgieron diversas hipótesis explicativas. De todas formas, la causa de este fenómeno permanece aún desconocida y no existe evidencia suficiente que permita afirmar que este hecho refleje un verdadero aumento de la inteligencia. Entre las hipótesis propuestas se encuentran las siguientes: la heterosis, es decir, la mejora en las condiciones de vida en las distintas poblaciones y una alimentación o nutrición más adecuada; la expansión del sistema educativo; la reducción del tamaño de las familias nucleares, con el consecuente incremento del desarrollo psicológico; la adquisición progresiva de ciertas habilidades para responder satisfactoriamente a los tests; un mayor desarrollo de la escolarización y educación de los niños, y la creciente tecnificación de la cultura, desde los juegos de video hasta el acceso cada vez más irrestricto a los medios de comunicación a través de internet, que ofrecen otros aspectos de estimulación (Sundet et al., 2004; Baker et al., 2015; Passig, 2015). Investigaciones realizadas en países que han sufrido guerras podrían ser consideradas casos que sostienen estas hipótesis. Tal es el caso de Kuwait, Sudán y Siria (Dutton, Bakhiet, Alhossein, & Essa, 2019; Dutton, Bakhiet et al., 2018; Dutton, Essa et al., 2018), donde luego de conflictos militares (la Primera Guerra del Golfo en el caso de

Kuwait y la Guerra Civil en Sudán y Siria) que trajeron como consecuencia un detrimento en la calidad de vida de sus habitantes, se detectó una baja en los puntajes de los test de inteligencia. Estos descubrimientos llevaron a parte del equipo de investigación a suponer una correlación positiva entre el efecto Flynn y el crecimiento del PBI de un país. Dicha suposición fue corroborada por [Rindermann y Becker \(2018\)](#).

Por su parte, [Armstrong y Woodley \(2014\)](#) afirman que hay evidencia neurológica que señala que el efecto Flynn está asociado a un aumento en el tamaño del cerebro que ha potenciado las funciones del hipocampo, aunque no descartan que otra causa para este fenómeno sea la posibilidad de que los tests que miden la inteligencia fluida y la capacidad eductiva, como el Test de Matrices Progresivas de Raven, contengan algún tipo de *andamio cognitivo* que oriente el uso de habilidades no dependientes del factor *g* para resolver los problemas. Es decir, que el mismo test podría guiar al sujeto en la adquisición de dichas habilidades.

La observación de una detención en el aumento de los puntajes medios en los tests nos invita a una nueva consideración de las posibles causas de este fenómeno. En un intento por abordar esta cuestión, [Bratsberg y Rogeberg \(2018\)](#), a partir de datos similares a los empleados por [Sundet et al. \(2004\)](#), pusieron a prueba las hipótesis causales para el efecto Flynn y su reverso, clasificándolas en dos grupos: aquellas que apelan al origen genético de este efecto y las que postulan una causalidad ambiental. Al no encontrar diferencias significativas en el desarrollo de este efecto entre sujetos pertenecientes a la misma familia con respecto a la población general, los autores concluyen que el primer grupo de hipótesis carece de sustento suficiente, y se decantan entonces por el segundo. A través de un relevamiento realizado con 70 expertos en la materia, [Rindermann,](#)

[Becker y Coyle \(2017\)](#) arribaron a la misma conclusión.

En este punto, consideramos que las perspectivas ligadas a las dimensiones generacionales pueden aportar elementos para elucidar la presentación actual del efecto Flynn. Los sujetos que integran la muestra del presente estudio se corresponden con las denominadas Generación Y y Z, la cohorte de personas nacidas en la década de los años 80 y de los años 90, respectivamente ([Strauss & Howe, 1991](#)). [Navós \(2014\)](#) retoma la tradicional clasificación generacional según el criterio de año de nacimiento y establece los siguientes grupos: a) Tradicionalistas: nacidos antes de 1945; b) Baby boomers: nacidos entre 1945 y 1965; c) Generación X: nacidos entre 1966 y 1980; d) Generación Y: nacidos entre 1981 y 1995, los famosos *millennials*; e) Generación Z: nacidos a partir de 1996.

Ahora bien, dichos grupos suponen, según nos plantea el autor, diferencias relativas al contexto social y político en el cual nacieron y se criaron (códigos, aspiraciones, valores), diferencias que actualmente coexisten en diversos ámbitos de nuestra sociedad. En efecto, [Navós \(2014\)](#) resume las características correspondientes a los grupos generacionales de la siguiente manera:

Los *tradicionalistas* valorizan fundamentalmente la educación y el esfuerzo en tanto camino hacia el progreso; la palabra y la autoridad denotan para ellos respeto tanto en las relaciones familiares como también frente a las instituciones, jerarquías y superiores, y procuran la proyección a largo plazo durante sus vidas, etc.

El grupo generacional *baby boomers* se constituyó en un contexto de mayor incertidumbre tras el paso de las dos Guerras Mundiales. En ese sentido, bajo los efectos de las coyunturas políticas y sociales, comienzan a discutir las jerarquías, lo establecido, buscan lograr un título y profesionalizarse a los fines de alcanzar un me-

nor grado de dependencia, se oponen a los valores impuestos y apuestan a la paz y la libertad.

La *generación X* se compone de sujetos competitivos e independientes, que aceptan la diversidad y procuran contextos informales. Buscan su autodesarrollo, tomando dimensión de la inmediatez de sus vidas y, en consecuencia, acelerando sus ritmos profesionales, procurando el disfrute de otros planos sociales.

El grupo generacional Y o *millennials* atravesó un contexto signado por la globalización y el acceso a internet. Las tecnologías se presentan para ellos como inherentes a sus crianzas y prácticas cotidianas. Están siempre conectados y tienen la posibilidad de acceder a información y a realidades más amplias que las generaciones anteriores. Priorizan lo inmediato por sobre la construcción de proyectos a largo plazo. Procuran disfrutar de una buena calidad de vida, de lo que hacen, buscar nuevos horizontes, sostener un pensamiento independiente por fuera de los patrones establecidos.

Por último, la *Generación Z* se caracteriza por haber nacido en una época en donde el consumo adquiere su mayor auge y el acceso a las tecnologías se presenta desde el nacimiento. Su comunicación se encuentra predominantemente mediada por herramientas digitales. La inmediatez signa sus vidas, y priorizan el disfrute por sobre el trabajo.

De esta manera, los procesos histórico-sociales y los desarrollos tecnológicos se sintetizan en patrones de comportamiento específicos en cada grupo generacional. Con el arribo y difusión masiva de la tecnología digital a fines del siglo XX, se inaugura una discontinuidad generacional substancial, identificada por Prensky (2001) con la oposición entre nativos e inmigrantes digitales. Prensky designa *nativos digitales* a quienes han nacido y se han formado utilizando la particular “lengua digital” de juegos por ordenador, vídeo

e Internet (tales como las generaciones Y y Z) en tanto que a quienes no han nacido en ese contexto, pero que deben ajustarse en la actualidad a su uso, los denomina *inmigrantes digitales*. En ese sentido, más allá de las diferencias entre las generaciones, encontramos entre ambos grupos un modo de pensar y procesar la información significativamente diferente. Los nativos digitales se caracterizan por un acceso rápido e inmediato a la información dadas las disponibilidades que genera internet, optando por textos que disminuyan la complejidad del conocimiento y tiendan a ser intuitivos.

Sartori (1998) señala que la televisión y el conjunto de nuevas tecnologías, que se inscribieron en la crianza de quienes hoy se presentan como generación Y, altera, empobrece el aparato cognoscitivo del *homo sapiens* y menoscaba la naturaleza simbólica del hombre. Dado que la televisión ocupó un lugar privilegiado en la crianza de quienes componen la generación *millennial*, se presenta una prevalencia del ver sobre el hablar, con lo que arribamos a una sustitución que modifica profundamente la relación entre entender y ver ya que actualmente el relato (su explicación) está supeditado a las imágenes que aparecen en la pantalla. Esto, según señala Sartori, reduce la capacidad de abstracción y entendimiento. En ese sentido, el autor postula un proceso de suplantación del *homo sapiens* por el *homo videns*, en tanto en este último predomina el lenguaje perceptivo por sobre otras capacidades cognoscitivas.

Las características enumeradas por estos autores convergen en referencia a la inmediatez como elemento que signa la vida de estos sujetos. Esto remite a los desarrollos de Bauman (2002) sobre la modernidad líquida. Este autor se apoyó en la cualidad que los líquidos tienen para cambiar su forma y para desplazarse, por oposición a los sólidos, como metáfora para referirse al impacto que ha tenido en la sociedad y en el individuo la

pérdida actual de ciertos códigos y conductas que solían servir como puntos estables de orientación. Este autor asocia la cualidad de “líquido” con las ideas de levedad, movilidad e inconstancia. La forma que un líquido tenga dependerá del momento en que se lo observe, por lo que “para ellos lo que cuenta es el flujo del tiempo más que el espacio que pueden ocupar” (Bauman, 2002, p. 8). Más adelante en su obra habla del *síndrome de la impaciencia*, producto de la concepción del acceso a los atajos como emblema de privilegio en la escala social. Según el autor, la posición de cada uno en dicha escala “se mide por la capacidad (o la ineptitud) para reducir o hacer desaparecer por completo el espacio de tiempo que separa el deseo de su satisfacción” (Bauman, 2008, p. 22).

El objetivo de este artículo es presentar los resultados finales del proyecto de investigación dedicado a la construcción de normas de percentiles para el Test de Matrices Progresivas de Raven, Escala General, para las edades de 19 a 30 años en la ciudad de La Plata, Argentina. De modo complementario, y sobre la base de reflexiones introducidas anteriormente (Rossi-Casé, Doná, Garzaniti, Biganzoli, & Llanos-Barja, 2018; Rossi-Casé, Doná, Biganzoli, & Garzaniti, 2019), se elaborarán algunas consideraciones sobre las particularidades de la inteligencia en las nuevas generaciones. Las mismas se desprenden de la indagación de las diferencias entre las poblaciones en las cuales el efecto Flynn estaba aún vigente y las actuales, en las que el mismo ya no se observa.

Método

Participantes

La muestra estuvo conformada por 979 participantes de ambos sexos, 543 mujeres y 426 varones, con edades comprendidas dentro del rango de 19 a 30 años, considerando la edad en años

cumplidos al momento de la administración del test. El total de casos quedó determinado una vez que fueron desestimados los protocolos con discrepancias significativas de los puntajes directos.

Según la edad, se dividió la muestra en cuatro intervalos: 19-20 años, 21-22 años, 23-24 años y 25-30 años. La amplitud de los rangos de edad obedece al hecho constatado de que los puntajes de la prueba se incrementan con el aumento de edad, en intervalos iguales de dos años desde los 13 a los 24 años, y en intervalos de edad diferentes, a partir de los 25 años. La versión original de la prueba y las revisiones argentinas del año 1964 y 2000 presentan esta segmentación de edades, por lo que mantenerlas nos permite realizar comparaciones de los resultados alcanzados (Raven et al., 2003; Rossi-Casé et al., 2001, 2017).

Además de ser caracterizada por la edad, la muestra fue estratificada con respecto a la variable sexo, teniendo como referencia los datos del último censo poblacional para la ciudad de La Plata y sus alrededores de la República Argentina, realizado en el año 2010 (Instituto Nacional de Estadística y Censos [INDEC], 2012). Todas las personas que conformaron la muestra, habían alcanzado el nivel de escolarización media.

Instrumentos

Se utilizó el Test de Matrices Progresivas de Raven; Escala General (Raven, 2005). Es una prueba no verbal, independiente del lenguaje hablado o escrito, de la cultura o escolarización de las personas. Esta prueba mide la capacidad intelectual general mediante la comparación de formas y el razonamiento por analogía, sin intervención de los conocimientos adquiridos previamente. Brinda información sobre la capacidad intelectual, midiendo dos de sus componentes: educación, en tanto razonamiento lógico; y repro-

ducción de la información.

El test está compuesto por sesenta problemas organizados en cinco series (A; B; C; D; E) de doce ítems cada una. Cada una de las series comienza con problemas sencillos y luego paulatinamente va aumentando la complejidad. A su vez, la complejidad de las series también va en aumento, desde la serie A hasta la E. La puntuación directa máxima es de 60 puntos.

Tabla 1

Composición de la muestra.

Edad	Sexo		Total de casos por edad
	Mujer	Varón	
19-20 años	203	133	336
21-22 años	116	107	223
23-24 años	121	113	234
25-30 años	103	83	186
Total de casos por Sexo	543	436	979

Nota. N = 979.

Procedimiento

Se obtuvo la autorización previa de los participantes en la investigación, mediante un formulario de consentimiento informado, el cual incluía una explicación del propósito de la investigación, una aclaración sobre la confidencialidad de los datos, y un agradecimiento por la participación.

El test se administró de manera individual o grupal, según el caso, con presencia del examinador y sin límite de tiempo para su ejecución. Sin embargo, se observó que los participantes más rápidos completaron la prueba dentro del rango de 30 minutos, y los participantes más lentos, dentro del rango de 60 minutos.

Para la consigna, se siguieron de manera rigurosa las instrucciones dadas por el autor para la autoadministración o administración colectiva. La prueba se tomó en distintos períodos de los años 2016 a 2019: junio a noviembre de 2016, marzo a septiembre de 2017, marzo a agosto de 2018, y abril a agosto de 2019. La administración estuvo a cargo de este Equipo de Investigación. El número mayor de participantes por grupo fue de 32. La mayoría de las pruebas fueron realizadas en el espacio físico de la Facultad de Psicología de la Universidad Nacional de La Plata, lugar al que los participantes fueron citados por ser la institución a la que se vincula esta investigación.

Al conformar la base de datos, se excluyeron 58 protocolos que mostraban una discrepancia mayor a la esperada en la composición de los puntajes. Se analizó una base de datos que quedó conformada por 979 protocolos.

Procesamiento de los datos

Con la información recogida se construyó una base utilizando la planilla de cálculos Microsoft Excel (Microsoft Corporation, 2010). Luego se exportaron los datos al software InfoStat (Di Rienzo et al., 2018) para obtener los estadísticos descriptivos que permitieron elaborar las normas para cada grupo de edad.

Se definió un intervalo de confianza del 95%, y se estableció que el error muestral máximo sería del 3% para el tamaño de esta muestra. Los resultados obtenidos en el estudio actual se compararon con los baremos obtenidos en los años 1964 y 2000, para los mismos grupos etarios de la ciudad de La Plata, Argentina, a excepción de las puntuaciones directas obtenidas en el año 1964 para el rango de edad 25-30 años, que no se encuentran disponibles (Rossi-Casé et al., 2017).

La comparación de dichos rendimientos

promedio por rango de edad y por sexo se realizó aplicando el estadístico z para la prueba de hipótesis de la diferencia de medias entre muestras aleatorias, normales e independientes, con un intervalo de confianza del 95%.

Resultados

El análisis de las puntuaciones directas obtenidas muestra que en los cuatro grupos de edad se alcanzó la puntuación máxima posible (ver Tabla 2). La puntuación mínima observada fue de 34 puntos y se registró en el grupo de menor edad. Los resultados mostraron que, a medida que aumenta la edad de los participantes, el rendimiento promedio de los grupos mejora; y que la heterogeneidad de las respuestas resulta similar entre grupos.

La prueba z de diferencia de medias entre dos muestras permitió observar que entre las puntuaciones promedio obtenidas en el año 2000 y las del año 2019 no se constatan diferencias significativas. Para ningún rango de edad estudiado el estadístico z calculado pudo superar el valor crítico de z , para un nivel de significación $\alpha = .05$. Las variaciones encontradas se deberían al azar.

De igual manera, los varones obtuvieron puntuaciones levemente superiores que las mujeres en todos los rangos de edad; con un mínimo de respuestas correctas también mayor, y alcanzando siempre la mayor puntuación posible. Las variaciones encontradas entre los rendimientos de ambos sexos no fueron estadísticamente significativas ($p < .05$). Las diferencias encontradas también se deberían al azar.

Resultados para el grupo de edad de 19 a 20 años

En los participantes que integran este grupo

Tabla 2

Descripción de los promedios y variaciones de las puntuaciones directas, por rango de edades.

Rango de Edades	Promedio	Desviación estándar	Rango de Respuestas
19-20 años	50.16	4.77	34 a 60 ptos.
21-22 años	51.78	4.39	40 a 60 ptos.
23-24 años	51.51	5.52	35 a 60 ptos.
25-30 años	52.89	4.83	38 a 60 ptos.

Nota. Puntuación máxima: 60 puntos . N = 979.

se observan los resultados más bajos de todos los que integran la muestra. La puntuación promedio es de 50.16 puntos y la desviación estándar es de 4.77 puntos. El rango de respuestas correctas fue de 34 para los puntajes más bajos y de 60 para los más altos. Entre las mujeres, el rango de respuestas correctas fue de 34 a 59, y entre los varones, de 37 a 60.

El rendimiento de los varones ($M = 51.13$ puntos; $DE = 4.73$ puntos) es mejor que el de las mujeres ($M = 49.63$ puntos; $DE = 4.71$ puntos). La prueba de diferencia de medias mostró que tales variaciones no son estadísticamente significativas ($p < .05$).

La mitad de los sujetos que integran este grupo de edad ha obtenido 50 puntos o menos (ver Tabla 3). Si bien esto representa un incremento de 11 puntos respecto del baremo del año 1964 (Rossi-Casé et al., 2001), el resultado observado es 1 punto menor que la norma alcanzada en el año 2000 (Rossi-Casé et al., 2017). Este comportamiento muestra un incremento significativo de

las puntuaciones respecto de las normas 1964 ($p < .05$). En comparación con los baremos del año 2000, se constata una igualdad para los percentiles 75 o superior, y una disminución en todos los valores percentilares de la mediana e inferiores.

Tabla 3

Edades 19-20 años: Comparación de la distribución percentilar de las puntuaciones directas, años 1964, 2000 y 2019.

Percentil	Baremos		Valores
	año 1964	año 2000	percentilares año 2019
P99	53	59	59
P95	53	57	57
P90	51	56	56
P75	45	54	54
P50	40	52	51
P25	34	48	47
P10	27	46	44
P5	22	43	41

Nota. Puntuación máxima: 60 puntos. $n = 336$.

Resultados para el grupo de edad de 21-22 años

La puntuación promedio observada es de 51.78 puntos y la desviación estándar es de 4.35 puntos (ver Tabla 4). El rango de respuestas correctas fue de 40 para los puntajes más bajos y de 60 para los más altos. Entre las mujeres, el rango de respuestas correctas fue de 40 a 59; y entre los varones, de 41 a 60.

El rendimiento de los varones ($M = 51.93$ puntos; $DE = 3.99$ puntos) es levemente mejor que el de las mujeres ($M = 51.67$ puntos; $DE = 4.71$ puntos), pero no se observa una diferencia de medias estadísticamente significativa ($p < .05$).

La mitad de los participantes que integran este grupo de edad ha obtenido al menos 52 pun-

tos. Esto representa un incremento de 12 puntos respecto del baremo del año 1964 (Rossi-Casé et al., 2001), y 1 punto respecto de la norma alcanzada en el año 2000 (Rossi-Casé et al., 2017). Este incremento significativo en las puntuaciones respecto de las normas 1964 se observó en todos los valores percentilares calculados ($p < .05$). La comparación de resultados con los baremos del año 2000 demuestra que se conservan iguales los puntajes directos necesarios para alcanzar el percentil 90 y superiores; y aumentan en un punto los necesarios para alcanzar los percentiles 75 e inferiores.

Tabla 4

Edades 21-22 años: Comparación de la distribución percentilar de las puntuaciones directas, años 1964, 2000 y 2019.

Percentil	Baremos		Valores
	año 1964	año 2000	percentilares año 2019
P99	53	59	59
P95	53	58	58
P90	51	57	57
P75	45	54	55
P50	40	51	52
P25	34	48	49
P10	27	45	46
P5	22	42	43

Nota. Puntuación máxima: 60 puntos. $n = 223$.

Resultados para el grupo de edad de 23 a 24 años

En los participantes que integran este grupo se observan los resultados con mayor variabilidad de todos los que integran la muestra. La puntuación promedio es de 51.51 puntos y la desviación estándar de 5.52 puntos (ver Tabla 5). El rango de respuestas correctas fue de entre 35 y 59 en mujeres; y entre 37 y 60 entre varones.

El rendimiento de los varones ($M = 53.76$ puntos; $DE = 5.55$ puntos) es mejor que el de las mujeres ($M = 50.18$ puntos; $DE = 5.09$ puntos). La diferencia de medias muestra que los rendimientos promedio no son estadísticamente significativos y que tales variaciones se deberían al azar ($p < .05$).

La mitad de los sujetos que integran el grupo de edad ha obtenido 53 puntos o menos. El resultado observado representa un incremento de 13 puntos respecto del baremo del año 1964 (Rossi-Casé et al., 2001), al tiempo que permanece igual respecto de la norma alcanzada en el año 2000 (Rossi-Casé et al., 2017). Este incremento significativo de las puntuaciones respecto de las normas 1964 ($p < .05$) se observó en todos los valores percentilares calculados. Respecto de los baremos del año 2000, no hay aumento de las puntuaciones directas sino que se constata la igualdad o, incluso, la disminución de 1 y 2 puntos para los valores de los percentiles 99 y 25, respectivamente.

Resultados para el grupo de edad de 25 a 30 años

En los sujetos que integran este grupo se observan los resultados más altos de todos los que integran la muestra. La puntuación promedio es de 52.89 puntos, y la desviación estándar de 4.83 puntos (ver Tabla 6). El rango de respuestas correctas fue de 38 para mujeres y 42 en varones; hasta los 60 puntos para los más altos, valor alcanzado en ambos sexos.

La mitad de los sujetos que integran el grupo de edad ha obtenido 54 puntos o menos. El resultado observado representa un incremento de 4 puntos respecto del baremo del año 2000 (Rossi-Casé et al., 2017).

El rendimiento de los varones ($M = 55.07$ puntos; $DE = 4.84$ puntos) es mejor que el de las

Tabla 5

Edades 23-24 años: Comparación de la distribución percentilar de las puntuaciones directas, años 1964, 2000 y 2019.

Percentil	Baremos		Valores
	año 1964	año 2000	percentilares
			año 2019
P99	53	60	59
P95	53	59	59
P90	51	58	58
P75	45	55	56
P50	40	53	53
P25	34	49	47
P10	27	44	44
P5	22	42	42

Nota. Puntuación máxima: 60 puntos. $n = 234$.

mujeres ($M = 51.76$ puntos; $DE = 4.44$ puntos), aunque la diferencia de medias no es estadísticamente significativa ($p < .05$).

Respecto de dichos baremos, se constata el aumento de al menos 1 punto en todos los percentiles más altos; y de 4 ó 5 puntos para los valores percentilares medio y más bajos.

Baremos

La Tabla 7 presenta los baremos 2019 del Test de Raven, Escala General, para las edades de 19 a 30 años, para la ciudad de La Plata, Argentina.

Discusión

En las casi dos décadas transcurridas desde la elaboración de las normas anteriores del test hasta la actualidad, los resultados obtenidos

Tabla 6

Edades 25-30 años: Comparación de la distribución percentilar de las puntuaciones directas, años 2000 y 2019.

Percentil	Baremos	Valores percentilares
	año 2000	año 2019
P99	59	60
P95	58	59
P90	57	58
P75	55	57
P50	50	54
P25	45	50
P10	42	47
P5	39	43

Nota. Puntuación máxima: 60 puntos. n = 186.

permiten mostrar, en consonancia con investigaciones realizadas en otros países (Bratsberg & Rogeberg, 2018; Brouwers, van de Vijver, & van Hemert, 2009; Flynn, 2013; Sundet et al., 2004; Teasdale & Owen, 2007), una detención en el aumento de los puntajes directos necesarios para alcanzar el rendimiento promedio en cada rango de

edad ($p < .05$). Las variaciones encontradas entre los valores actuales y los baremos del año 2000 (Rossi-Casé et al., 2017) no resultaron ser estadísticamente significativas, por lo que se constata la detención del efecto Flynn, característico de la segunda mitad del siglo XX (ver Tabla 8).

Los autores antes citados han mencionado como posible causa de este fenómeno un *efecto techo*. Esto se podría explicar a partir de los mismos factores que se tuvieron en cuenta para intentar explicar el anterior aumento: heterosis, nutrición más adecuada, expansión del sistema educativo. Para fundamentarlo empíricamente, Brouwers et al. (2009), argumentaron que en los países con menor PBI per cápita los puntajes promedio seguían aumentando, mientras que el efecto techo podía observarse en los países escandinavos. De la misma forma, Teasdale y Owen (2007) afirmaron que los países en vías de desarrollo estarían lejos de llegar al efecto techo.

Sin embargo, la situación de Argentina representa un ejemplo del estancamiento de los puntajes promedio en un país no tan desarrollado y cuyos indicadores sociales están lejos de aquellos de los países escandinavos. Consideramos pertinente, en este punto, aclarar que *efecto te-*

Tabla 7

Edades 19-30 años: Escala percentilar y puntuaciones directas, año 2019, ambos sexos, para cada grupo etario.

Percentil	Edad			
	19-20 años	21-22 años	23-24 años	25-30 años
P99	59	59	59	60
P95	57	58	59	59
P90	56	57	58	58
P75	54	55	56	57
P50	51	52	53	54
P25	47	49	47	50
P10	44	46	44	47
P5	41	43	42	43

Nota. Puntuación máxima: 60 puntos. N = 979.

cho y *efecto meseta* no son sinónimos. El primero aparece cuando se han alcanzado los puntajes más altos que era posible obtener en la prueba, por lo que los mismos ya no pueden continuar aumentando. En cambio, el segundo se da cuando se produce un estancamiento en los puntajes sin haber alcanzado el máximo posible. Hecha esta diferenciación, observamos que mientras que lo descrito en los países escandinavos fue caracterizado como efecto techo, el caso argentino muestra más bien un efecto meseta.

El hecho de que la recolección de datos del presente estudio se haya realizado a lo largo de tres años podría acarrear el riesgo de la introducción de un sesgo indeseado en el análisis de los resultados, si se sostuviera la vigencia local del efecto Flynn. No obstante, el hecho de que aún en estos años los resultados no hayan mostrado diferencias significativas con respecto a aquellos de la estandarización anterior del Test, refuerza la constatación del efecto meseta, lejos de invalidarla.

Por otro lado, si bien el PBI per cápita es menor en Argentina que en los países donde se observó el efecto techo, la globalización ha logrado que las nuevas generaciones en el mundo oc-

cidental accedan a los mismos medios de comunicación y redes sociales. Proponemos entonces retomar el concepto de *homo videns*, postulado por Sartori (1998), especialmente en lo atinente al empobrecimiento del aparato cognitivo, y considerar que, además, para estas nuevas generaciones la inmediatez es una virtud. La administración del Test de Raven puede tomar unos 40 minutos o más (requiriendo bastante esfuerzo cognitivo en las últimas series), lo cual podría generar un desgano a la hora de concluir con el test que explique el amesetamiento de los puntajes en el mismo.

Consideramos que una línea a indagar es si se encuentra una explicación posible del amesetamiento del efecto Flynn a partir de postulados de Piscitelli (2006), Sartori (1998) y Prensky (2001), entre otros, sobre los grupos generacionales y la discontinuidad substancial para los grupos Y y Z, ya que suponen cambios cognitivos signados por la inmediatez, como la disminución de las capacidades de abstracción y entendimiento, y nuevos modos de pensar y procesar la información que motivarían el desgano de estas generaciones frente al Test de Raven en tanto el mismo les demanda un esfuerzo cognitivo superior al que emplean en la actualidad, lo cual se manifestaría en los resul-

Tabla 8

Edades 19-30 años: Comparación de la distribución percentilar de las puntuaciones directas, años 2000 y 2019.

Percentil	Edad							
	19-20 años		21-22 años		23-24 años		25-30 años	
	año 2000	año 2019	año 2000	año 2019	año 2000	año 2019	año 2000	año 2019
P99	59	59	59	59	60	59	59	60
P95	57	57	58	58	59	59	58	59
P90	56	56	57	57	58	58	57	58
P75	54	54	54	55	55	56	55	57
P50	52	51	51	52	53	53	50	54
P25	48	47	48	49	49	47	45	50
P10	46	44	45	46	44	44	42	47
P5	43	41	42	43	42	42	39	43

Nota. Puntuación máxima: 60 puntos. N = 979.

tados obtenidos.

Considerando que el presente estudio analiza los resultados para el grupo etario 19-30 años, es decir, quienes nacieron entre 1988 y 1999, en este trabajo nos ocupamos de las generaciones Y, o *millennials*, y Z, o *centennials*. La generación Y o *millennials* es la primera integrada por nativos digitales. Se constituye sustancialmente a partir de la aparición de internet y la globalización, expresándose en una conectividad a escala global y cuya velocidad equivale a la inmediatez. Todos son nativos digitales, y la brecha generacional con los grupos anteriores no se circunscribe exclusivamente a cuestiones del orden social, histórico, tecnológico, simbólico específicas de cada grupo, sino que también supone una brecha cognitiva. En ese sentido, las modificaciones en las capacidades cognitivas que postulan dichos autores, signadas por la inmediatez, podrían contribuir al detenimiento del efecto Flynn. Pensemos que estos tests son tomados en el contexto áulico, para los menores de 18 años, y en universidades e institutos terciarios para los mayores. Con esto queremos señalar que quienes realizan la prueba no lo hacen condicionados por el futuro resultado, como podría pasar en una entrevista laboral, por ejemplo. Esto nos hace pensar que quienes están resolviendo la prueba podrían cuestionar la motivación para llevar adelante el esfuerzo que conlleva esta tarea. Tal vez aparezcan dificultades para concentrarse durante varios minutos seguidos en una misma tarea, sumándose a ello las ansiedades por obtener un resultado de manera inmediata. Estas condiciones pueden caracterizar a los sujetos que son evaluados con el Test de Raven, y reflejarse directamente sobre los resultados finales.

Posibles líneas de investigación futura podrían incluir la indagación de si existe un correlato, en otros aspectos de la inteligencia ligados a la motricidad y el lenguaje, de la mencionada

diferencia cognitiva intergeneracional observada a nivel de la capacidad educativa.

Referencias

- Armstrong, E. L., & Woodley, M. A. (2014). The rule-dependence model explains the commonalities between the Flynn effect and IQ gains via retesting. *Learning and Individual Differences, 29*, 41-49. doi: [10.1016/j.lindif.2013.10.009](https://doi.org/10.1016/j.lindif.2013.10.009)
- Baker, D. P., Eslinger, P. J., Benavides, M., Peters, E., Dieckmann, N. F., & Leon, J. (2015). The cognitive impact of the education revolution: A possible cause of the Flynn effect on population IQ. *Intelligence, 49*, 144-158. doi: [10.1016/j.intell.2015.01.003](https://doi.org/10.1016/j.intell.2015.01.003)
- Bauman, Z. (2002). *Modernidad líquida*. Buenos Aires, Argentina: Fondo de Cultura Económica.
- Bauman, Z. (2008). *Los retos de la educación en la modernidad líquida*. Barcelona, España: Gedisa.
- Bratsberg, B., & Rogeberg, O. (2018). Flynn effect and its reversal are both environmentally caused. *Proceedings of the National Academy of Sciences of the United States of America, 115*(26), 6674-6678. doi: [10.1073/pnas.1718793115](https://doi.org/10.1073/pnas.1718793115)
- Brouwers, S. A., van de Vijver, F. J., & van Hemert, D. A. (2009). Variation in Raven's Progressive Matrices scores across time and place. *Learning and Individual Differences, 19*(3), 330-338. doi: [10.1016/j.lindif.2008.10.006](https://doi.org/10.1016/j.lindif.2008.10.006)
- Di-Rienzo, J. A., Casanoves, F., Balzarini, M. G., Gonzalez, L., Tablada, M., & Robledo, C. W. (2018). InfoStat (versión 2018) [software de cómputo]. Recuperado de <http://www.infostat.com.ar>
- Dutton, E., Bakhiet, S. F. A., Alhossein, A. H., & Essa, Y. A. S. (2019). A Flynn effect in Kuwait, 1985-1998. *Personality and Individual Differences, 138*, 355-357. doi: [10.1016/j.paid.2018.10.029](https://doi.org/10.1016/j.paid.2018.10.029)
- Dutton, E., Bakhiet, S. F. A., Osman, H. A., Becker, D., Essa, Y. A. S., Blahmar, T. A. M., & Hakami, S. M. (2018). A Flynn effect in Khartoum, the Sudanese

- capital, 2004-2016. *Intelligence*, 68, 82-86. doi: [10.1016/j.intell.2018.03.007](https://doi.org/10.1016/j.intell.2018.03.007)
- Dutton, E., Essa, Y. A. S., Bakhiet, S. F., Ali, H. A. A., Alqafari, S. M., Alfaleh, A. S. H., & Becker, D. (2018). Brain drain in Syria's ancient capital: No Flynn effect in Damascus, 2004-2013/14. *Personality and Individual Differences*, 125, 10-13. doi: [10.1016/j.paid.2017.12.025](https://doi.org/10.1016/j.paid.2017.12.025)
- Fernández-Liporace, M., Ongarato, P., Saavedra, E., & Casullo, M. M. (2004). El Test de Matrices Progresivas, Escala General: Un análisis psicométrico. *Revista Evaluar*, 4, 50-69. Recuperado de <https://revistas.unc.edu.ar/index.php/revaluar/index>
- Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932-1978. *Psychological Bulletin*, 95(1), 29-51. doi: [10.1037/0033-2909.95.1.29](https://doi.org/10.1037/0033-2909.95.1.29)
- Flynn, J. R. (2013). The "Flynn effect" and Flynn's paradox. *Intelligence*, 41(6), 851-857. doi: [10.1016/j.intell.2013.06.014](https://doi.org/10.1016/j.intell.2013.06.014)
- Flynn, J. R., & Rossi-Casé, L. (2011). Modern women match men on Raven's Progressive Matrices. *Personality and Individual Differences*, 50(6), 799-803. doi: [10.1016/j.paid.2010.12.035](https://doi.org/10.1016/j.paid.2010.12.035)
- Flynn, J. R., & Rossi-Casé, L. (2012). IQ gains in Argentina between 1964 and 1998. *Intelligence*, 40(2), 145-150. doi: [10.1016/j.intell.2012.01.006](https://doi.org/10.1016/j.intell.2012.01.006)
- Gignac, G. E. (2015). Raven's is not a pure measure of general intelligence: Implications for g factor theory and the brief measurement of g. *Intelligence*, 52, 71-79. doi: [10.1016/j.intell.2015.07.006](https://doi.org/10.1016/j.intell.2015.07.006)
- Hartshorne, J. K., & Germine, L. T. (2015). When does cognitive functioning peak? The asynchronous rise and fall of different cognitive abilities across the life span. *Psychological Science*, 26(4), 433-443. doi: [10.1177/0956797614567339](https://doi.org/10.1177/0956797614567339)
- Hayes, T. R., Petrov, A. A., & Sederberg, P. B. (2015). Do we really become smarter when our fluid-intelligence test scores improve? *Intelligence*, 48, 1-14. doi: [10.1016/j.intell.2014.10.005](https://doi.org/10.1016/j.intell.2014.10.005)
- Instituto Nacional de Estadística y Censos. (2012). *Censo Nacional de Población, Hogares y Viviendas 2010*. Censo del Bicentenario. Resultados definitivos. Serie B N° 2. Tomo 1. Buenos Aires, Argentina: Autor. Recuperado de <https://www.indec.gov.ar>
- Microsoft Corporation. (2010). Microsoft Office Excel (Versión 2010) [software]. Redmond, WA: Microsoft.
- Navós, O. T. (2014). *Nuevas generaciones en universidades privadas: ¿Qué hacer? Algunas propuestas para la gestión desde el punto de vista del marketing*. Buenos Aires: Dunker.
- Passig, D. (2015). Revisiting the Flynn effect through 3D immersive virtual reality (IVR). *Computers & Education*, 88, 327-342. doi: [10.1016/j.compedu.2015.05.008](https://doi.org/10.1016/j.compedu.2015.05.008)
- Piscitelli, A. (2006). Nativos e inmigrantes digitales. ¿Brecha generacional, brecha cognitiva, o las dos juntas y más aún? *Revista Mexicana de Investigación Educativa*, 11(28), 179-185. Recuperado de <https://www.comie.org.mx/revista/v2018/rmie/index.php/nrmie>
- Prensky, M. (2001). Digital natives, digital immigrants part 1. *On the Horizon*, 9(5), 1-6. doi: [10.1108/10748120110424816](https://doi.org/10.1108/10748120110424816)
- Raven, J. C. (2005). *Test de Matrices Progresivas, Escala General, Carpeta de Evaluación*. Buenos Aires, Argentina: Paidós.
- Raven, J. C., Raven, J., & Court, J. H. (2003). *Test de Matrices Progresivas: Escala General*. Buenos Aires, Argentina: Paidós." *Test de Matrices Progresivas: Escala General*. Buenos Aires, Argentina: Paidós." *Test de Matrices Progresivas: Escala General*. Buenos Aires, Argentina: Paidós.
- Rindermann, H., & Becker, R. (2018). Flynn-effect and economic growth: Do national increases in intelligence lead to increases in GDP? *Intelligence*, 69, 87-93. doi: [10.1016/j.intell.2018.05.001](https://doi.org/10.1016/j.intell.2018.05.001)
- Rindermann, H., Becker, D., & Coyle, T. R. (2017). Survey of expert opinion on intelligence: The Flynn effect and the future of intelligence. *Personality and Individual Differences*, 106, 242-247. doi: [10.1016/j.paid.2016.10.061](https://doi.org/10.1016/j.paid.2016.10.061)
- Rossi-Casé, L., Doná, S. M., Biganzoli, B., & Garzaniti, R.

- (2019). Evaluando a los Millennials. Apreciaciones sobre la inteligencia a partir del Test de Raven. *Perspectivas en Psicología*, 16(1), 14-25. Recuperado de <http://200.0.183.216/revista/index.php/pep/issue/view/24>
- Rossi-Casé, L., Doná, S. M., Garzaniti, R., Biganzoli, B., & Llanos-Barja, C. (2018). La inteligencia a través de las generaciones: Millennials y centennials. *Acta de Investigación Psicológica*, 8(2), 90-100. Recuperado de <http://www.psicologia.unam.mx/acta-de-investigacion-psicologica>
- Rossi-Casé, L., Neer, R., & Lopetegui, S. (2001). Test de Matrices Progresivas de Raven: Comparación de baremos. El aumento de los puntajes directos a través del tiempo. *Revista Evaluar*, 1, 39-51.
- Rossi-Casé, L., Neer, R., & Lopetegui, S. (2002). Test de Matrices Progresivas de Raven: Construcción de baremos y constatación del “efecto Flynn”. *Orientación y Sociedad*, 3, 181-187. Recuperado de http://www.psico.unlp.edu.ar/articulo/2015/11/16/revista_orientacion_y_sociedad
- Rossi-Casé, L., Neer, R., & Lopetegui, S. (2017). Normas de la ciudad de La Plata, Argentina (2000 y 2003-2004). En J. C. Raven (Ed.), *Test de Matrices Progresivas. Carpeta de Evaluación, Escala General, Normas* (pp. 21-25). Buenos Aires, Argentina: Paidós.
- Rossi-Casé, L., Neer, R., Lopetegui, S., Doná, S. M., Biganzoli, B., & Garzaniti, R. (2014). Matrices Progresivas de Raven: Efecto Flynn y actualización de baremos. *Revista de Psicología*, 23(2), 3-13. doi: [10.5354/0719-0581.2014.36144](https://doi.org/10.5354/0719-0581.2014.36144)
- Rossi-Casé, L., Neer, R., Lopetegui, S., Doná, S., Biganzoli, B., & Garzaniti, R. (2016). Test de Raven: Actualización de baremos en adolescentes argentinos y análisis del efecto Flynn. *Revista Iberoamericana de Diagnóstico y Evaluación - e Avaliação Psicológica*, 42(2), 3-13. doi: [10.21865/RIDEP42_3](https://doi.org/10.21865/RIDEP42_3)
- Sartori, G. (1998). *Homovidens. La sociedad teledirigida*. Buenos Aires, Argentina: Taurus.
- Strauss, W., & Howe, N. (1991). *Generations. The History of America's Future, 1584 to 2069*. New York, NY: William Morrow.
- Sundet, J. M., Barlaug, D. G., & Torjussen, T. M. (2004). The end of the Flynn Effect? A study of secular trends in mean intelligence test scores of Norwegian conscripts during half a century. *Intelligence*, 32(4), 349-362. doi: [10.1016/s0160-2896\(04\)00052-2](https://doi.org/10.1016/s0160-2896(04)00052-2)
- Teasdale, T. W., & Owen, D. R. (2007). Secular declines in cognitive test scores: A reversal of the Flynn effect. *Intelligence*, 36(2), 121-126. doi: [10.1016/j.intell.2007.01.007](https://doi.org/10.1016/j.intell.2007.01.007)
- Wechsler, D. (2008). *Wechsler Adult Intelligence Scale* (4^a ed.). San Antonio, TX: Pearson Assessment. doi: [10.1037/t15169-000](https://doi.org/10.1037/t15169-000)

Desarrollo y validación de la Escala de Distancia Social hacia Personas Trans

Development and Validation of the Social Distance towards Trans Persons Scale

Caleb Esteban ^{* 1}, Juan Aníbal González-Rivera ², Margarita Francia-Martínez ³,
Zahira Lespier ⁴

Introducción
Método
Resultados
Discusión
Referencias

1 - Ponce Health Sciences University, Ponce, Puerto Rico.

2 - Ponce Health Sciences University, San Juan University Center, Puerto Rico.

3 - Universidad Carlos Albizu, San Juan, Puerto Rico.

4 - Administración de Servicios de Salud Mental y Contra la Adicción, San Juan, Puerto Rico.

Recibido: 17/01/2019 Revisado: 25/02/2019 Aceptado: 10/03/2019

Resumen

El presente estudio analiza las propiedades psicométricas de la Escala de Distancia Social hacia Personas Trans (DS-T) en una muestra de profesionales y estudiantes graduados en psicología. Además, se analizó si existen diferencias significativas en la distancia social hacia la comunidad trans entre estudiantes y profesionales de la psicología. La muestra contó con 127 participantes reclutados por disponibilidad. Los resultados confirman que la escala posee una estructura unidimensional. Los 13 ítems desarrollados cumplieron con los criterios de discriminación y obtuvieron cargas factoriales apropiadas. El índice alfa de Cronbach de la DS-T fue de .99. No se encontraron diferencias significativas en la distancia social hacia personas trans entre estudiantes y profesionales de la psicología. Se determinó cómo la distancia y la cercanía constituyen extremos de la empatía, por lo que cuanto mayor sea la distancia hacia una persona o comunidad, menor será la probabilidad de desarrollar empatía hacia estas personas.

Palabras clave: distancia social, transgénero, transexual, género, validación

Abstract

The present study analyzes the psychometric properties of the Social Distance Scale toward Trans Persons (DS-T) in a sample of psychology professionals and graduate students. In addition, it analyzes whether there are significant differences in the social distance to the trans community between psychology students and professionals. The sample had 127 participants, who were recruited based on their availability. The results confirm that the scale has a one-dimensional structure. The 13 items developed by the researchers met the criteria of discrimination and showed appropriate factor loads. The Cronbach alpha internal consistency index for DS-T was .99. There were no significant differences in social distance towards trans people between psychology students and professionals. It was determined how distance and closeness constitute the extremes of empathy, so the larger the distance from a person or community, the lower the probability of being empathetic toward them.

Keywords: social distance, transgender, transsexual, gender, validation

* **Correspondencia a:** Dr. Caleb Esteban, Ph.D., Clinical Psychology Program, School of Behavioral and Brain Sciences, Ponce Health Sciences University, PO BOX 7004, Ponce, Puerto Rico 00732-7004. Email: cesteban@psm.edu. ORCID: <https://orcid.org/0000-0002-0960-6311>

Cómo citar este artículo: Esteban, C., González-Rivera, J. A., Francia-Martínez, M., & Lespier, Z. (2020). Desarrollo y validación de la Escala de Distancia Social hacia Personas Trans. *Revista Evaluar*, 20(1), 32-48. Recuperado de <https://revistas.unc.edu.ar/index.php/revaluar>

Introducción

Las sociedades y las culturas actuales son, en su mayoría, cisnormalistas (American Psychological Association [APA], 2015; Coleman et al., 2012; Ramos-Pibernus, Rodríguez-Madera, Ramos, Padilla, & Varas-Díaz, 2016). Esto implica que existen relaciones sociopolíticas de poder que normalizan y reglamentan el género y el sexo biológico de cada persona. Se tiene una visión dicotoma inflexible, en donde se ignora completamente la diversidad y se asume que toda persona es cisgénero y cissexual. Es decir, que toda persona está conforme con los constructos sociales del género que se le inculcó por su sexo biológico y que toda persona está conforme con los genitales y características sexuales con los que nació (Asociación de Psicología de Puerto Rico [APPR], 2014; Rodríguez-Madera, 2012; LGBT Advisory Committee & San Francisco Human Right Commission, 2011).

A consecuencia de estas actitudes y creencias negativas, se producen culturas transnegativistas. En estas culturas se crea un estigma que culmina en el repudio, prejuicio y discriminación hacia aquellas personas cuyo sexo y/o su género no está conforme con los constructos polarizados de hombre-mujer o masculino-femenino (APA, 2015; APPR, 2014; Ramos-Pibernus et al., 2016). La cultura crea entonces una división en la cual se observa a las personas diversas en sus identidades de sexo y de género como «los otros». *Los otros* pasan a ser un grupo de personas que, dado que sus identidades no se alinean con lo socialmente establecido, comienzan a adquirir características usualmente negativas y estereotipos que se convierten en esquemas conflictivos para muchas personas. Estos esquemas crean una intolerancia a todo hombre con masculinidades diversas y a aquel que traspase a cualquier nivel de la feminidad. Igualmente, a toda mujer con feminidades

diversas y a toda aquella que traspase a lo masculino. No obstante, este ideal y estos límites de lo femenino y lo masculino son mayormente ambiguos y, como todo constructo social, cambian de acuerdo a la cultura, el tiempo y el contexto (APA, 2012).

A pesar de tal panorama, los humanos son sumamente diversos y todos en su mayoría transgreden los constructos del género de alguna u otra manera. Como ejemplo de esto podemos mencionar instancias que van desde el hombre que cocina, hasta la mujer que trabaja en construcción. Sin embargo, a pesar de que estamos conscientes de la señalada transgresión, la sociedad define los límites y estos están altamente influenciados por controles sociopolíticos, ideales religiosos y posturas conservadoras (Francia-Martínez, Esteban, & Lespier, 2017; Horn, 2013). Por lo tanto, solo se ven perjudicadas aquellas personas que no están dentro de esos límites socialmente establecidos (APPR, 2014; Rodríguez-Madera, 2009).

El panorama es más nocivo y perjudicial cuando hablamos de las personas cuya «transgresión» no pretende ser pasajera, sino que de cierta manera deciden renunciar a los constructos establecidos para transitarlos, combinarlos o simplemente alejarse de ellos. Este es el caso de la comunidad trans, una comunidad compuesta por personas no-cisgénero y/o no-cissexuales. En otras palabras, personas que no se encuentran conformes con el género que se les asignó al nacer y/o que manifiestan inconformidad parcial o total con el sexo de nacimiento. Este grupo de personas son mejor conocidas como personas trans, término que conglomerar a las personas transgénero, transsexuales, andróginas, agénero, bigénero, entre otras (APA, 2015; APPR, 2014).

Investigaciones recientes han encontrado que las personas trans han sido mucho más estigmatizadas, discriminadas y marginadas que las personas gays, lesbianas y bisexuales. De hecho,

estudios han encontrado que la comunidad trans ha sido también la más invisibilizada y menos estudiada a nivel científico dentro la comunidad de personas lesbianas, gays, bisexuales y trans (Coleman et al., 2012; McCann & Sharek, 2014; Rodríguez-Madera, 2012). Este dato es alarmante, sobre todo si tomamos en consideración que dada la naturaleza de sus necesidades físicas y psicológicas son un sector que amerita atención multidisciplinaria. Se ha encontrado que esta comunidad presenta múltiples determinantes sociales que impactan su salud, especialmente en lo que a su salud mental se refiere (Rodríguez-Madera et al., 2016). Francia-Martínez et al. (2017), encontraron en una muestra de psicoterapeutas que, aunque presentaron en promedio actitudes positivas hacia la comunidad trans, informaron a su vez un grado moderado de distancia social. Es decir, por un lado informan no tener actitudes negativas hacia esta población, pero por otro lado informan que no vivirían con una persona trans bajo el mismo techo. Es en estas medidas incongruentes en donde mejor se puede observar el prejuicio consciente versus el inconsciente.

Marco teórico

Según Wark y Galliher (2007) Robert Park presentó a Emory Bogardus el concepto de distancia social, lo cual lo inspiró a construir la primera la escala que medía distancia social en 1926 con la discriminación por raza como motivo principal. Históricamente, el término *distancia social* se ha utilizado para reducir a categorías medibles los grados de comprensión e intimidad que caracterizan las relaciones personales y sociales en general (Wark & Galliher, 2007). También, se ha definido como el grado de intimidad y comprensión que existe entre individuos o grupos

sociales (Hughes, Johnson, Masuoka, Redfield, & Wirth, 1950). En este sentido, se puede definir distancia social como una medida de separación social entre grupos causada por diferencias reales o percibidas entre grupos de personas; estas diferencias suelen ser categorías socialmente establecidas por grupos mayoritarios que suelen percibirse como la norma. Consistentemente, la literatura científica vincula la distancia social con el prejuicio, entendiendo este último como la disposición más o menos instintiva y espontánea a mantener distancias sociales con otros grupos o individuos con ciertas características o atributos (Choi, Lee, Lee, & Kim, 2017; Maurer & Keim, 2018; Wark & Galliher, 2007; Weaver, 2008).

La distancia social es además un trecho que se forma entre dos grupos por sus diferencias, usualmente marcado por estigma, estereotipos y/o prejuicios. Esta distancia entre los grupos produce una falta de contacto suficiente o prolongado entre estos. Esto propicia el desarrollo de la separación de los mismos, pudiendo culminar en actos de violencia y de agresión, tales como microagresiones, discriminación, marginación, violencia sexual, entre otros. Por lo tanto, la distancia social es una variable de medición de los vínculos de unión o segregación de grupos sociales y mide el conflicto que se genera en estas relaciones (Arteaga-Botello & Lara-Carmona, 2004).

Según Navalles-Gómez (2011, p. 173), “la distancia y la cercanía se develan como los extremos humanos de la empatía”, implicando que cuando existe distancia hacia algún grupo se recrea un espacio de pugna, recalcando un ideal, un discurso, un sentimiento que, al final, expone una distinción. Existe evidencia de que se puede tener un bajo nivel de prejuicio hacia algún grupo en particular y, no obstante, mostrar una conducta prejuiciada (Devine, Monteith, Zuwe-rink, & Elliot, 1991), creando así una disonancia cognitiva entre cognición-conducta. Por ejemplo,

mostrar expresiones favorables hacia las personas LGB en lo cotidiano, pero hacer chistes ridiculizando la homosexualidad en el lugar de trabajo. Según [Fernández-Rodríguez y Calderón-Squiabro \(2014, p. 53\)](#), "...parece haber una discrepancia en lo que se piensa y cómo se actúa cuando se trata de los prejuicios enraizados social y culturalmente".

En Puerto Rico, el constructo de distancia social se ha utilizado mayormente para medir distancia con respecto a minorías sexuales y de género. [García \(1984\)](#), comparó los prejuicios y la distancia social hacia los hombres gay y las mujeres lesbianas entre diferentes facultades universitarias, y encontró distancia social en todas ellas, pero más distancia en la facultad de Ciencias Naturales. Estos resultados implican que la distancia social hacia una comunidad específica podría depender del tipo de educación que recibe la persona o del área de estudio en que se desempeña. [Toro-Alfonso y Varas-Díaz \(2004\)](#), [Fernández-Rodríguez y Calderón-Squiabro \(2014\)](#), y [Nieves-Rosa \(2012\)](#) repitieron este estudio en diversas universidades públicas del país y encontraron resultados similares. Además, encontraron que esta distancia era mucho mayor en los hombres y menor en aquellas personas que conocían a una persona gay o lesbiana.

Desde la disciplina de la psicología, [Vázquez-Rivera, Nazario-Serrano y Sayers-Montalvo \(2012\)](#) y [Vázquez-Rivera, Esteban y Toro-Alfonso \(2018\)](#) encontraron algunas actitudes negativas y niveles moderados de distancia social por parte de profesionales de la psicología clínica y estudiantes de dicha disciplina. [Francia-Martínez et al. \(2017\)](#), encontraron también algunas actitudes negativas y distancia social moderada por parte de estudiantes y profesionales de psicología hacia la comunidad trans. A su vez, [Rodríguez-Polo et al. \(2018\)](#) hallaron bajos niveles de actitudes negativas y baja distancia social por parte de per-

sonas empleadas en la Isla hacia personas gay y lesbianas. Por otro lado, desde la disciplina de la Salud Pública, [González-Guzman \(1998\)](#), también midió actitudes de prejuicio y distancia social hacia personas gays y lesbianas, y encontró mayores niveles de distancia social en estudiantes de segundo y tercer año de dicha concentración académica.

Propósito del estudio

En la actualidad no existen escalas para medir la distancia social hacia las personas trans. La variable de distancia social fue utilizada inicialmente para medir la distancia hacia personas con trastornos mentales ([Grandón-Fernández, Cova-Solar, Saldivia-Bórquez, & Bustos-Navarrete, 2015](#); [Link, Cullen, Frank, & Wozniak, 1987](#)). Recientemente, se ha utilizado para medir la distancia hacia personas gays y lesbianas ([Coppari et al., 2014](#); [Fernández-Rodríguez & Calderón-Squiabro, 2014](#); [García, 1984](#); [Rodríguez-Polo et al., 2018](#); [Toro-Alfonso & Varas-Díaz, 2004](#); [Vázquez-Rivera et al., 2018](#)). Tomando esto en consideración, el presente trabajo tuvo como propósito desarrollar y validar una escala para medir distancia social hacia la comunidad trans en una muestra de profesionales y estudiantes graduados de psicología. Con este fin, se examinó la validez de constructo del instrumento y se evaluó la confiabilidad del instrumento para corroborar si los ítems son consistentes en la medida del constructo distancia social hacia las personas trans. Como objetivo secundario, se evaluó si existían diferencias en la distancia social entre estudiantes y profesionales de la psicología.

Método

Diseño de investigación

Esta investigación fue de tipo cuantitativo no-experimental, de alcance exploratorio, con un diseño de estudio instrumental (Ato, López-García, & Benavente, 2013; Montero & León, 2007). En ella se desarrolló y validó preliminarmente la Escala de Distancia Social hacia Personas Trans (DS-T) con una muestra de profesionales y estudiantes de psicología.

Participantes

Se trabajó con una muestra no probabilística constituida por 127 participantes (68 profesionales de la psicología y 59 estudiantes graduados de psicología), que decidieron confidencial y voluntariamente participar del estudio. La edad promedio de la muestra fue de 37.17 años (DE = 12.23). En la Tabla 1 se desglosa el resto de los datos sociodemográficos de los participantes. Los criterios de inclusión para ambos grupos fueron: (a) ser mayor de 21 años de edad y (b) tener experiencia o práctica supervisada en intervenciones de psicoterapia y/o evaluaciones psicológicas.

Instrumentos

Cuestionario de datos generales. Para identificar las características sociodemográficas de la muestra, se desarrolló un cuestionario de datos generales que recogía información importante en cuanto a la edad, el sexo, el género, el estado civil, la orientación sexual y la afiliación religiosa.

Escala de Distancia Social hacia Personas Trans (DS-T). Este instrumento fue construido por los

Tabla 1

Datos sociodemográficos de la muestra.

Variable	f	%
Sexo		
Hombre	40	31.5
Mujer	87	68.5
Género		
Masculino	40	31.5
Femenino	86	67.7
Queer	1	0.8
Orientación sexual		
Heterosexual	100	78.7
Homosexual	19	15.0
Bisexual	5	3.9
Pansexual	1	0.8
Valores perdidos	2	1.6
Estado civil		
Solteros	45	35.4
Casados	33	26.0
Convivencia	18	14.2
Divorciados	16	12.6
Noviazgo	13	10.2
Viudos	2	1.6
Religión		
Católica	59	46.5
Protestante	6	4.7
Evangélica	13	10.2
Otra	12	9.4
Ninguna	35	27.6
Valores perdidos	2	1.6
Rama en la psicología		
Psicología clínica	46	36.2
Consejería psicológica	18	14.2
Psicología escolar	4	3.1
Estudiante psicología clínica	52	40.9
Estudiante consejería psicológica	6	4.7
Estudiante psicología escolar	1	0.8

Nota. (N = 127).

investigadores con el propósito de desarrollar una justa medida que examine el distanciamiento social hacia personas trans. Operacionalmente, la *distancia social hacia personas trans* se define como la distancia que una persona indica que existe entre ella y las personas trans, mediante el respaldo de ciertas declaraciones (en este caso, los ítems de la DS-T). Con este propósito se desarrollaron 13 ítems que recogen varios aspectos de la comunidad trans en general. Los ítems fueron sometidos a la opinión de 10 jueces conocedores y estudiosos de la comunidad LGBT en Puerto Rico con el objetivo de identificar si los ítems del instrumento son pertinentes al distanciamiento social hacia personas trans (método de Lawshe). Para rechazar o retener los ítems se utilizó el *content validity ratio* (CVR). Para interpretar los resultados, utilizamos los valores críticos recalculados por [Wilson, Pan y Schumsky \(2012\)](#). De acuerdo con estos autores, el valor mínimo requerido para 10 jueces, según la tabla de Schipper, es .62 para aceptar un ítem como esencial. Todos los ítems obtuvieron valores aceptables, por lo cual permanecieron en el instrumento. La escala se responde seleccionando la alternativa de preferencia en cuanto a la distancia social: 1 (*Viviría con ellos/as bajo la misma habitación*), 2 (*Viviría con ellos/as bajo el mismo techo*), 3 (*Los/as aceptaría como amigos/as*), 4 (*Preferiría tenerlos/as como conocidos/as y no como amigos/as*), 5 (*Los/as excluiría de mis relaciones sociales*), y 6 (*No quisiera ni verles*).

Procedimientos

La investigación fue aprobada por el Institutional Review Board (IRB) de la Universidad Carlos Albizu de Puerto Rico (Fall-14-09). Luego se procedió a recopilar la información de manera

presencial y en línea, a través del portal *survey-monkey.com*. Se utilizó una hoja de consentimiento informado, donde no era necesario identificarse con nombre, para notificar a los participantes: el propósito del estudio, la naturaleza voluntaria del estudio, los posibles riesgos y beneficios, así como su derecho a retirarse del estudio en cualquier momento. Además, se informó sobre la duración de su participación y su derecho a obtener los resultados del estudio. Los acercamientos presenciales se realizaron mediante visitas a centros de trabajo de profesionales de la psicología, así como visitas a centros universitarios donde cursan estudiantes graduados de psicología.

Estrategia de análisis

Una vez recopilados los datos, analizamos los mismos utilizando el sistema para análisis estadísticos IBM SPSS Statistics (versión 23.0; [IBM Corp., 2014](#)). Específicamente, realizamos análisis descriptivos de la muestra (medias y desviaciones típicas), análisis de los ítems para conocer su índice de discriminación, análisis de factores y análisis de confiabilidad para calcular el coeficiente alfa de Cronbach y división en mitades de Spearman-Brown. Se realizó un análisis de comparación de medias prueba t con el propósito de evaluar si existen diferencias estadísticamente significativas entre el estudiantado y los profesionales de la psicología.

Para el análisis factorial exploratorio se utilizó el método de extracción de mínimos cuadrados no ponderados (*unweighted least squares*) con rotación oblicua para identificar las variables latentes que subyacen en los ítems. Se utilizó este procedimiento de ajuste por dos razones principales: (1) el método de extracción de mínimos cuadrados no ponderados es el más recomendado

actualmente para trabajar con muestras pequeñas, especialmente si el número de factores a retener es pequeño (Jung, 2013) y evita la aparición de casos Heywood (saturaciones mayores que la unidad y varianzas de error negativas), más frecuentes con otros métodos de estimación (Lloret-Segura, Ferreres-Traver, Hernández-Baeza, & Tomás-Marco, 2014); y (2) la rotación oblicua es más precisa y provee más información que la rotación octogonal (Schmitt, 2011). Como criterios de aceptación, consideramos aquellos ítems con una carga factorial mayor a .50 en un solo factor (Stevens, 2002). Para conocer la capacidad de discriminación de los ítems de los factores se calculó el índice de correlación ítem-total (r_{bis}). Para conocer el grado de confiabilidad de la escala, se calcularon el coeficiente alfa de Cronbach, alfa de Cronbach estandarizado y el coeficiente de división en mitades de Spearman-Brown. Field (2013) y Kline (2000) indican que índices mayores a .70 son aceptables.

Resultados

Análisis descriptivos

Con la intención de conocer la distancia social de los profesionales y estudiantes de psicología hacia personas trans, se organizaron descriptivamente (por frecuencia) las respuestas de cada uno de los ítems de la DS-T (véase Tabla 2). Los datos muestran que casi la mitad de profesionales y estudiantes de psicología aceptarían como amistades a personas trans. No obstante, no vivirían bajo la misma habitación ni vivirían bajo el mismo techo con personas que de alguna manera transgreden el género o el sexo. Es decir, no presentan cercanía ni aceptación incondicional. Ninguno de los participantes respondió que no quisieran ni verlos/as, y solo una persona señaló

que los excluiría de sus relaciones sociales.

Análisis factorial exploratorio

Para determinar la validez de construcción lógica del instrumento, llevamos a cabo un análisis de factores utilizando el método de extracción de mínimos cuadrados no ponderados con rotación oblicua. Los resultados mostraron una estructura unifactorial que explicaba el 92% de la varianza de los datos originales. Todos los ítems de la escala original cumplieron con cargas factoriales adecuadas (ver Tabla 3), por lo cual cumplieron con los criterios de aceptación detallados anteriormente. Las pruebas de Kaiser-Meyer-Olkin ($KMO = .903$) y la prueba de esfericidad Bartlett ($\chi^2_{(78)} = 4026.526, p < .001$) apoyaron la adecuación de los datos de muestreo para el análisis.

Análisis de discriminación de ítems

Con los datos obtenidos se llevó a cabo un análisis de los ítems para determinar sus índices de discriminación. Los resultados mostraron unos índices de discriminación que fluctuaban de .94 a .98. Al observar los valores obtenidos, identificamos que todos los ítems obtuvieron índices de discriminación dentro de los puntos de corte ($r_{bis} > .30$) sugeridos en la literatura (Field, 2013; Kline, 2000). En la Tabla 3 presentamos los índices de discriminación por ítem.

Análisis de confiabilidad

Los 13 ítems fueron sometidos a un análisis de consistencia interna para determinar el

índice de confiabilidad de la escala. Para ello utilizamos dos métodos: el coeficiente alfa de Cronbach y división en mitades de Spearman-Brown. Los resultados mostraron un coeficiente alfa de .99 y de .98 para la prueba Spearman-Brown. En la Tabla 4 se presentan el alfa de Cronbach, el coeficiente de la prueba Spearman-Brown, la media y la desviación estándar de la DS-T.

Análisis de comparación

Se realizó un análisis de comparación de medias prueba *t* con el propósito de evaluar si existen diferencias estadísticamente significativas entre las medias de distancia social hacia personas trans entre profesionales de la psicología y estudiantes graduados de psicología (ver Tabla 5). Los análisis reflejaron que no existen diferencias estadísticamente significativas en distancia social hacia personas trans entre profesionales de la psicología y estudiantes graduados de psicología.

Discusión

En Puerto Rico, así como en muchos países del mundo, la comunidad trans es víctima de distanciamiento social, discriminación, prejuicio y rechazo en todos los niveles. No obstante, la comunidad científica del país no contaba con un instrumento diseñado para medir este fenómeno en la Isla. De hecho, como se ha evidenciado, no existe un instrumento en español que mida este constructo. Es por ello que el presente estudio tuvo como propósito desarrollar, validar y examinar las propiedades psicométricas de la Escala de Distancia Social hacia las Personas Trans (DS-T) con una muestra de profesionales y estudiantes graduados de psicología. Los resultados

obtenidos confirman que el instrumento posee las propiedades psicométricas adecuadas para medir distanciamiento social hacia las personas trans.

En términos teóricos, el análisis factorial exploratorio reveló la existencia de un solo factor latente en los ítems del instrumento. Esto evidencia cómo el prejuicio y la distancia social hacia la comunidad trans son consistentes en todas sus expresiones y formas; a tal grado que casi la mitad de los participantes, a pesar de estar vinculados con la profesión de la psicología, cuyo trabajo es de una naturaleza que requiere niveles bajos de distancia social, no presentan cercanía ni aceptación incondicional hacia esta comunidad. Por esta razón, mantener un discurso académico y profesional de inclusión y tener un perfil de no prejuicio no asegura cercanía social hacia la comunidad trans (Devine et al., 1991).

Por otra parte, no se observaron diferencias estadísticamente significativas en la distancia social entre estudiantes y profesionales de la psicología. Esto podría implicar, entre otras cosas, que la distancia social hacia las personas trans no se debe a un asunto generacional. También nos provee evidencia de que la experiencia en la profesión de la psicología no disminuye el distanciamiento hacia una comunidad marginada. Según Navalles-Gómez (2011), la distancia y la cercanía son los extremos de la empatía, por lo que mientras mayor es la distancia hacia una persona o comunidad, menor será la probabilidad de empatía. La empatía es una destreza básica y esencial que debe tener toda aquella persona que provee ayuda, especialmente aquellas que proveen servicios de salud mental. Un servicio sin empatía es posiblemente un servicio que no aporta a la salud, el bienestar y la calidad de vida de la persona que lo necesita. Este dato cobra aún más relevancia e importancia en el caso de la comunidad trans que, como se ha indicado previamente en este escrito, representa un grupo vulnerable de clientes

Tabla 2

Frecuencias de las respuestas de los participantes en la DS-T.

	Viviría con ellos/as bajo la misma habitación	Viviría con ellos/as bajo el mismo techo	Los/as aceptaría como amigos/as	Preferiría tenerlos/as como conocidos/as y no como amigos/as	Los/as excluiría de mis relaciones sociales	No quisiera ni verlos/as
Hombres que se visten de mujer	18 (14%)	48 (38%)	53 (42%)	8 (6%)	0 (0%)	0 (0%)
Mujeres que se visten de hombre	21 (17%)	45 (35%)	52 (41%)	9 (7%)	0 (0%)	0 (0%)
Una mujer transexual en hormonas feminizantes	21 (17%)	47 (37%)	51 (40%)	7 (5%)	1 (1%)	0 (0%)
Un hombre transexual en hormonas masculinizantes	21 (17%)	48 (38%)	48 (38%)	9 (7%)	1 (1%)	0 (0%)
Una mujer transexual en procesos quirúrgicos corporales	20 (16%)	44 (35%)	56 (44%)	6 (5%)	1 (1%)	0 (0%)
Un hombre transexual en procesos quirúrgicos corporales	22 (17%)	43 (34%)	52 (41%)	9 (7%)	1 (1%)	0 (0%)
Una mujer transexual que posea sus genitales de nacimiento	22 (17%)	48 (38%)	50 (39%)	6 (5%)	1 (1%)	0 (0%)
Un hombre transexual que posea sus genitales de nacimiento	22 (17%)	45 (35%)	52 (41%)	6 (5%)	1 (1%)	0 (0%)
Una mujer transexual o transgénero que todavía parezca hombre físicamente	19 (15%)	46 (36%)	55 (43%)	6 (5%)	1 (1%)	0 (0%)
Un hombre transexual o transgénero que todavía parezca mujer físicamente	20 (16%)	45 (35%)	55 (43%)	6 (5%)	1 (1%)	0 (0%)
Una persona que no parezca ni hombre ni mujer (andrógina)	18 (14%)	51 (40%)	49 (39%)	8 (6%)	1 (1%)	0 (0%)
Una mujer que su expresión de género sea masculina	22 (17%)	46 (36%)	51 (40%)	7 (5%)	1 (1%)	0 (0%)
Un hombre que su expresión de género sea femenina	22 (17%)	47 (37%)	52 (41%)	6 (5%)	0 (0%)	0 (0%)

con necesidades de tratamiento únicas y específicas. La empatía es un elemento clave para poder establecer una relación terapéutica efectiva, no obstante, se ha reportado en la literatura falta de competencias clínicas y actitudes de prejuicio y rechazo dentro de la misma profesión de psicología (Chang & Singh, 2016; Francia-Martínez et al., 2017; Sue & Sue, 2008). Por lo tanto, un instrumento que nos brinda información sobre la distancia social que tiene un profesional hacia una comunidad, paralelamente nos brinda información sobre la empatía que tendrá con esta población en el ejercicio de su profesión.

En cuanto a la confiabilidad de la escala, se obtuvieron índices superiores al mínimo recomendado por la literatura científica (Field, 2013; Kline, 2000), lo cual sugiere que la DS-T es un instrumento estable, reproducible y consistente en la medida de distanciamiento social hacia las personas trans. Asimismo, las correlaciones de cada ítem con la puntuación total manifiestan una consistencia interna notable. Esto sugiere que los ítems de la versión final discriminan adecuadamente y son capaces de diferenciar personas con distintos niveles de distancia social hacia la comunidad trans.

Tabla 3

Cargas factoriales de los ítems y sus índices de discriminación.

Ítems	Factor	r_{bis}
1 Hombres que se visten de mujer	.944	.94
2 Mujeres que se visten de hombre	.961	.96
3 Una mujer transexual en hormonas feminizantes	.940	.94
4 Un hombre transexual en hormonas masculinizantes	.956	.95
5 Una mujer transexual en procesos quirúrgicos corporales	.968	.96
6 Un hombre transexual en procesos quirúrgicos corporales	.959	.96
7 Una mujer transexual que posea sus genitales de nacimiento	.960	.96
8 Un hombre transexual que posea sus genitales de nacimiento	.962	.98
9 Una mujer transexual o transgénero que todavía parezca hombre físicamente	.979	.98
10 Un hombre transexual o transgénero que todavía parezca mujer físicamente	.979	.98
11 Una persona que no parezca ni hombre ni mujer (andrógina)	.950	.95
12 Una mujer que su expresión de género sea masculina	.949	.95
13 Un hombre que su expresión de género sea femenina.	.948	.95

Nota. Factor = carga factorial obtenida por el ítem en el análisis factorial; r_{bis} = índice de discriminación. (N = 127).

Tabla 4

Confiabilidad y estadística descriptiva de la DS-T.

Escala	Ítems	Alfa	Alfa _{est}	Spearman Brown	M	DE
DS-T	13	.99	.99	.98	30.97	10.48

Nota. Alfa_{est} = Alfa estandarizado; M = media; DE = desviación estándar. (N = 127).

Tabla 5Resultados del análisis de comparación de medias prueba *t* y tamaño del efecto.

Distancia Social	N	M	DE	<i>t</i>	<i>df</i>	<i>p</i>	<i>d</i>
Profesionales	67	31.43	10.58	.531	123	.60	.10
Estudiantes	58	30.43	10.42				

Nota. N = participantes; M = medias; *gl* = grados de libertad; DE = desviación estándar; *t* = valor de la prueba *t*; *p* = significancia; *d* = tamaño del efecto *d* de Cohen. (N = 127).

En cuanto a las implicaciones prácticas, se demostró que la DS-T puede ser utilizada para el desarrollo de nuevas investigaciones en el campo de la psicología caribeña. Este es un gran avance, dado que en Latinoamérica no existía un instrumento que examine el distanciamiento social hacia la comunidad trans. Aunque el concepto de

distancia social hacia la comunidad gay y lesbiana ha sido estudiado y replicado, el concepto no se había adaptado a la comunidad trans; comunidad sobre la que recae un mayor nivel de prejuicio por parte de la sociedad en general e inclusive por parte de la misma comunidad LGBT+. Del mismo modo, el concepto de distancia social se

puede adaptar a cualquier comunidad desaventajada tales como las que incluyen personas con necesidades especiales, diversidad funcional, usuarias de sustancias, diversidad étnica, con VIH/SIDA, entre otras. Conocer el grado de cercanía o de distancia social hacia estos grupos permitiría trabajar en recomendaciones y planes de acción para la reducción y/o eliminación de esta barrera terapéutica.

En términos de corrección, las puntuaciones de la DS-T se calculan mediante la sumatoria de los 13 ítems para obtener un índice general. Para obtener la puntuación total de la escala se deben sumar las puntuaciones alcanzadas en cada ítem (no hay ítems inversos). Al finalizar la sumatoria se clasifican las puntuaciones de acuerdo a los cuartiles: *baja o ninguna distancia* (13-26), *distancia moderada baja* (31-45), *distancia moderada alta* (46-62) y *distancia alta* (63-78).

Limitaciones y recomendaciones

Varias limitaciones fueron identificadas en este estudio. Una de las limitaciones principales fue el tamaño de la muestra, la cual fue pequeña para un estudio instrumental de esta envergadura, considerando la cantidad de variables en el mismo. No obstante, se tomó en consideración el tamaño de la muestra y el nivel de significancia al momento de tomar decisiones sobre los ítems, así como un método de extracción adecuado a dicha limitación (Jung, 2013). Esto permite sostener la calidad y confiabilidad de los resultados obtenidos (Stevens, 2002). Cabe añadir que los análisis se llevaron a cabo con una sola muestra, lo cual no es recomendable, ya que se pueden repetir los mismos errores de muestreo. Otra limitación fue que no se logró establecer la confiabilidad del instrumento a través del tiempo, solo se pudo hacer a

través de sus elementos.

Tomando en consideración las limitaciones identificadas, se recomienda realizar un nuevo estudio y administrar la DS-T a una muestra mayor de participantes, que incluya participantes de la población general, para examinar si se sostienen las propiedades psicométricas del instrumento y realizar el proceso de validación cruzada. De igual forma, se debe llevar a cabo un análisis de factores confirmatorio y examinar la confiabilidad temporal a través de la técnica de prueba-reprueba. Por último, se recomienda validar la DS-T en otras poblaciones latinoamericanas para examinar sus propiedades psicométricas en distintos contextos nacionales e internacionales.

En síntesis, la Escala de Distancia Social hacia las Personas Trans (DS-T) mostró preliminarmente indicadores adecuados de validez y confiabilidad, lo que sugiere que la DS-T tiene el potencial de contribuir al avance de la investigación con miras a comprender con mayor amplitud el fenómeno del distanciamiento social y el prejuicio hacia la comunidad trans. Se espera que la DS-T sea de beneficio en el ámbito de la investigación y facilite nuevos conocimientos sobre este fenómeno en Puerto Rico y en América Latina.

Referencias

- American Psychological Association. (2012). Guidelines for psychological practice with lesbian, gay, and bisexual clients. *American Psychologist*, 67(1), 10-42. doi: [10.1037/a0024659](https://doi.org/10.1037/a0024659)
- American Psychological Association. (2015). Guidelines for psychological practice with transgender and gender nonconforming people. *American Psychologist*, 70(9), 832-864. doi: [10.1037/a0039906](https://doi.org/10.1037/a0039906)
- Arteaga-Botello, N., & Lara-Carmona, V. L. (2004). Violencia y distancia social: Una revisión. *Papeles de Población*, 10(40), 169-191. Recuperado de: [ht-](http://www.inecubator.org/revista/inecubator/inecubator_10_40_169-191.pdf)

[tps://rppoblacion.uaemex.mx/index](https://rppoblacion.uaemex.mx/index)

- Asociación de Psicología de Puerto Rico. (2014). *Estándares para el trabajo e intervención en comunidades lesbianas, gay, bisexuales e identidades trans (LGBT)*. San Juan, Puerto Rico: Autor.
- Ato, M., López-García, J. J., & Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en psicología. *Anales de Psicología*, 29(3), 1038-1059. doi: [10.6018/analesps.29.3.178511](https://doi.org/10.6018/analesps.29.3.178511)
- Chang, S. C., & Singh, A. A. (2016). Affirming psychological practice with transgender and nonconforming people of color. *Psychology of Sexual Orientation and Gender Diversity*, 3(2), 140-147. doi: [10.1037/sgd0000153](https://doi.org/10.1037/sgd0000153)
- Choi, O., Lee, K., Lee, K., & Kim, J. (2017). Influences of stereotype and social distance on prejudice toward African Americans. *Journal of Psychology in Africa*, 27(1), 13-17. doi: [10.1080/14330237.2016.1268262](https://doi.org/10.1080/14330237.2016.1268262)
- Coleman, E., Bockting, W., Botzer, M., Cohen-Kettenis, P., DeCuypere, G., Feldman, J., ... Zucker, K. (2012). Standards of care for the health of Transsexual, transgender and gender-nonconforming people, Version 7. *International Journal of Transgenderism*, 13(4), 165-232. doi: [10.1080/15532739.2011.700873](https://doi.org/10.1080/15532739.2011.700873)
- Coppari, N., Arcondo, G., Bagnoli, L., Chaves, M., Corvalán, M., Enciso, A., ... Rodríguez, X. (2014). Prejuicio y distancia social hacia la homosexualidad en universitarios de psicología de Paraguay. *Salud y Sociedad*, 5(3), 240-252. doi: [10.22199/s07187475.2014.0003.00001](https://doi.org/10.22199/s07187475.2014.0003.00001)
- Devine, P. G., Monteith, M. J., Zuwerink, J. R., & Elliot, A. J. (1991). Prejudice with and without compunction. *Journal of Personality & Social Psychology*, 60(6), 817-830. doi: [10.1037/0022-3514.60.6.817](https://doi.org/10.1037/0022-3514.60.6.817)
- Fernández-Rodríguez, M. C., & Calderón-Squiabro, J. (2014). Prejuicio y distancia social hacia personas homosexuales por parte de jóvenes universitarios. *Revista Puertorriqueña de Psicología*, 25(1), 52-60. Recuperado de <http://www.ojs.repsasppr.net/index.php/reps/index>
- Field, A. P. (2013). *Discovering statistics using SPSS*. Londres, Inglaterra: SAGE.
- Francia-Martínez, M., Esteban, C., & Lespier, Z. (2017). Actitudes, conocimiento y distancia social de psicoterapeutas con la comunidad transgénero y transexual. *Revista Puertorriqueña de Psicología*, 28(1), 98-113. Recuperado de <http://www.ojs.repsasppr.net/index.php/reps/index>
- García, H. (1984). *Actitudes hacia homosexuales y las lesbianas de cuatro grupos de estudiantes universitarios* (Tesis de maestría inédita). Universidad de Puerto Rico, Río Piedras, Puerto Rico.
- González-Guzmán, M. (1998). *Actitud de los estudiantes de salud pública general y educación en salud de la escuela graduada de salud pública de la Universidad de Puerto Rico, Recinto de Ciencias Médicas, hacia los homosexuales y lesbianas* (Tesis de maestría inédita). Universidad de Puerto Rico, Recinto de Ciencias Médicas, Puerto Rico.
- Grandón-Fernández, P., Cova-Solar, F., Saldívia-Bórquez, S., & Bustos-Navarrete, C. (2015). Análisis psicométrico y adaptación de la Escala de Distancia Social (DS) en una muestra chilena. *Salud Mental*, 38(2), 117-122. doi: [10.17711/SM.0185-3325.2015.016](https://doi.org/10.17711/SM.0185-3325.2015.016)
- Horn, S. (2013). Attitudes about sexual orientation. En C. Patterson & A. D'Angelis (Eds.), *Handbook of psychology and sexual orientation* (pp. 239-251). New York, N.Y.: Oxford University. doi: [10.1093/acprof:oso/9780199765218.003.0017](https://doi.org/10.1093/acprof:oso/9780199765218.003.0017)
- Hughes, E. C., Johnson, C. S., Masuoka, J., Redfield, R., & Wirth, L. (1950). *The collected papers of Robert E. Park*. Glencoe, IL: The Free.
- IBM Corp. (2014). IBM SPSS Statistics for Windows (version 23.0). [software de cómputo]. Armonk, NY: IBM Corp.
- Jung, S. (2013). Exploratory factor analysis with small sample sizes: A comparison of three approaches. *Be-*

- havioural Processes*, 97(1), 90-95. doi: [10.1016/j.beproc.2012.11.016](https://doi.org/10.1016/j.beproc.2012.11.016)
- Kline, P. (2000). *The Handbook of Psychometric Testing* (2^a ed.). New York, NY: Routledge.
- LGBT Advisory Committee, & San Francisco Human Rights Commission. (2011). *Bisexual invisibility: Impacts and recommendations*. San Francisco, CA: San Francisco Human Rights Commission. Recuperado de <https://sf-hrc.org>
- Link, B. G., Cullen, F. T., Frank, J., & Wozniak, J. F. (1987). The social rejection of former mental patients: Understanding why labels matter. *American Journal of Sociology*, 92(6), 1461-1500. doi: [10.1086/228672](https://doi.org/10.1086/228672)
- Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., & Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: Una guía práctica, revisada y actualizada. *Anales de Psicología*, 30(3), 1151-1169. doi: [10.6018/analesps.30.3.199361](https://doi.org/10.6018/analesps.30.3.199361)
- Maurer, T. W., & Keim, C. (2018). Teaching about prejudice with a Bogardus Social Distance Scale activity: Replication and extension. *International Journal for the Scholarship of Teaching and Learning*, 12(1), 1-10. doi: [10.20429/ijstol.2018.120107](https://doi.org/10.20429/ijstol.2018.120107)
- McCann, E., & Sharek, D. (2014). Survey of lesbian, gay, bisexual and transgender people's experiences of mental health services in Ireland. *International Journal of Mental Health Nursing*, 23(2), 118-127. doi: [10.1111/inm.12018](https://doi.org/10.1111/inm.12018)
- Montero, I., & León, O. G. (2007). A guide for naming research studies in psychology. *International Journal of Clinical and Health Psychology*, 7(3), 847-862. Recuperado de <https://www.redalyc.org>
- Navalles-Gómez, J. (2011). Acercamientos a la distancia social. *Athenea Digital: Revista de Pensamiento e Investigación Social*, 11(2), 173-190. doi: [10.5565/rev/athenead/v11n2.763](https://doi.org/10.5565/rev/athenead/v11n2.763)
- Nieves-Rosa, L. E. (2012). Homofobia al estilo universitario. *Revista Puertorriqueña de Psicología*, 23(2), 62-76. Recuperado de <http://www.ojs.repsasppr.net>
- Ramos-Pibernus, A. G., Rodríguez-Madera, S. L., Padilla, M., Varas-Díaz, N., & Vargas-Molina, R. (2016). Intersections and evolution of 'Butch-trans' categories in Puerto Rico: Needs and barriers of an invisible population. *Global Public Health*, 11(7), 966-980. doi: [10.1080/17441692.2016.1180703](https://doi.org/10.1080/17441692.2016.1180703)
- Rodríguez-Madera, S. L. (2009). *Género Trans: Transitando por las zonas grises*. San Juan, Puerto Rico: Terranova.
- Rodríguez-Madera, S. L. (2012). Trans-acciones de la carne: Criminalización de mujeres trans que ejercen el trabajo sexual. En S. Serrano (Ed.), *Registros criminológicos contemporáneos* (pp. 107-150). San Juan, Puerto Rico: Situm.
- Rodríguez-Madera, S., Ramos, A., Padilla, M., & Varas-Díaz, N. (2016). Radiografías de las comunidades Trans en Puerto Rico: Visibilizando femineidades y masculinidades alternas. En M. Vázquez-Rivera, A. Martínez-Taboas, M. Francia-Martínez & J. Toro-Alfonso (Eds.), *LGBT 101: Una mirada introductoria al colectivo* (pp. 315-342). Hato Rey, Puerto Rico: Publicaciones Puertorriqueñas.
- Rodríguez-Polo, J., Ayvar, A., Dávila, A., Andino, P., Quiñones, C., Rodríguez, L., ... Pacheco, T. (2018). Prejuicio y distancia social hacia las personas gays y lesbianas en una muestra de empleados en Puerto Rico. *Revista Griot*, 11(1), 16-33. Recuperado de <https://revistas.upr.edu/index.php/griot/index>
- Schmitt, T. A. (2011). Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment*, 29(4), 304-321. doi: [10.1177/0734282911406653](https://doi.org/10.1177/0734282911406653)
- Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences* (4^a ed.). Mahwah, NJ: Lawrence Erlbaum.
- Sue, D., & Sue, D. M. (2008). *Foundations of counseling and psychotherapy: Evidence based practice for a diverse society*. New York, NY: J. Wiley and Sons.

- Toro-Alfonso, J., & Varas-Díaz, N. (2004). Los otros: Prejuicio y distancia social hacia hombres gay y lesbianas en una muestra de estudiantes de nivel universitario. *International Journal of Clinical and Health Psychology, 4*(3), 537-551. Recuperado de <https://aepc.es/ijchp/busca.php>
- Vázquez-Rivera, M., Nazario-Serrano, J., & Sayers-Montalvo, S. K. (2012). La deseabilidad social en estudiantes graduados/as de psicología clínica y psicólogos/as clínicos con licencia. *Revista Puertorriqueña de Psicología, 23*(2), 94-106. Recuperado de <http://www.ojs.repsasppr.net>
- Vázquez-Rivera, M., Esteban, C., & Toro-Alfonso, J. (2018). Hacia una psicología libre de paños tibios: Actitudes, prejuicio y distancia social de psicoterapeutas hacia gays y lesbianas. *Perspectivas en Psicología: Revista de Psicología y Ciencia Afines, 15*(1), 17-30. Recuperado de <https://www.redalyc.org/revista.oa?id=4835&tipo=coleccion>
- Wark, C., & Galliher, J. F. (2007). Emory Bogardus and the origins of the Social Distance Scale. *The American Sociologist, 38*(4), 383-395. doi: [10.1007/s12108-007-9023-9](https://doi.org/10.1007/s12108-007-9023-9)
- Weaver, C. N. (2008). Social distance as a measure of prejudice among ethnic groups in the United States. *Journal of Applied Social Psychology, 38*(3), 779-795. doi: [10.1111/j.1559-1816.2007.00326.x](https://doi.org/10.1111/j.1559-1816.2007.00326.x)
- Wilson, F. R., Pan, W., & Schumsky, D. A. (2012). Recalculation of the critical values for Lawshe's content validity ratio. *Measurement and Evaluation in Counseling and Development, 45*(3), 197-210. doi: [10.1080/07481756.2017.1308227](https://doi.org/10.1080/07481756.2017.1308227)
-

-
13. Un hombre que su expresión de género sea femenina
-
-

Método de corrección sumaria

Puntuación regular: los ítems se puntúan de 1 (*Viviría bajo la misma habitación*) a 6 (*No quisiera ni verles*).

Puntuación inversa: no hay ítems inversos.

Estructura dimensional de la Escala Clinical Outcomes in Routine Evaluation-Outcome Measure (CORE-OM) en universitarios mexicanos

Dimensional Structure of the Clinical Outcomes in Routine Evaluation - Outcome Measure (CORE-OM) in Mexican University Students

Introducción
Método
Resultados
Discusión
Referencias

Janett Esmeralda Sosa-Torralba *¹, Martha Patricia Romero-Mendoza²,
María Elena Medina-Mora-Icaza², Enrique Méndez-Ríos², Bertha Blum-Grynberg¹

1- Facultad de Psicología, Universidad Nacional Autónoma de México.

2- Dirección de Investigaciones Epidemiológicas y Sociales, Instituto Nacional de Psiquiatría "Ramón de la Fuente Muñiz", México.

Recibido: 17/09/2019 Revisado: 11/11/2019 Aceptado: 03/12/2019

Resumen

Este trabajo verifica la estructura dimensional del Clinical Outcomes in Routine Evaluation-Outcome Measure (CORE-OM; Evans et al., 2000) mediante un estudio cuantitativo-transversal en una muestra de 466 estudiantes universitarios elegidos por muestreo de conveniencia; 181 de 18 a 30 años que solicitaron atención psicológica, y 285 de 17 a 29 años que no estaban en búsqueda de atención. A través de los análisis factoriales exploratorio y confirmatorio se identificaron tres dimensiones para el instrumento con adecuados niveles de consistencia interna, lo cual replica sustancialmente los resultados previos de la versión original. Se hallaron diferencias estadísticamente significativas entre varones y mujeres y entre la población consultante y no consultante. Finalmente, se comprobó si existía funcionamiento diferencial del ítem por sexo y por población consultante. Cinco reactivos se hallaron en la primera categoría y tres en la segunda. La estructura interna del CORE-OM explora un amplio conjunto de síntomas de malestar psicológico y resulta útil para su evaluación en jóvenes universitarios.

Palabras clave: CORE-OM, malestar psicológico, universitarios, AFC, DIF

Abstract

This work verifies the dimensional structure of the Clinical Outcomes in Routine Evaluation - Outcome Measure (CORE-OM; Evans et al., 2000) through a quantitative-cross-sectional study in a sample of 466 university students who were selected by convenience sampling. 181 ages 18 to 30 who had requested psychological attention; and 285 ages 17 to 29 who were not seeking care. Through exploratory and confirmatory factor analyses, three dimensions for the instrument were identified, which presented adequate levels of internal consistency. This substantially replicates the previous results of the original version. Statistically significant differences were found in relation to sex, as well as of the consulting and non-consulting population. Differential item functioning (DIF) was carried out to explore these differences, and five items were found in the first category and three in the second. The internal structure of CORE-OM explores a wide range of symptoms of psychological distress and it is useful for evaluation in young university students.

Keywords: CORE-OM, psychological distress, university, CFA, DIF

*Correspondencia a: Dra. Janett Esmeralda Sosa Torralba. Facultad de Psicología, UNAM. Avenida Universidad 3004, Cubículo 2. Edificio D. Copilco Universidad, 04510, Ciudad de México. Tel: 56220555, ext. 41242. E-mail: jan_est@comunidad.unam.mx

Cómo citar este artículo: Sosa-Torralba, J. E., Romero-Mendoza, M. P., Medina-Mora-Icaza, M. E., Méndez-Ríos, E., & Blum-Grynberg, B. (2020). Estructura dimensional de la Escala Clinical Outcomes in Routine Evaluation - Outcome Measure (CORE-OM) en estudiantes universitarios mexicanos. *Revista Evaluar*, 20(1), 49-66. Recuperado de <https://revistas.unc.edu.ar/index.php/revaluar>

Introducción

En México, la salud integral de la población estudiantil de adolescentes y jóvenes adultos ha sido reconocida como un tema preponderante en instituciones educativas (López-Bárcena & González de Cossío-Ortiz, 2005). En el caso de los jóvenes, se han detectado síntomas somáticos, de depresión, ansiedad, tristeza, soledad; conductas de riesgo, e incluso ideación suicida (Guevara-Ruiseñor, 2008; Pérez-Padilla, Ponce-Rojo, Hernández-Contreras, & Márquez-Muñoz, 2010; Rivera-Ledesma, Caballero-Suárez, Pérez-Sánchez, & Montero-López-Lena, 2013; Rosales-Pérez, Córdova-Osnaya, & Ramos-Clatempa, 2012). Ante estas situaciones, se ha hecho énfasis en identificar el malestar emocional en los jóvenes, así como en detectar las áreas de atención para dar solución a las consecuencias de este (Guevara-Ruiseñor, 2008; Pérez-Padilla et al., 2010).

El malestar emocional se identifica por sus manifestaciones en el estado de ánimo, el tono afectivo, la percepción y los pensamientos, que alteran la funcionalidad y las representaciones que conforman el mundo interno del sujeto (Espíndola-Hernández et al., 2006). Con el propósito de su evaluación se han diseñado varias escalas, inventarios e instrumentos, uno de ellos es el CORE-OM (Clinical Outcomes in Routine Evaluation-Outcome Measure; Evans et al., 2000). Fue presentado por el CORE System Group en 1998 en el Reino Unido como un medio estandarizado y viable para la evaluación de los resultados de los servicios de terapias psicológicas, dentro del marco de la práctica basada en la evidencia (Mellor-Clark, Barkham, Connel, & Evans, 1999). Evalúa áreas de síntomas, funcionamiento, bienestar subjetivo y riesgo hacia sí mismo y

hacia otros, por lo que es considerado adecuado como herramienta de detección inicial (Margison et al., 2000; Mellor-Clark et al., 1999) en servicios de atención primaria y secundaria de la salud (Barkham, Gilbert, Connell, Marshall, & Twigg, 2005; Beck, Burdett, & Lewis, 2015; Navascués, Calvo-Medel, & Bombin-Martín, 2016).

Este instrumento ha sido evaluado, traducido y aplicado a diferentes poblaciones. En el país donde fue desarrollado, se han valorado sus propiedades psicométricas y su utilidad clínica en pacientes con diferentes diagnósticos y en diferentes escenarios terapéuticos; por ejemplo, en estudiantes universitarios, pacientes con diagnóstico psiquiátrico, pacientes en psicoterapia sexual, en ambientes hospitalarios y en intervenciones con foco en la pareja (Branney & Barkham, 2006; Casey et al., 2017; Connell, Barkham, & Mellor-Clarck, 2007; Perry, Barkham, & Evans, 2013).

En algunas de las traducciones realizadas se ha visto que es un medio útil, pertinente y adecuado para la valoración del malestar psicológico, tanto en la práctica clínica como en la investigación (Dias-Sales, de Matos-Moleiro, Evans, & Gomes-Alves, 2012; Elfström et al., 2013; Juhová et al., 2018; Uji, Sakamoto, Adachi, & Kitamura, 2012; Viliū-nienė et al., 2013; Zhang et al., 2019).

La versión castellana del CORE-OM en España ha mostrado propiedades psicométricas aceptables (Trujillo-Terán, 2012), y la versión validada en Costa Rica ha sido sugerida como confiable y aplicable para medir la efectividad psicoterapéutica (Vargas-Fallas & Prada-Villalobos, 2013). Sin embargo, en comparación con los análisis de la versión original (Evans et al., 2002), los análisis exploratorios para estas versiones han presentado variaciones en la estructura de sus componentes. Por otro lado, en ambas versio-

nes se hallaron diferencias a nivel de género y de población clínica y no clínica, motivo por el cual se ha considerado necesario tomar en cuenta estas variables cuando se interpretan puntuaciones individuales en cada una de las dimensiones, en el caso de datos clínicos (Evans et al., 2002; Trujillo-Terán, 2012).

En México no se ha documentado investigación acerca de este instrumento, y dada la necesidad de contar con herramientas válidas y confiables que permitan medir el malestar psicológico en los jóvenes, el objetivo del presente estudio se centró en verificar la estructura interna del CORE-OM para identificar la complejidad del constructo, así como para obtener evidencias de validez y confiabilidad del instrumento que avalen su uso en el contexto universitario. Esta herramienta permitiría conocer las manifestaciones que experimentan los estudiantes con malestar emocional con el propósito de emprender programas de tratamiento que permitan mejorar su calidad de vida.

Método

Participantes

Se trata de un estudio cuantitativo, transversal-descriptivo (Echevarría, 2016), que incluyó a 466 estudiantes universitarios de la Universidad Nacional Autónoma de México (UNAM), seleccionados mediante muestreo por conveniencia (Otzen & Manterola, 2017). El criterio de inclusión considerado fue que pertenecieran al Campus de Ciudad Universitaria: 181 buscaron atención en centros de servicios psicológicos de la misma Universidad, los otros 285 alumnos no buscaban atención psicológica al momento de responder al instrumento. Los estudiantes no consultantes te-

nían entre 18 y 30 años ($M = 21$, $DE = 2$), 40% eran hombres, 58% mujeres y el 2% no proporcionó esa información. Aquellos que acudieron a solicitar atención tenían entre 17 y 29 años ($M = 23$, $DE = 3.8$), 61% eran mujeres, 38% hombres y 2 personas (1%) no dieron este dato.

Instrumento

CORE-OM. Se utilizó la versión en castellano que se encuentra disponible en la página web del CORE Information Management Systems disponible para hombres y mujeres (CORE IMS, 2001). Incluye 34 reactivos de autoaplicación que exploran 4 subescalas: 1) *bienestar subjetivo* (4 ítems), 2) *síntomas/problemas* (12 ítems que evalúan ansiedad, depresión, problemas físicos y trauma), 3) *funcionamiento* (12 ítems, incluyen funcionamiento general, en relaciones cercanas y en relaciones sociales), y 4) *riesgo* (6 ítems de daño hacia sí mismo y actos de agresión hacia otros). En la Tabla 1 se muestran los reactivos correspondientes a cada una de estas subescalas.

Cada ítem cuestiona la frecuencia de malestar durante los últimos siete días y se valora en una escala de 5 puntos tipo Likert según la respuesta elegida (0 = *Nunca*, 4 = *Siempre o casi siempre*). La mayoría de los reactivos reflejan reacciones negativas. Los reactivos 3, 4, 7, 12, 19, 21, 31 y 32 reflejan reacciones positivas de la persona; en consecuencia, la puntuación que se obtiene de estos ítems está invertida. De este modo, también en la escala de bienestar subjetivo mayores puntuaciones indican un mayor nivel de malestar.

Se obtienen calificaciones por subescala y de la prueba total; las puntuaciones, en ambos casos, varían de 0 a 40. La escala está diseñada de tal manera que las calificaciones más altas indi-

can mayores problemas o síntomas. Para la puntuación total del instrumento, también se considera que se sumen todas las subescalas menos la de Riesgo (T-R), por ser considerada como anexa al instrumento, y cualquiera de sus ítems que puntúe por encima de 0 es utilizado como indicador clínico de presencia de ideación suicida, autolesiones o actos de agresión a terceros.

El análisis de consistencia interna, tanto para el total de la escala como para cada una de las dimensiones, ha generado valores de entre $\alpha = .6$ y $\alpha = .9$ en la versión original y en las versiones en español (Evans et al., 2002; Trujillo-Terán, 2012; Vargas-Fallas & Prada-Villalobos, 2013).

Procedimiento

El CORE-OM fue evaluado en un análisis de contenido por ocho expertos cuya trayectoria se caracteriza por su experiencia clínica en salud mental (siete psicólogos y un psiquiatra). Mediante un cuestionario escrito cada juez valoró cualitativamente el grado de pertinencia de los ítems con respecto al constructo de malestar emocional. Los aportes cualitativos de los jueces indicaron que la escala era apropiada para su propósito de evaluación. Posteriormente, el instrumento se aplicó en dependencias de la Universidad a los estudiantes no consultantes. Asimismo, se pidió autorización en dos Centros de Atención Psicológica y se aplicó a universitarios que llegaron a solicitar el servicio. En ambos grupos se entregaba el instrumento a los sujetos, siguiendo la distinción de la versión para hombres y para mujeres, y era completado en presencia del aplicador para atender dudas o problemas que pudieran presentarse.

Análisis de los datos

Los datos se analizaron en tres fases consecutivas. Para las dos primeras fases se utilizó el programa estadístico SPSS Statistics 21.0 (IBM Corporation, 2012). En la primera fase se analizaron los estadísticos descriptivos de los ítems (media, desviación estándar, asimetría y curtosis). Con el fin de analizar la estructura interna del CORE-OM, se realizó un análisis factorial exploratorio (AFE). Ante el incumplimiento de la normalidad de algunos de los ítems se utilizó el método de extracción de mínimos cuadrados no ponderados, ya que este método no hace suposiciones acerca de la distribución y es adecuado para las variables ordinales (Forero, Maydeu-Olivares, & Gallardo-Pujol, 2009), y rotación oblicua (promax) por ser adecuada para factores que están relacionados (Campo-Arias, Herazo, & Oviedo, 2012). Para la extracción del número de factores, aparte de tener en cuenta los resultados de los estudios previos de la versión original (Evans et al., 2002; Lyne, Barrett, Evans, & Barkham, 2006), se consideraron criterios teóricos y metodológicos (Campo-Arias et al., 2012): el primer criterio parte del supuesto de que un factor es importante si lo forman entre tres y cinco reactivos. El segundo fue aplicar el criterio de Kaiser (valor *eigen* superior a la unidad) y el criterio de Gorsuch o de interpretabilidad de las diferentes estructuras factoriales resultantes. El tercer criterio establecía que el modelo factorial garantizara la máxima explicación de varianza sin perder validez de contenido y que los ítems saturaran en su factor de pertenencia con valores que alcanzaran coeficientes iguales o superiores a .30. Se procedió con múltiples análisis factoriales exploratorios, entre los que se ajustó el número de factores a extraer para encontrar una estructura de factores consistente.

En la segunda fase se realizó un análisis factorial confirmatorio (AFC) a través del programa AMOS (integrado en el SPSS 21.0; [IBM Corporation, 2012](#)) para ratificar el modelo obtenido en el exploratorio. Se contrastaron tres modelos diferentes de la adecuación de las estructuras dimensionales propuestas de la versión original: modelo de dos factores, un componente general de malestar psicológico que incluye el conjunto de reactivos con reacciones negativas y positivas, y otro componente de riesgo (modelo 1; [Lyne et al., 2006](#)); modelo de tres factores, un componente de los reactivos de reacciones negativas, otro grupo de reactivos de reacciones positivas, y un último componente que agrupa los reactivos de riesgo (modelo 2; [Evans et al., 2002](#)); modelo de cuatro factores, buscando coincidencia con las subescalas del instrumento ([CORE IMS, 2001](#)).

Para el AFC, el método de estimación utilizado fue el de mínimos cuadrados no ponderados, ya que al indagar la normalidad multivariante se encontraron valores de curtosis multivariada superiores a 5.99 ($CR = 242.463$). Para la determinación del ajuste del modelo, tal como es recomendado, se evaluaron varios índices de ajuste ([Byrne, 2009](#)). Se utilizaron: el estadístico χ^2 , el índice de bondad de ajuste (GFI), el índice corregido de bondad de ajuste (AGFI), el índice de ajuste normativo (NFI) y la raíz cuadrada de los residuos al cuadrado (RMR). En el caso del GFI, AGFI y NFI se consideraron aceptables valores superiores a .90 ([Bentler, 1992](#); [Byrne, 2009](#)). Para la RMR se consideraron valores aceptables en el rango de menores o iguales que .05 ([Byrne, 2009](#)).

Se realizó un análisis de correlaciones entre las dimensiones resultantes, utilizando el coeficiente de correlación de Spearman. La consistencia interna fue analizada mediante alfa de Cronbach. Se analizaron diferencias de sexo y entre

quienes buscaron y quienes no buscaron atención psicológica mediante la prueba *t* de Student.

En la tercera fase se examinó el funcionamiento diferencial de los ítems (DIF) para analizar posibles sesgos producidos por el efecto del sexo y por el efecto de la población consultante de atención psicológica. Este estadístico trata de comparar la ejecución de un ítem entre el grupo de referencia y el grupo focal a través de distintos niveles de un determinado criterio de equiparación, asumiendo que en cada nivel los sujetos de uno y otro grupo son comparables, por lo que si no hay DIF ambos lo ejecutarán por igual. Un ítem presenta DIF cuando la probabilidad de ser resuelto correctamente por individuos con el mismo nivel de rasgo varía en función de su grupo de pertenencia (sexo, nivel socioeconómico, etc.) que genera una falta de equivalencia métrica entre las puntuaciones ([Elosua-Oliden, 2003](#)). En este trabajo se empleó el test generalizado de Mantel-Haenszel (GMH; [Mantel & Haenszel, 1959](#)). Para el análisis por sexo se consideró como grupo de referencia a las mujeres. Para el análisis por población consultante se consideró como grupo de referencia a quienes buscaron atención psicológica.

Este análisis estadístico se implementó mediante dos pruebas de significancia: en primer lugar, la prueba $M - \chi^2$ con $gl = 1$, para ítems politómicos. Un resultado $M - \chi^2 > 3.84$ detecta un DIF en el nivel $p < .05$, utilizando para el presente estudio 6.63 en el nivel $p < .01$. En segundo lugar, se estimó la significancia práctica del DIF usando los resultados obtenidos con el estimador de Liu-Agresti (LA LOR; [Penfield, 2013](#)). La magnitud del DIF para ítems politómicos mediante LA LOR se divide en los niveles de insignificante ($< .43$), moderado (entre $.43$ y $.64$) y grande ($\geq .64$; [Penfield & Algina, 2006](#)). Los valores positivos indican DIF a favor del grupo de referencia y los

valores negativos indican DIF a favor del grupo focal (Penfield, 2013). De igual manera, se aplicó la estimación estandarizada del índice común de probabilidad acumulativa de LA LOR (LOR Z), en el cual valores más grandes a 2.0 o menores a -2.0 son considerados como evidencia de DIF (Penfield, 2013). Para este estadístico se utilizó el programa informático DIFAS 5.0 (Penfield, 2009).

Consideraciones éticas

El protocolo de investigación fue aprobado por el Comité de Ética del Programa de Maestría y Doctorado en Psicología de la Facultad de Psicología de la UNAM. A los participantes se les explicó el objetivo del estudio y se requirió su consentimiento informado.

Resultados

En la Tabla 1 se presentan los estadísticos descriptivos de centralidad, dispersión y normalidad para los 34 ítems de la escala CORE-OM en la muestra total, referidos a la media, desviación estándar, asimetría y curtosis. Como se puede apreciar, las medias más bajas corresponden a los ítems pertenecientes a indicadores de riesgo de daño hacia sí mismo, el ítem 16 (*He hecho planes para acabar con mi vida*) es el de menor puntuación ($M = .19$). Respecto a la normalidad de las variables, cinco de los ítems presentan índices de asimetría y curtosis mayores al rango ± 1.5 (reactivos 9, 16, 22, 24 y 34), lo cual se considera inadecuado para el supuesto de normalidad (George & Mallery, 2001, como se citó en Pérez & Medrano, 2010).

Análisis factorial exploratorio

Puesto que todavía no ha sido referido ningún estudio factorial del CORE-OM en población mexicana, pareció conveniente realizar un análisis factorial de tipo exploratorio (AFE). La adecuación muestral de Kaiser-Meyer-Olkin (KMO) mostró un valor de .95 y el test de esfericidad de Bartlett resultó estadísticamente significativo ($\chi^2 = 7562.562$, $gl = 561$, $p < .001$); en consecuencia, la aplicación del análisis factorial resultaba pertinente.

Tras el análisis, se obtuvo una estructura dimensional de tres factores con capacidad para explicar el 48% de la varianza total. Los reactivos *He sido violento(a) físicamente hacia los demás* (ítem 6) y *He amenazado e intimidado a otra persona* (ítem 22) tuvieron cargas factoriales menores a .30 con comunalidades de .10. Los demás ítems alcanzaron pesos factoriales de entre .41 y .80, y comunalidades de entre .19 y .64. En la Tabla 2, se muestran la matriz factorial rotada, las saturaciones de los pesos factoriales y los valores de comunalidad.

Esta estructura de tres dimensiones, al igual que la distribución de los ítems, coincidió parcialmente con la propuesta de tres factores de la versión inglesa (Evans et al., 2002). La dimensión de mayor peso factorial fue donde se agruparon los ítems que reflejan reacciones negativas e indican malestar psicológico. Esta dimensión explica el 36% de la varianza total. A continuación, le sigue la dimensión que agrupó los reactivos de reacciones positivas relacionados con el bienestar psicológico, y explica el 7% de la varianza. La tercera dimensión, que hace referencia al riesgo hacia sí mismo y hacia otros, es la que menos varianza explica, con un 5%.

Tabla 1

Análisis descriptivo de los ítems de la versión en castellano del CORE-OM

Ítems	M	DE	Asimetría	Curtosis
1. Me he sentido muy solo(a) y aislado(a) (F-rc)	1.33	1.10	0.51	-0.40
2. Me he sentido tenso(a), ansioso(a) o nervioso(a) (S-a)	2.34	0.98	-0.39	-0.11
3. He sentido que tenía a alguien en quien apoyarme cuando lo he necesitado (F-rc)	1.24	1.14	0.65	-0.41
4. Me he sentido bien conmigo mismo(a) (B)	1.17	1.02	0.62	-0.24
5. Me he sentido totalmente sin energía ni entusiasmo (S-d)	1.64	1.08	0.27	-0.61
6. He sido violento(a) físicamente hacia los demás (R-o)	0.61	0.90	1.38	1.02
7. Me he sentido capaz de afrontar las cosas cuando han ido mal (F-g)	1.20	1.00	0.60	-0.18
8. He tenido molestias, dolores y otros problemas físicos (S-f)	1.59	1.09	0.19	-0.60
9. He pensado en hacerme daño a mí mismo(a) (R-s)	0.30	0.73	2.82	8.17
10. Me ha costado mucho hablar con la gente (F-rs)	1.26	1.16	0.66	-0.39
11. La tensión y la ansiedad me han impedido hacer cosas importantes (S-a)	1.24	1.17	0.62	-0.61
12. Me he sentido satisfecho(a) con las cosas que he hecho (F-g)	1.19	0.99	0.58	-0.19
13. Me han inquietado pensamientos y sentimientos no deseados (S-t)	1.61	1.17	0.21	-0.89
14. He tenido ganas de llorar (B)	1.80	1.15	0.02	-0.85
15. He sentido pánico o terror (S-a)	1.00	1.05	0.85	0.01
16. He hecho planes para acabar con mi vida (R-s)	0.19	0.62	4.03	18.11
17. Me he sentido agobiado(a) por mis problemas (B)	1.56	1.15	0.26	-0.84
18. He tenido dificultad para conciliar el sueño o permanecer dormido(a) (S-f)	1.39	1.26	0.47	-0.91
19. He sentido afecto o cariño por alguien (F-rc)	0.77	0.91	1.22	1.26
20. Me ha sido imposible dejar a un lado mis problemas (S-a)	1.53	1.15	0.47	-0.53
21. He sido capaz de hacer la mayoría de las cosas que tenía que hacer (F-g)	1.24	1.09	0.63	-0.45
22. He amenazado o intimidado a otra persona (R-o)	0.32	0.70	2.36	5.55
23. Me he sentido desesperado(a) o sin esperanza (S-d)	1.05	1.16	0.89	-0.18
24. He pensado que sería mejor que estuviera muerto(a) (R-s)	0.36	0.81	2.52	6.19
25. Me he sentido criticado(a) por los demás (F-rs)	1.46	1.12	0.54	-0.45
26. He pensado que no tengo amigos (F-rc)	0.84	1.12	1.19	0.45
27. Me he sentido infeliz (S-d)	1.16	1.17	0.75	-0.40
28. Me han angustiado imágenes o recuerdos no deseados (S-t)	1.41	1.24	0.48	-0.86
29. Me he sentido irritable cuando estaba con otras personas (F-rs)	1.52	1.08	0.19	-0.81
30. He pensado que yo tengo la culpa de mis problemas y dificultades (S-d)	1.93	1.19	1.25	-0.94
31. Me he sentido optimista sobre de mi futuro (B)	1.14	1.10	0.69	-0.39
32. He conseguido las cosas que quería (F-g)	1.38	0.97	0.41	-0.25
33. Me he sentido humillado(a) o avergonzado(a) por otras personas (F-rs)	0.86	1.03	1.19	0.76
34. Me he hecho daño físicamente o he puesto en peligro mi salud (R-s)	0.22	0.63	3.50	13.31

Nota. B = Bienestar; F-g = Funcionamiento general, F-rc = Funcionamiento en relaciones cercanas, F-rs = Funcionamiento en relaciones sociales; S-a = Síntomas/problemas ansiedad, S-d = Síntomas/problemas depresión, S-f = Síntomas/problemas físicos, S-t = Síntomas/problemas trauma; R-o = Riesgo de actos de agresión hacia otros, R-s = Riesgo de daño hacia sí mismo.

Tabla 2

Matriz de cargas factoriales, comunalidades y varianza explicada de cada uno de los factores.

	Ítems	Factores			Comunalidad
		1	2	3	
17	Me he sentido agobiado(a) por mis problemas (B)	.78			.62
28	Me han angustiada(a) imágenes o recuerdos no deseados (S)	.73			.54
23	Me he sentido desesperado(a) o sin esperanza (S)	.73			.59
11	La tensión y la ansiedad me han impedido hacer cosas importantes (S)	.72			.53
33	Me he sentido humillado(a) o avergonzado(a) por otras personas	.68			.47
13	Me han inquietado pensamientos y sentimientos no deseados (S)	.68			.47
15	He sentido pánico o terror (S)	.68			.47
1	Me he sentido muy solo(a) y aislado(a) (F)	.66			.47
14	He tenido ganas de llorar (B)	.66			.43
25	Me he sentido criticado(a) por los demás (F)	.65			.44
5	Me he sentido totalmente sin energía ni entusiasmo (S)	.64			.43
27	Me he sentido infeliz (S)	.64			.42
29	Me he sentido irritable cuando estaba con otras personas (F)	.64			.41
20	Me ha sido imposible dejar a un lado mis problemas (S)	.62			.39
26	He pensado que no tengo amigos (F)	.60			.40
30	He pensado que yo tengo la culpa de mis problemas y dificultades (S)	.57			.33
10	Me ha costado mucho hablar con la gente (F)	.56			.33
2	Me he sentido tenso(a), ansioso(a) o nervioso(a) (S)	.56			.34
18	He tenido dificultad para conciliar el sueño o permanecer dormido(a) (S)	.55			.30
8	He tenido molestias, dolores y otros problemas físicos (S)	.41			.19
12	Me he sentido satisfecho(a) con las cosas que he hecho (F)		.80		.64
31	Me he sentido optimista sobre mi futuro (F)		.78		.63
32	He conseguido las cosas que quería (F)		.78		.62
4	Me he sentido bien conmigo mismo(a) (B)		.76		.60
7	Me he sentido capaz de afrontar las cosas cuando han ido mal (F)		.69		.48
21	He sido capaz de hacer la mayoría de las cosas que tenía que hacer (F)		.57		.33
3	He sentido que tenía a alguien en quien apoyarme cuando lo he necesitado (F)		.45		.23
19	He sentido afecto o cariño por alguien (F)		.43		.21
9	He pensado en hacerme daño a mí mismo(a) (R)			.80	.64
24	He pensado que sería mejor que estuviera muerto(a) (R)			.78	.63
16	He hecho planes para acabar con mi vida (R)			.67	.47
34	Me he hecho daño físicamente o he puesto en peligro mi salud (R)			.63	.42
22	He amenazado o intimidado a otra persona (R)			.29	.10
6	He sido violento(a) físicamente hacia los demás (R)			.25	.10
	Varianza explicada	36%	7%	5%	

Nota. B = Bienestar, F = Funcionamiento, S = Síntomas/problemas, R = Riesgo.

Análisis factorial confirmatorio

Posteriormente, se realizó un análisis factorial confirmatorio (AFC) de los ítems mediante mínimos cuadrados no ponderados para ratificar el modelo que se obtuvo en el análisis exploratorio. Al retomar las propuestas de la estructura de la escala original, se realizaron comparaciones entre tres modelos de dos, tres y cuatro factores. Según puede observarse en la Tabla 3, en los modelos de dos y cuatro factores, los ajustes en el índice RMR están lejos de los criterios recomendados ($\leq .05$). El ajuste es más adecuado en el modelo 2 de tres factores, lo que confirma que es el que mejor se ajusta a los datos ($\chi^2_{(524)} = 1325.33$). Puede observarse que los valores de GFI, AGFI y NFI se sitúan por encima del valor de .90, mientras que el valor de RMR se encuentra en .05. Las cargas factoriales de los reactivos fueron superiores a .30. La representación gráfica del modelo, junto con los valores de interrelación entre las variables, se aprecia en la Figura 1.

En la Tabla 4 se recogen la media, la desviación estándar y el valor alfa de Cronbach para las dimensiones resultantes del CORE - OM, junto con las correlaciones entre estas, las diferencias de género y las diferencias entre consultantes y no consultantes. La consistencia interna fue de .93 para la primera dimensión, .86 para la segunda, .74 para la tercera y .94 para la escala total. La media para la dimensión de malestar emocional fue de 28.49 (DE = 15.05), para bienestar emocional de 9.31 (DE = 5.86) y para riesgo de 2.0 (DE = 2.92).

Tabla 3

Índices de ajuste de los modelos factoriales probados para la Escala CORE-OM.

Modelo factorial	χ^2	gl	GFI	AGFI	RMR	NFI
Modelo 1 (2 factores)	1655.20 ^a	526	.97	.97	.07	.97
Modelo 2 (3 factores)	1325.33 ^a	524	.98	.98	.05	.98
Modelo 3 (4 factores)	1838.43 ^a	521	.98	.97	.07	.97

Nota. GFI = Índice de bondad de ajuste; AGFI = Índice de bondad de ajuste corregido según los grados de libertad; RMR = Raíz cuadrada de los residuos al cuadrado; NFI = Índice de ajuste normativo. ^a: $p < .001$

La correlación r de Spearman mostró intercorrelaciones fuertes y positivas entre las tres dimensiones, con valores de entre .43 y .62 ($p \leq .01$).

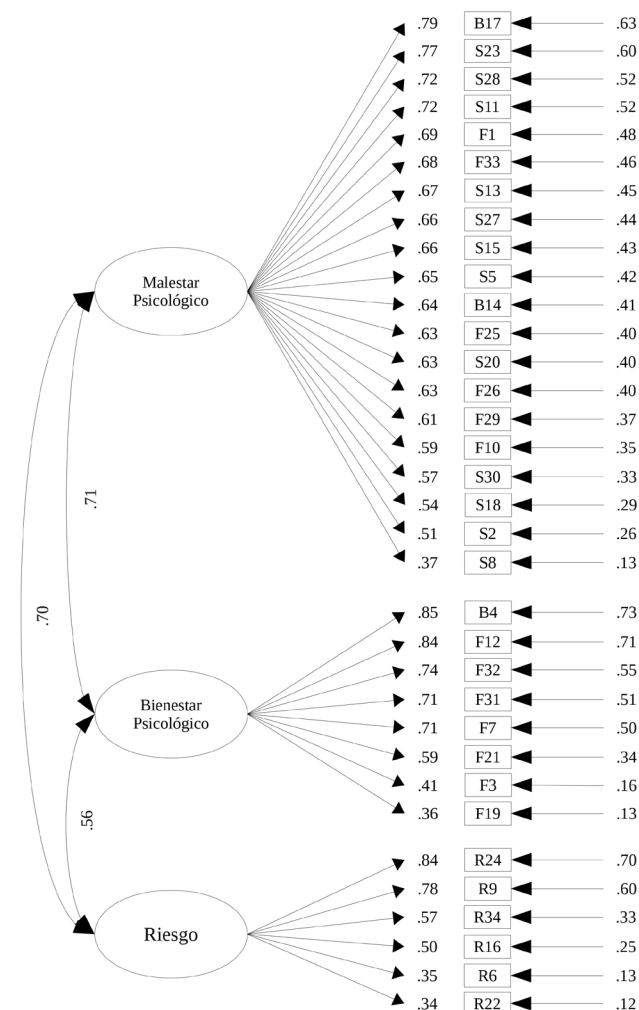


Figura 1

Modelo factorial confirmatorio de la escala CORE-OM.

Nota. B = Bienestar, F = Funcionamiento, S = Síntomas/problemas, R = Riesgo.

Tabla 4

Estadísticos descriptivos, valores de fiabilidad alfa de Cronbach, valores de correlación para las dimensiones del CORE-OM y valores *t*.

Dimensión	Sexo										Búsqueda de atención					
	Hombres					Mujeres					Buscaron			No buscaron		
	M	DE	α	Bienestar emocional	Riesgo	M	DE	M	DE	<i>t</i>	M	DE	M	DE	<i>t</i>	
Malestar emocional	28.49	15.05	.93	.62**	.53**	25.54	14.74	30.43	15.00	-3.45**	36.64	15.94	23.33	11.87	-9.67**	
Bienestar emocional	9.31	5.86	.86	-	.43**	9.29	5.94	9.31	5.82	-0.32	11.99	5.91	7.60	5.15	-8.21**	
Riesgo	2.0	2.92	.74	-	-	1.98	2.96	2.00	2.92	-0.71	3.07	3.69	1.31	2.04	-5.86**	

Nota. ** $p \leq .01$

En el análisis de las diferencias de género, los resultados indican que las mujeres puntúan con una media mayor que los hombres en la dimensión de malestar emocional y que esta diferencia es estadísticamente significativa ($p \leq .01$), en tanto que los que buscaron atención puntuaron con una media significativamente mayor en todas las dimensiones con un nivel de significación de $p \leq .01$.

Análisis del funcionamiento diferencial de los ítems

En la Tabla 5 se muestran los resultados del análisis DIF para cada ítem. El análisis reveló que cinco ítems del CORE-OM mostraron un funcionamiento diferencial en función del género. Los ítems 2 (*Me he sentido tenso(a), ansioso(a) o nervioso(a)*), 14 (*He tenido ganas de llorar*) y 17 (*Me he sentido agobiado(a) por mis problemas*) favorecen al grupo focal, es decir, a los hombres. Los reactivos 19 (*He sentido afecto o cariño por alguien*) y 30 (*He pensado que yo tengo la culpa de mis problemas*) favorecieron a las mujeres (grupo de referencia), mientras que tres reactivos tuvieron funcionamiento diferencial en función de solicitar atención psicológica. Los reactivos 2 (*Me he sentido tenso(a), ansioso(a) y nervio-*

so(a)) y 5 (*Me he sentido totalmente sin energía y ni entusiasmo*) favorecieron a los consultantes (grupo de referencia) y el reactivo 4 (*Me he sentido bien conmigo mismo(a)*) a los no consultantes (grupo focal).

Discusión

Mediante el presente estudio, examinamos la estructura dimensional y otras propiedades psicométricas de la Escala de Resultados Clínicos de Evaluación Psicológica (CORE-OM) en población mexicana. Esta escala no solo valora el malestar psicológico, sino que fue desarrollada con el objetivo de evaluar los resultados de las psicoterapias. Un instrumento con adecuadas propiedades psicométricas que provea una interpretación teórica coherente en un contexto de uso bien delimitado permite su utilización más efectiva en la práctica clínica y en la investigación psicopatológica. Dada la importancia de lo anterior, se examinó la estructura interna del CORE-OM, se estimó la fiabilidad de las puntuaciones y se obtuvieron diferentes evidencias de validez a partir de una muestra de estudiantes universitarios. Este objetivo permite actualizar las propiedades métricas del CORE-OM para su uso en población de jóvenes, población perteneciente a un grupo de

Tabla 5
Resultados del funcionamiento diferencial del ítem (DIF).

	Sexo			Solicitó atención		
	$M - X^2$	LOR_{L-A}	Z_{LOR}	$M - X^2$	LOR_{L-A}	Z_{LOR}
<i>Malestar emocional</i>						
Ítem 1	0.69	-0.17	-0.78	0.01	0.02	0.09
Ítem 2	14.91	-0.83	-3.39	6.70	0.62	2.61
Ítem 5	4.59	-0.45	-2.04	10.35	0.77	3.29
Ítem 8	6.41	-0.52	-2.43	6.20	0.56	2.45
Ítem 10	0.03	0.04	0.18	1.77	0.3	1.28
Ítem 11	0.17	-0.1	-0.43	0.22	-0.11	0.49
Ítem 13	0.34	0.12	0.57	0.03	0.04	0.18
Ítem 14	56.02	-1.72	-6.96	1.24	-0.25	-1.12
Ítem 15	0.43	-0.15	-0.70	1.54	0.29	1.32
Ítem 17	13.95	-0.81	-3.59	0.05	-0.05	-0.24
Ítem 18	0.07	-0.05	-0.25	0.40	-0.14	-0.64
Ítem 20	0.29	-0.12	-0.55	0.08	0.07	0.29
Ítem 23	0.26	-0.12	-0.53	0.07	-0.06	-0.26
Ítem 25	0.53	0.15	0.71	0.45	-0.15	-0.67
Ítem 26	0.19	-0.10	-0.44	4.84	-0.57	-2.24
Ítem 27	2.82	0.39	1.72	0.50	0.17	0.78
Ítem 28	0.08	-0.06	-0.29	0.52	-0.16	-0.73
Ítem 29	1.68	-0.27	-1.31	1.63	-0.29	-1.31
Ítem 30	9.41	0.63	3.08	0.22	-0.10	-0.48
Ítem 33	4.02	0.45	1.90	2.69	-0.39	-1.58
<i>Bienestar emocional</i>						
Ítem 3	5.11	0.49	2.44	0.14	-0.08	-0.38
Ítem 4	0.09	-0.06	-0.31	8.27	-0.70	-2.88
Ítem 7	0.00	0.01	0.05	3.46	-0.46	-1.88
Ítem 12	1.09	0.24	1.11	0.38	-0.14	-0.61
Ítem 19	15.02	0.86	3.82	0.35	0.14	0.56
Ítem 21	3.62	0.41	1.87	0.29	0.12	0.56
Ítem 31	3.54	0.41	1.89	1.13	0.24	1.04
Ítem 32	4.04	0.42	1.95	2.13	0.35	1.46
<i>Riesgo</i>						
Ítem 6	0.32	0.13	0.56	0.01	-0.02	-0.08
Ítem 9	0.00	0.01	0.05	0.56	-0.29	-0.78
Ítem 16	1.81	0.58	1.33	0.48	0.33	0.84
Ítem 22	1.50	0.38	1.22	0.00	0.01	0.04
Ítem 24	0.70	0.32	0.89	0	-0.00	-0.01
Ítem 34	0.26	0.22	0.56	0.28	0.24	0.52

Nota. M - X2: Chi-cuadrado de Mantel, LORL-A: Índice de probabilidad acumulativa de Liu-Agresti de Liu y Agresti, ZLOR: Transformación Z de LORL-A

edad de especial vulnerabilidad en su salud mental (Auerbach et al., 2016).

Procediendo con análisis factorial exploratorio y confirmatorio, los datos muestran que la estructura de tres factores es la que mejor se ajusta para el CORE-OM y es similar a la sugerida en la versión original (Evans et al., 2002). Este modelo está justificado psicométricamente y desde una perspectiva teórica y de congruencia de significado y contenido. El primer factor hace referencia a lo que se ha definido como *malestar psicológico*, o las alteraciones psíquicas resultantes de la interpretación que la persona hace del mundo y su manera particular de resolver un conflicto externo y/o interno (Espíndola-Hernández et al., 2006). El segundo factor puede interpretarse como indicadores de *bienestar psicológico*, es decir, de la expresión del desarrollo de las capacidades y el crecimiento personal en un sentido de funcionamiento positivo (Díaz et al., 2006). Y el tercer factor hace referencia a los pensamientos, conductas o planeaciones de daño hacia sí mismo u otros, es decir, a *conductas de riesgo* (Vargas-Fallas & Prada-Villalobos, 2013).

Una debilidad psicométrica del instrumento consiste en la ambigüedad factorial en el AFE de los reactivos 6 (*He sido violento(a) físicamente hacia los demás*) y 22 (*He amenazado o intimidado a otra persona*); ambos ítems fueron elaborados para medir actos de agresión hacia otros en la subescala de riesgo. En este estudio, ambos saturaban con cargas factoriales de entre .20 y .30 tanto en el factor de malestar emocional como en el de riesgo. Estos resultados evidencian que, además de tener saturaciones pobres, estos reactivos son los que menos se relacionan con el conjunto de la dimensión de riesgo. En esta dimensión se exploran dos clasificaciones de problemas: el autodaño como un problema internalizado y el daño hacia otros como un problema externalizado. Esta

dualidad podría influir en los resultados obtenidos. Estos pesos factoriales sugieren hacer más análisis que ayuden a explicar la funcionalidad de estos ítems.

Respecto a la confiabilidad de la escala, considerando que en esta puede influir el número de reactivos del instrumento y la variabilidad de las personas a quienes se aplicó, la consistencia interna de las dimensiones y de la escala total ha sido adecuada con valores superiores a .70 (Campo-Arias & Oviedo, 2008), es decir, la escala muestra adecuada confiabilidad al aplicarla sobre esta muestra de jóvenes universitarios. Las correlaciones entre las dimensiones del CORE-OM resultaron significativas en todos los casos. Estos resultados muestran que a valores altos de una dimensión corresponden valores altos en la otra, valores de asociación que son lógicos y coherentes con el planteamiento teórico: un sujeto con malestar psicológico mostrará bajos niveles de bienestar.

En los resultados encontrados respecto a las diferencias estudiadas en las dimensiones en función del sexo, son las mujeres quienes informan un mayor malestar emocional. El análisis del DIF permite establecer si esas diferencias encontradas son parte legítima del constructo que se mide en el instrumento o, por el contrario, se deben a otras variables. Para comprender mejor la causa de sesgo en la respuesta de hombres y mujeres, cabe realizar un análisis de contenido de los ítems con DIF. Los ítems que favorecen a las mujeres plantean: *He sentido afecto o cariño por alguien*, *He pensado que yo tengo la culpa de mis problemas y dificultades*, los contenidos de estos hacen referencia a sintomatología que afecta el funcionamiento en relaciones cercanas, en el caso del primer reactivo, y a sintomatología depresiva en el segundo reactivo. Mientras que los ítems que favorecen a los hombres se refieren a: *Me he sen-*

tido tenso, ansioso o nervioso, He tenido ganas de llorar, Me he sentido agobiado por mis problemas, el primer reactivo hace referencia a sintomatología ansiosa y los otros dos a sintomatología que afecta el bienestar subjetivo. Dada la propia formulación de los ítems, donde independientemente del sexo de los jóvenes se les pregunta por su estado de malestar psicológico general, cabría esperar un comportamiento similar en hombres y mujeres. Sin embargo, debido a la infiltración de factores psicosociales de la cultura, puede suceder que las mujeres tiendan a puntuar más alto en cuestiones de externalización de síntomas de malestar emocional (Berenzon, Lara, Robles, & Medina-Mora, 2013). Por ejemplo, se ha visto que las mujeres informan experimentar prevalencias más altas de problemas de salud mental con síntomas más graves y discapacitantes (Freeman & Freeman, 2013), y en población mexicana, la sintomatología y los trastornos depresivos son más prevalentes en adolescentes (Benjet et al., 2009) y mujeres adultas (Fleiz-Bautista et al., 2012). Por el contrario, los hombres tienden a mostrar patrones internalizados de expresión emocional, siendo esta inhibida o restrictiva, lo que implica que sus emociones no son reconocidas o están enmascaradas por otras más aceptables por la condición de género (Jansz, 2000). Esto podría explicar el DIF hallado en esos reactivos. Se requiere estudiar a mayor detalle la manera en la que los hombres y las mujeres manifiestan su malestar emocional para, de esta manera, tener explicaciones más amplias sobre la funcionalidad de los ítems hallados con DIF.

Respecto a las diferencias halladas entre quienes solicitan atención y quienes no lo habían hecho hasta el momento de la aplicación del instrumento, son los consultantes quienes puntuaron más alto en todas las dimensiones, mostrando así que el malestar emocional está más presente en

aquellos que buscan atención psicológica. En relación con la evaluación del sesgo, el análisis del funcionamiento diferencial de los ítems sugiere diferencias existentes en los ítems *Me he sentido tenso(a), ansioso(a) o nervioso(a), Me he sentido totalmente sin energía ni entusiasmo y Me he sentido bien conmigo mismo(a)*. En este caso, parece lógico pensar que los jóvenes consultantes serán más propensos a presentar mayor sintomatología de ansiedad o depresión, que es lo que valoran los dos primeros reactivos, mientras que los jóvenes no consultantes puntuarán más alto en indicadores de bienestar subjetivo, que es lo que evalúa el tercer reactivo hallado con DIF.

Las diferencias encontradas tanto por sexo como por población consultante indican que el instrumento es capaz de discriminar entre grupos en los cuales se esperaba encontrar diferencias según lo publicado en los artículos de la versión original y la versión castellana (Evans et al., 2002; Trujillo-Terán, 2012), lo que confirma la generalidad del constructo bienestar como estado de la salud mental (Organización Mundial de la Salud, 2013). Así, el patrón de relación entre los indicadores del constructo fue similar, estando las medidas globales de este más relacionadas entre sí que las específicas. Los resultados de los análisis del funcionamiento diferencial ponen de manifiesto que, independientemente de que puedan existir diferencias reales en la expresión del malestar emocional entre hombres y mujeres, y entre población consultante y no consultante, es necesario realizar una revisión de los ítems que muestran un funcionamiento diferencial y complementarla con análisis cualitativos que exploren la expresión del malestar emocional en los distintos grupos comparados.

Los datos encontrados en este trabajo son preliminares y esclarecen evidencias de validez de estructura interna y confiabilidad del CORE-OM.

En la versión original, para la obtención de las características psicométricas y datos normativos de la escala, se incluyeron dos muestras: una población clínica compuesta por usuarios que se encontraban en la lista de espera o estaban ya en un proceso terapéutico en el área de salud mental, y una población no clínica, es decir, personas que no cumplían con las dos condiciones anteriores. En ambas muestras se encontraron tres factores; sin embargo, en la muestra no clínica los reactivos se agruparon de forma más clara con respecto a un componente de reactivos de reacciones negativas, otro de reacciones positivas y un último de reactivos que reflejan riesgo (Evans et al., 2002). En la validación española, se tomaron los resultados de ambas muestras en el análisis de la estructura del instrumento y se obtuvo un componente donde se situaron la mayoría de los reactivos de riesgo, y otros dos componentes en los cuales no se identificaron agrupaciones relacionadas con alguna escala específica o con el tipo de reacciones positivas o negativas (Trujillo-Terán, 2012). La validación realizada en Costa Rica mostró ocho factores en la muestra clínica y nueve en la muestra no clínica (Vargas-Fallas & Prada-Villalobos, 2013). Ante las diferencias en la estabilidad factorial y los resultados de este estudio, es importante considerar que la estructura factorial no determina por qué se genera la agrupación de los reactivos, a lo cual se suma que las diferencias encontradas podrían deberse al sesgo de información que se asocia a su aplicación en culturas diferentes (Ramada-Rodilla, Serra-Pujadas, & Delclós-Clanchet, 2013). Por ello, son convenientes estudios que indaguen al respecto y permitan obtener mayores elementos explicativos de las diferencias encontradas.

Los resultados permiten concluir que el CORE-OM es un instrumento de medida confiable que permite explorar un amplio conjunto de síntomas de malestar psicológico en jóvenes uni-

versitarios. En consonancia con la importancia de valorar la sintomatología que los estudiantes experimentan, este instrumento puede resultar útil para adecuar programas de intervención de acuerdo con sus síntomas y problemáticas, además de valorar su impacto.

Estos resultados deben ser interpretados a la luz de las siguientes limitaciones: en primer lugar, los resultados encontrados pueden estar influidos por las características del muestreo por conveniencia (jóvenes universitarios), lo cual no permite generalizarlos a otras poblaciones de interés; incluso sería útil ampliar y variar la muestra y realizar otros estudios para obtener mayores evidencias de validez y funcionalidad y para aportar más información sobre la estabilidad de la estructura del instrumento. Entre otras evidencias, se considera incluir la relación del instrumento con otras mediciones del mismo constructo (validez convergente), la relación con mediciones de constructos diferentes con el objetivo de mostrar que efectivamente puede realizarse una distinción entre ellos (evidencia discriminante) y la capacidad del instrumento para predecir distintos resultados relevantes (evidencia de predicción de criterio). En segundo lugar, están los problemas inherentes a este tipo de estudios basados en autoinformes, con la posible falta de comprensión de ciertos ítems o la poca capacidad de introspección por parte de los estudiantes, por lo que puede resultar útil emplear otras medidas, por ejemplo, entrevistas, para analizar con mayor profundidad la percepción que tienen los jóvenes sobre el malestar emocional.

Referencias

Auerbach, R., Alonso, J., Axinn, W., Cuijpers, P., Ebert, D.,

- Greif, J., ... Bruffaerts, R. (2016). Mental disorders among college students in the World Health Organization World Mental Health Surveys. *Psychological Medicine*, 46(14), 2955-2970. doi: [10.1017/S0033291716001665](https://doi.org/10.1017/S0033291716001665)
- Barkham, M., Gilbert, N., Connell, J., Marshall, C., & Twigg, E. (2005). Suitability and utility of the CORE-OM and CORE-A for assessing severity of presenting problems in psychological therapy services based in primary and secondary care settings. *The British Journal of Psychiatry*, 186(3), 239-246. doi: [10.1192/bjp.186.3.239](https://doi.org/10.1192/bjp.186.3.239)
- Beck, A., Burdett, M., & Lewis, H. (2015). The association between waiting for psychological therapy and therapy outcomes as measured by the CORE-OM. *British Journal of Clinical Psychology*, 54(2), 233-248. doi: [10.1111/bjc.12072](https://doi.org/10.1111/bjc.12072)
- Benjet, C., Borges, G., Medina-Mora, M. E., Méndez, E., Fleiz, C., Rojas, E., & Cruz, C. (2009). Diferencias de sexo en la prevalencia y severidad de trastornos psiquiátricos en adolescentes de la Ciudad de México. *Salud Mental*, 32(2), 155-163. Recuperado de https://revistasaludmental.mx/index.php/salud_mental/index
- Bentler, P. M. (1992). On the fit of models to covariances and methodology to the Bulletin. *Psychological Bulletin*, 112(3), 400-404. doi: [10.1037/0033-2909.112.3.400](https://doi.org/10.1037/0033-2909.112.3.400)
- Berenzon, S., Lara, M. A., Robles, R., & Medina-Mora, M. E. (2013). Depresión: Estado del conocimiento y la necesidad de políticas públicas y planes de acción en México. *Salud Pública de México*, 55(1), 74-80. doi: [10.1590/s0036-36342013000100011](https://doi.org/10.1590/s0036-36342013000100011)
- Branney, P., & Barkham, M. (2006). Core outcomes in psychosexual therapy: A feasibility study of the CORE-OM. *Sexual and Relationship Therapy*, 21(1), 15-26. doi: [10.1080/14681990500281414](https://doi.org/10.1080/14681990500281414)
- Byrne, B. (2009). *Structural equation modeling with AMOS: Basic concepts, applications and programming* (2ª ed.). New York, NY: Routledge.
- Campo-Arias, A., & Oviedo, H. (2008). Propiedades psicométricas de una escala: La consistencia interna. *Revista de Salud Pública*, 10(5), 831-839. doi: [10.1590/s0124-00642008000500015](https://doi.org/10.1590/s0124-00642008000500015)
- Campo-Arias, A., Herazo, E., & Oviedo, H. (2012). Análisis de factores: Fundamentos para la evaluación de instrumentos de medición en salud mental. *Revista Colombiana de Psiquiatría*, 41(3), 659-671. doi: [10.1016/s0034-7450\(14\)60036-6](https://doi.org/10.1016/s0034-7450(14)60036-6)
- Casey, P., Cowan, P., Cowan, C., Draper, L., Mwamba, N., & Hewison, D. (2017). Parents as partners: A U.K. trial of a U.S. couples-based parenting intervention for at-risk low-income families. *Family Process*, 56(3), 589-606. doi: [10.1111/famp.12289](https://doi.org/10.1111/famp.12289)
- Connell, J., Barkham, M., & Mellor-Clark, J. (2007). CORE-OM mental health norms of students attending university counselling services benchmarked against an age-matched primary care sample. *British Journal of Guidance & Counselling*, 35(1), 41-57. doi: [10.1080/03069880601106781](https://doi.org/10.1080/03069880601106781)
- CORE IMS. (2001). Reino Unido: In CORE Information Management Systems. Recuperado el 2 de febrero de 2014 de <https://www.coreims.co.uk>
- Dias-Sales, C. M., de Matos-Moleiro, C. M., Evans, C., & Gomes-Alves, P. C. (2012). Versão portuguesa do CORE-OM: Tradução, adaptação e estudo preliminar das suas propriedades psicométricas. *Revista de Psiquiatria Clínica*, 39(2), 54-59. doi: [10.1590/S0101-60832012000200003](https://doi.org/10.1590/S0101-60832012000200003)
- Díaz, D., Rodríguez-Carvajal, R., Blanco, A., Moreno-Jiménez, B., Gallardo, I., Valle, C., & van Dierendonck, D. (2006). Adaptación española de las escalas de bienestar psicológico de Ryff. *Psicothema*, 18(3), 572-577. Recuperado de <https://www.psicothema.com>
- Echevarría, H. D. (2016). *Los diseños de investigación*

- cuantitativa en psicología y educación. Río Cuarto: UniRío. Recuperado de <https://www.unrc.edu.ar>
- Elfström, M., Evans, C., Lundgren, J., Johansson, B., Hakeberg, M., & Carlsson, S. (2013). Validation of the Swedish version of the Clinical Outcomes in Routine Evaluation Outcome Measure (CORE-OM). *Clinical Psychology and Psychotherapy*, 20(5), 447-455. doi: [10.1002/cpp.1788](https://doi.org/10.1002/cpp.1788)
- Elosua-Oliden, P. (2003). Sobre la validez de los tests. *Psicothema*, 15(2), 315-321. Recuperado de <https://www.psicothema.com>
- Espíndola-Hernández, J. G., Morales-Carmona, F., Díaz, E., Pimentel, D., Meza, P., Henales, C., ... Ibarra, C. (2006). Malestar psicológico: Algunas de sus manifestaciones clínicas en la paciente gineco-obstétrica hospitalizada. *Perinatología y Reproducción Humana*, 20(4), 112-122. Recuperado de <https://www.medigraphic.com>
- Evans, C., Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J., & Audin, K. (2002). Towards a standardised brief outcome measure: Psychometric properties and utility of the CORE-OM. *The British Journal of Psychiatry*, 180(1), 51-60. doi: [10.1192/bjp.180.1.51](https://doi.org/10.1192/bjp.180.1.51)
- Evans, C., Mellor-Clark, J., Margison, F., Barkham, M., Audin, K., Connell, J., & McGrath, G. (2000). CORE: Clinical Outcomes in Routine Evaluation. *Journal of Mental Health*, 9(3), 247-255. doi: [10.1080/jmh.9.3.247.255](https://doi.org/10.1080/jmh.9.3.247.255)
- Fleiz-Bautista, C., Villatoro-Velázquez, J., Medina-Mora-Icaza, M. E., Moreno-López, M., Gutiérrez-López, M. L., & Oliva-Robles, N. (2012). Sociodemographic and personal factors related to depressive symptomatology in the Mexican population aged 12 to 65. *Revista Brasileira de Psiquiatria*, 34(4), 395-404. doi: [10.1016/j.rbp.2012.03.004](https://doi.org/10.1016/j.rbp.2012.03.004)
- Forero, C., Maydeu-Olivares, A., & Gallardo-Pujol, D. (2009). Factor analysis with ordinal indicators: A Monte Carlo study comparing DWLS and ULS estimation. *Structural Equation Modeling: A Multidisciplinary Journal*, 16(4), 625-641. doi: [10.1080/10705510903203573](https://doi.org/10.1080/10705510903203573)
- Freeman, D., & Freeman, J. (2013). *The stressed sex: Uncovering the truth about men, women, and mental health*. Oxford, Reino Unido: Oxford University.
- Guevara-Ruiseñor, E. S. (2008). Vida sexual y malestar emocional. Las contradicciones de la modernidad. *REMO*, 6(14), 2-11. Recuperado de <https://remo.ws>
- IBM Corporation. (2012). IBM SPSS Statistics for Windows (Version 21.0) [software de cómputo]. Armonk, NY: IBM.
- Jansz, J. (2000). Masculine identity and restrictive emotionality. En A. Fischer (Ed.), *Gender and Emotion. Social Psychological Perspectives*, pp. 166-186. Cambridge, Reino Unido: Cambridge University. doi: [10.1017/cbo9780511628191.009](https://doi.org/10.1017/cbo9780511628191.009)
- Juhová, D., Řiháček, T., Cígler, H., Dubovská, E., Saic, M., Černý, M., ... Evans, C. (2018). Česká adaptace dotazníku CORE-OM: Vybrané psychometrické charakteristiky. *Československá Psychologie*, 62(1), 59-74.
- López-Bárcena, J. J., & González de Cossío-Ortiz, M. G. (2005). Modelo universitario de salud integral, en la Universidad Nacional Autónoma de México. *FACMED*, 48(6), 224-231. Recuperado de <https://www.revistas.unam.mx>
- Lyne, J., Barret, P., Evans, C., & Barkham, M. (2006). Dimensions of variation on the CORE-OM. *The British Journal of Clinical Psychology*, 45(2), 185-203. Recuperado de <https://onlinelibrary.wiley.com/journal/20448260>
- Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, 22(4), 719-

748. doi: [10.1093/jnci/22.4.719](https://doi.org/10.1093/jnci/22.4.719)

- Margison, F. R., Barkham, M., Evans, C., McGrath, G., Mellor-Clark, J., Audin, K., & Connell, J. (2000). Measurement and psychotherapy. Evidence-based practice and practice-based evidence. *The British Journal of Psychiatry*, *177*(2), 123-130. doi: [10.1192/bjp.177.2.123](https://doi.org/10.1192/bjp.177.2.123)
- Mellor-Clark, J., Barkham, M., Connell, J., & Evans, C. (1999). Practice-based evidence and standardized evaluation: Informing the design of the CORE system. *European Journal of Psychotherapy and Counselling*, *2*(3), 357-374. doi: [10.1080/13642539908400818](https://doi.org/10.1080/13642539908400818)
- Navascués, A., Calvo-Medel, D., & Bombin-Martín, A. (2016). Efectos del bienestar subjetivo y psicológico en los resultados terapéuticos de un hospital de día. *Acción Psicológica*, *13*(2), 143-156. doi: [10.5944/ap.13.2.15818](https://doi.org/10.5944/ap.13.2.15818)
- Otzen, T., & Manterola, C. (2017). Técnicas de muestreo sobre una población de estudio. *International Journal of Morphology*, *35*(1), 227-232. doi: [10.4067/S0717-95022017000100037](https://doi.org/10.4067/S0717-95022017000100037)
- Organización Mundial de la Salud. (2013). Salud mental: Un estado de bienestar. Recuperado el 1 de diciembre de 2013 de <https://origin.who.int>
- Penfield, R. (2009). DIFAS 5. Differential Item Functioning Analysis System. Computer Program Exchange (Versión 1.00) [software de cómputo]. Recuperado de <https://soe.uncg.edu/academics/departments/erm/erm-software>
- Penfield, R. (2013). *DIFAS 5.0. Differential Item Functioning Analysis System. User's Manual*. Recuperado de <https://soe.uncg.edu>
- Penfield, R. D., & Algina, J. (2006). A generalized DIF effect variance estimator for measuring unsigned differential test functioning in mixed format tests. *Journal of Educational Measurement*, *43*(4), 295-312. doi: [10.1111/j.1745-3984.2006.00018.x](https://doi.org/10.1111/j.1745-3984.2006.00018.x)
- Pérez, E., & Medrano, L. (2010). Análisis Factorial Exploratorio: Bases conceptuales y metodológicas. *Revista Argentina de Ciencias del Comportamiento*, *2*(1), 58-66. Recuperado de <https://www.revistas.unc.edu.ar>
- Pérez-Padilla, M. L., Ponce-Rojo, A., Hernández-Contreras, J., & Márquez-Muñoz, B. A. (2010). Salud mental y bienestar psicológico en los estudiantes universitarios de primer ingreso de la Región Altos Norte de Jalisco. *Revista de Educación y Desarrollo*, *14*, 31-37. Recuperado de http://www.cucs.udg.mx/revistas/edu_desarrollo
- Perry, T., Barkham, M., & Evans, C. (2013). The CORE-OM and CORE-OM (SV) in secure settings: A template analysis of the experiences of male patients and their staff. *The British Journal of Forensic Practice*, *15*(1), 32-43. doi: [10.1108/14636641311299068](https://doi.org/10.1108/14636641311299068)
- Ramada-Rodilla, J. M., Serra-Pujadas, C., & Delclós-Clanchet, G. L. (2013). Adaptación cultural y validación de cuestionarios de salud: Revisión y recomendaciones metodológicas. *Salud Pública de México*, *55*(1), 57-66. doi: [10.1590/s0036-36342013000100009](https://doi.org/10.1590/s0036-36342013000100009)
- Rivera-Ledesma, A., Caballero-Suárez, N., Pérez-Sánchez, I., & Montero-López-Lena, M. (2013). SCL-90: Distrés psicológico, género y conductas de riesgo. *Universitas Psychologica*, *12*(1), 105-118. doi: [10.11144/javeriana.upsy12-1.dpgc](https://doi.org/10.11144/javeriana.upsy12-1.dpgc)
- Rosales-Pérez, J. C., Córdova-Osnaya, M., & Ramos-Clatempa, R. (2012). Ideación suicida en estudiantes mexicanos: Un modelo de relación múltiple con variables de identificación personal. *Psicología y Salud*, *22*(1), 63-74. Recuperado de <https://psicologiay-salud.uv.mx/index.php/psicysalud/index>
- Trujillo-Terán, L. A. (2012). *Análisis psicométrico preliminar de la versión en español del CORE-OM: Clinical Outcomes in Routine Evaluation-Outcome Measure* (Tesis de maestría). Universidad de Barcelona, España. Recuperado de <https://babel.banrepcultural.org>

- Uji, M., Sakamoto, A., Adachi, K., & Kitamura, T. (2012). Psychometric properties of the Japanese version of the Clinical Outcomes in Routine Evaluation-Outcome Measure. *Comprehensive Psychiatry*, 53(5), 600-608. doi: [10.1016/j.comppsy.2011.09.006](https://doi.org/10.1016/j.comppsy.2011.09.006)
- Vargas-Fallas, K., & Prada-Villalobos, S. (2013). Validación del instrumento: Resultados Clínicos en Evaluación de Rutina (CORE). *Revista Humanitas*, 10(10), 33-44. Recuperado de <https://www.unirioja.es>
- Viliū-nienė, R., Evans, C., Hilbig, J., Pakalniškienė, V., Danilevičiūtė, V., Laurinaitis, E., & Navickas, A. (2013). Translating the Clinical Outcomes in Routine Evaluation Outcome Measure (CORE-OM) into Lithuanian. *Nordic Journal of Psychiatry*, 67(5), 305-311. doi: [10.3109/08039488.2012.745599](https://doi.org/10.3109/08039488.2012.745599)
- Zhang, Y., Hu, J., Evans, C., Jin, L., Wu, M., Wang, C., ... Chen, G. (2019). Psychometric properties of the Chinese version of the Clinical Outcomes in Routine Evaluation-Outcome Measure (CORE-OM). *British Journal of Guidance & Counselling*, 1-11. doi: [10.1080/03069885.2019.1682120](https://doi.org/10.1080/03069885.2019.1682120)
-

Desarrollo de instrumentos para medir actitudes hacia la comunidad trans: Análisis preliminares

Development of Instruments to Measure Attitudes towards the Trans Community: Preliminary Analysis

Caleb Esteban ^{* 1}, Juan Aníbal González-Rivera ², Margarita Francia-Martínez ³, Zahira Lesprier ⁴

1- Ponce Health Sciences University, Ponce, Puerto Rico.

2- Ponce Health Sciences University, San Juan University Center, Puerto Rico.

3 - Universidad Carlos Albizu, San Juan, Puerto Rico.

4 - Administración de Servicios de Salud Mental y Contra la Adicción, San Juan, Puerto Rico.

Recibido: 23/10/2019 Revisado: 15/12/2019 Aceptado: 17/12/2019

Introducción
Método
Resultados
Discusión
Conclusión
Referencias

Resumen

El propósito de este estudio fue desarrollar y validar preliminarmente dos instrumentos para medir actitudes de psicoterapeutas hacia la comunidad trans. El primer instrumento mide las actitudes de psicoterapeutas hacia la comunidad transgénero y el segundo hacia la comunidad transexual. La muestra contó con 131 participantes reclutados según su disponibilidad. Los resultados confirmaron preliminarmente que ambas escalas poseen una estructura unidimensional. El índice de consistencia interna alfa de Cronbach para la primera escala fue de .75 y de .85 para la segunda. No se encontraron diferencias significativas en las actitudes hacia la comunidad transgénero y transexual entre estudiantes y profesionales de la psicología. Concluimos con que la salud de las personas trans es un derecho humano y universal, por tanto, es responsabilidad de todo psicoterapeuta educarse sobre el tema y es deuda de las instituciones educativas el integrarlo en sus currículos.

Palabras clave: *transgénero, transexual, psicoterapia, actitudes, validación de escala*

Abstract

The purpose of this study was to develop and preliminarily validate two instruments to measure attitudes of psychotherapists towards the trans community. The former instrument measures the attitudes of psychotherapists towards the transgender community and the latter, towards the transsexual community. The sample consisted of 131 participants recruited according to their availability. The results preliminarily confirmed that both scales have a one-dimensional structure. Cronbach's alpha internal consistency index for the first scale was .75 and .85 for the second. There were no significant differences in attitudes towards the transgender and transsexual community between students and psychology professionals. We conclude that the health of trans people is a human and universal right, therefore, it is the responsibility of all psychotherapists to educate themselves about the issue and it is a debt of the educational institutions to integrate it into their curriculum.

Key words: *transgender, transsexual, psychotherapy, attitudes, scale validation*

Nota de autor: Expresamos que no hubo conflictos de intereses al redactar el manuscrito. Este artículo es parte del Subcomité de Investigación del Comité de la Diversidad de Sexo, Género y Orientación Sexual de la Asociación de Psicología de Puerto Rico.

* **Correspondencia a:** Caleb Esteban, Ph.D., Clinical Psychology Program, School of Behavioral and Brain Sciences, Ponce Health Sciences University, PO BOX 7004, Ponce, Puerto Rico 00732-7004. **Email:** cesteban@psm.edu

Cómo citar este artículo: Esteban, C., González-Rivera, J. A., Francia-Martínez, M., & Lesprier, Z. (2020). Desarrollo de instrumentos para medir actitudes hacia la comunidad trans: Análisis preliminares. *Revista Evaluar*, 20(1), 67-82. Recuperado de <https://revistas.unc.edu.ar/index.php/revaluar>

Introducción

Las *actitudes* son respuestas naturales del ser humano (Stangor, 2012). Estas provienen primeramente de una creencia que brinda valores positivos o negativos a un fenómeno social. Por consiguiente, las actitudes son un reflejo de estas creencias que incluye un componente emocional hacia algo o alguien (Lilienfeld, Lynn, Namy, & Woolf, 2011). Las actitudes juegan un papel vital en la manera en que interactuamos con nuestro entorno. Las mismas han sido objeto de estudio del campo de la psicología social; disciplina que estudia, entre otras cosas, cómo lo cognitivo, lo conductual y lo afectivo interactúan y producen lo que conocemos como experiencia humana.

Con el paso del tiempo, los seres humanos desarrollamos un conjunto de conocimientos que contienen información acerca del yo, de las otras personas, de los grupos sociales, y de los modos de interactuar. Dentro de ese conjunto de conocimientos, se encuentran los esquemas y las actitudes. Los *esquemas* se refieren mayormente a información acerca de una persona o grupo, mientras que las actitudes se refieren mayormente a una representación que dice si nos gusta o no una persona o grupo social. Cada persona crea unos esquemas que culminan en unas actitudes, esto nos permite pasar juicio rápido sobre si una cosa nos gusta o no, si nos asusta o no, si debemos evadirla o acercarnos, etc. Entonces, las personas pueden interpretar un mismo evento de maneras muy diferentes. Esto hace que la experiencia humana sea interesante, pero también puede acarrear conflictos entre las personas. Entonces, tanto los esquemas como las actitudes influyen en cómo procesamos a nivel cognoscitivo la información social y en cómo nos comportamos (Stangor, 2012).

Según Stangor (2012), el afecto también desempeña un papel importante en la experiencia

social humana. Dentro del campo de la psicología social, el afecto hace referencia a los sentimientos que las personas experimentamos en nuestra vida cotidiana. A medida que transcurre el día, podemos experimentar sentimientos de bienestar, de tristeza, de agradecimiento, etc. Si mantenemos nuestros sentimientos regulados adecuadamente, pueden ejercer una función adaptativa, y alertarnos cuando estamos en una situación de potencial peligro. Cuando nos sentimos felices, podemos sentirnos inclinados a interactuar con otros, pero cuando nos sentimos molestos podemos atacar o alejarnos de los demás. De hecho, cuando estamos de buen humor estamos más dispuestos a ayudar a otros y nuestra creatividad aumenta.

Dentro de este marco teórico, las actitudes poseen tres componentes: el afectivo, el cognitivo y el conductual. Estos componentes entran en relación con el entorno sociocultural en el que se desenvuelven las personas y están íntimamente ligados a la manifestación de actitudes que exhibimos hacia los demás, incluyendo a los grupos marginados. Hay algunas actitudes cuyo componente afectivo puede ser mayor, mientras que en otras se puede percibir una prominencia de los aspectos cognitivos o de los comportamentales. Las actitudes también pueden variar su intensidad, aspecto que, en ocasiones, ayuda a predecir nuestras conductas. De ahí que se utilice también este principio para planificar estrategias de mercadeo y publicidad. Si se logra aumentar el sentimiento de bienestar hacia algo, es más probable que eso influya en nuestro comportamiento hacia ese objeto o persona. Si trasladamos estos principios al campo de la psicología clínica, las actitudes positivas hacia una persona o grupo podrían conllevar que nos acerquemos de manera positiva hacia los mismos (Stangor, 2012).

La comunidad Trans

La diversidad de las identidades de género es un tema muy discutido en la actualidad (Rodríguez-Madera, Ramos-Pibernus, Padilla, & Varas-Díaz, 2015). Como fenómeno de estudio, se ha encontrado que existen actitudes negativas hacia personas no-cisgénero o personas que no se identifican con el género asociado al sexo que se les asignó al nacer (Asociación de Psicología de Puerto Rico [APPR], 2014; Ramos-Pibernus, Rodríguez-Madera, Padilla, Varas-Díaz, & Vargas-Molina, 2016).

Bajo la sombrilla de lo que conocemos como trans—término con que nombramos a las categorías diversas entre los constructos tradicionales de los binomios sociales del sexo y/o el género—tenemos diversas categorías o identidades, tales como: la transgénero, la transexual, la androginia, el género no-binario, el transformismo, el travestismo y, en ocasiones, lo intersexual (APPR, 2014). Muchas de estas categorías han sido establecidas desde centros de investigación con fines analíticos, didácticos y/o categóricos, mientras que otras surgen desde la misma comunidad. Dentro de este grupo de investigación, las comunidades transgénero y transexuales de Puerto Rico han sido el mayor enfoque de estudio y se han encontrado actitudes negativas predominantes en la sociedad (Ramos-Pibernus, 2016; Rodríguez-Madera, & Toro-Alfonso, 2002, 2003). Sin embargo, cabe destacar que son pocas las investigaciones que se han hecho en esta comunidad y especialmente en el caso de los hombres trans (Ramos-Pibernus et al., 2016; Ramos-Pibernus, 2016).

La comunidad transgénero y la transexual tienen en común la transformación de los conceptos binomiales de nuestra sociedad. Las personas transgénero son aquellas cuya identidad de género, conducta o expresión de género difiere del género

relacionado al sexo que les fue asignado en el nacimiento. Las personas transexuales son aquellas que nacieron con una anatomía sexual con la que no se identifican (total o parcialmente), por lo que deciden transformar su corporalidad para experimentar un sentido de congruencia entre la percepción de su género y sus características sexuales primarias y secundarias (APPR, 2014).

Las actitudes negativas se pueden explicar a partir de la tendencia transnegativista de la sociedad en que vivimos (LGBT Advisory Committee, 2011). Esto implica que muchas personas sienten o pueden sentir miedo, odio o demostrar actitudes de prejuicio hacia las personas trans. El desconocimiento, la confusión (Rodríguez-Madera et al., 2015) y la invisibilización de la comunidad (McCann & Sharek, 2014; Ramos-Pibernus, 2016; Rodríguez-Madera, 2012) son también factores que se añaden al panorama desfavorable. Estas actitudes negativas se manifiestan constantemente en discriminación, aislamiento, hostigamiento o diversos tipos de violencia padecidos por los integrantes de estas comunidades (LGBT Advisory Committee, 2011). A esta visión social se añade el hecho de que muchos profesionales de la psicología y otras disciplinas interpretan la diversidad de la identidad de género como una patología clínica (Jorge, 2016).

Antecedentes de medición

En Argentina, Páez, Hevia, Pesci y Rabbia (2015) construyeron una escala que mide actitudes negativas hacia personas trans. La primera parte de este estudio contó con 203 participantes, en su mayoría estudiantes universitarios, y en ella se llevaron a cabo la construcción del instrumento y análisis exploratorios. La segunda parte contó con 408 participantes de la población general seleccionados mediante muestreo por cuotas. En esta

se hicieron análisis factoriales y de validación externa del instrumento. El resultado final fue una escala de 9 reactivos, con una consistencia interna aceptable ($\alpha = .88$). Además, se encontró relación entre actitudes negativas y religiosidad, mayor edad, mayor asistencia al culto, menor educación, mayor dominancia social y mayor autoritarismo. En los Estados Unidos de América, [Norton y Herrek \(2013\)](#) llevaron a cabo un estudio de muestra probabilística con personas heterosexuales. En este estudio se utilizaron varias escalas para medir las actitudes hacia las personas transgénero y otros grupos minoritarios. Para medir las actitudes hacia la comunidad trans, utilizaron una medición de termómetro. En este tipo de medición, se solicita a los participantes que indiquen su disposición hacia varias comunidades a partir de una escala de 0 a 100, donde 0 corresponde a las actitudes más negativas y 100 a las más positivas. Estos autores encontraron que las actitudes de los hombres heterosexuales eran más negativas hacia las personas trans. Además, estas actitudes negativas se asociaban con niveles altos de autoritarismo, conservacionismo político y religiosidad. En Canadá, [Ali, Fleisher y Erickson \(2015\)](#) midieron las actitudes de psiquiatras y residentes en psiquiatría. Mediante un estudio en línea, obtuvieron una muestra de 142 participantes. Utilizaron la Escala de Ideología de Género y Transfobia (GTS) que mide sentimientos, pensamientos y conductas antitransgénero. Este cuestionario había sido publicado anteriormente ([Hill & Willoughby, 2005](#)) y traducido al español ([Carrera-Fernández, Lameiras-Fernández, Rodríguez-Castro, & Vallejo-Medina, 2014](#)). Consta de 32 reactivos que se responden mediante una escala Likert, en la cual puntuaciones más altas significan mayor intolerancia hacia la no conformidad del género. [Ali et al. \(2016\)](#) encontraron puntuaciones más bajas que en un estudio con estudiantes universitarios publicado previamente. Además, encontra-

ron relaciones asociadas a la identidad de género, ideología política y religiosidad.

En Australia, [Riggs y Sion \(2016\)](#) exploraron las actitudes de personas cisgénero, profesionales de salud mental. En tres diferentes estudios, los autores examinaron estudiantes de bachillerato y, en los dos estudios restantes, profesionales de salud mental. Las muestras fueron de 173, 122 y 304 participantes, respectivamente. Para los primeros dos estudios se utilizó la Escala de Ideología de Género y Transfobia (GTS) adaptada al contexto australiano (GTS-RA, por sus siglas en inglés). En el primer estudio con estudiantes universitarios ($\alpha = .94$), los hombres resultaron con puntuaciones más altas en los tres factores (sentimientos, pensamientos y conductas). En el segundo estudio en que participaron psicólogos, trabajadores sociales y consejeros ($\alpha = .87$), los hombres también puntuaron más alto. Por último, en el tercer estudio que contó con psicólogos, trabajadores sociales, consejeros, psiquiatras y enfermeras en salud mental, se utilizó una versión modificada del GTS-RA, llamada Escala de Actitudes de Consejeros hacia Personas Transgénero. Esta escala utiliza una escala Likert de seis puntos y cuenta con 20 reactivos ($\alpha = .92$). Este estudio igualmente encontró que las actitudes de los hombres eran más negativas.

En Estados Unidos, [Walzer \(2015\)](#) investigó las actitudes y los conocimientos sobre el estudiantado transgénero. Participaron 121 psicólogos escolares que realizaban su práctica en un ambiente escolar en los estados de New Jersey, Delaware, Connecticut y Maryland. Para este estudio se creó una escala de 20 reactivos llamada Knowledge and Attitudes Toward Transgender Student Survey ($\alpha = .78$). En los resultados no se encontró una diferencia significativa por el sexo del psicólogo escolar ni tampoco una diferencia por la edad. Sin embargo, se encontró una diferencia significativa en los resultados de aquellos participantes que es-

taban familiarizados con la posición del National Association of School Psychologists en comparación con los que no.

Propósito del estudio

Grant et al. (2011) realizaron un estudio en EEUU con 6450 personas identificadas como trans y género no-binario, y encontraron que el 86% de su muestra informó haber experimentado abuso físico y sexual, y discriminación en el empleo, acoso escolar, hostigamiento, pérdida de relaciones, pérdida de hogares y negación a ofrecerle servicios médicos. Además, la comunidad trans suele sufrir de doble discriminación, tanto por personas heterosexuales como por parte de la comunidad de lesbianas, gays y bisexuales. Esta discriminación no solo proviene de la comunidad en general, sino que también se han registrado experiencias negativas de la comunidad trans en los espacios de psicoterapia. Por lo tanto, cuando acuden a buscar ayuda, en muchas ocasiones continúan recibiendo experiencias de estigma, discriminación y transfobia (Bockting, Miner, Swinburne-Romine, Hamilton, & Coleman, 2013). Por ende, se vuelve necesario medir las actitudes hacia la comunidad trans, con el propósito de evaluar qué se debe trabajar y educar.

Actualmente, según nuestra revisión de la literatura, no existen escalas para medir las actitudes de los psicoterapeutas hacia la comunidad trans. Tomando esto en consideración, el presente trabajo tuvo como propósito desarrollar y validar preliminarmente dos instrumentos para medir actitudes de psicoterapeutas hacia la comunidad trans. Se decidió crear dos escalas diferentes, ya que las actitudes respecto a la transgeneridad y a la transexualidad pueden variar porque refieren a constructos diferentes. La transgeneridad alude a los constructos sociales del género y la

transexualidad, a los biocorporales del sexo. El primer instrumento mide las actitudes de psicoterapeutas hacia la comunidad transgénero; es decir, aquellas personas cuya identidad, expresión de género o conductas (total o parcialmente) no se ajustan a las normas y expectativas típicamente asociadas con el sexo que se les ha asignado al nacer (APPR, 2014). Por su parte, el segundo instrumento evalúa las actitudes de psicoterapeutas hacia la comunidad transexual; es decir, personas que nacieron con una anatomía sexual con la que no se sienten identificadas (total o parcialmente), por lo que deciden o desean transformar su corporalidad para experimentar un sentido de congruencia entre la percepción de su género y sus características sexuales primarias y secundarias (APPR, 2014). Con este fin, se examinó la validez de constructo y la confiabilidad de los dos instrumentos. Y, como objetivo secundario, se propuso evaluar si existían diferencias en las actitudes hacia la comunidad trans entre estudiantes y profesionales de la psicología.

Método

Diseño de investigación

Esta investigación tiene un diseño de estudio instrumental (Ato, López, & Benavente, 2013; Montero & León, 2007). Se desarrollaron dos instrumentos para medir actitudes de psicoterapeutas hacia la comunidad trans y se obtuvieron sus propiedades psicométricas preliminares.

Participantes

Se trabajó con una muestra no probabilística constituida por profesionales de la psicología ($n = 92$) y estudiantes graduados de psicología ($n = 39$), que decidieron confidencial y voluntaria-

mente participar del estudio. La edad promedio de la muestra fue de 37.27 años (DE = 12.84). En la Tabla 1 se presentan los datos sociodemográficos de los participantes. Los criterios de inclusión fueron: (a) ser mayor de 21 años de edad y (b) tener experiencia o práctica supervisada en intervenciones de psicoterapia o evaluaciones psicológicas.

Tabla 1
Datos sociodemográficos de la muestra.

Variable	f	%
Sexo		
Hombre	42	32.1
Mujer	89	67.9
Género		
Masculino	42	32.1
Femenino	88	67.2
Otro	1	0.8
Orientación sexual		
Heterosexual	104	79.4
Homosexual	19	14.5
Bisexual	5	3.8
Pansexual	1	0.8
Valores perdidos	2	1.5
Estado Civil		
Solteros	48	36.6
Casados	34	26.0
Convivencia	18	13.7
Divorciados	16	12.2
Noviazgo	13	9.9
Viudos	2	1.5
Religión		
Católica	62	47.3
Protestante	6	4.6
Evangélica	12	9.2
Otra	12	9.2

Ninguna	37	28.2
Valores perdidos	2	1.5

Título académico en Psicología

Doctorado en Psicología (PsyD)	29	22.1
Doctorado en Filosofía (PhD)	33	25.2
Maestría en Ciencias (MS)	14	10.7
Maestría en Artes (MA)	15	11.5
Estudiante graduado	39	29.8
Valores Perdidos	1	0.8

Nota. n = 131.

Instrumentos

Cuestionario de datos generales. Para identificar las características sociodemográficas de la muestra se desarrolló un cuestionario de datos generales que recogía información importante con respecto a la edad, sexo, género, estado civil, orientación sexual y religión.

Escala de Actitudes de Psicoterapeutas hacia la Comunidad Transgénero. Este instrumento fue diseñado por los investigadores con el propósito de desarrollar una justa medida que examine las actitudes de estudiantes y profesionales de la psicología hacia la comunidad transgénero. Con este fin, se desarrollaron 43 ítems que recogen varias posibles actitudes hacia la comunidad transgénero, tales como: empatía, cercanía, incomodidad, conocimiento, apertura, prejuicios, aceptación, entre otros. Los ítems fueron sometidos a la opinión de 10 jueces conocedores y estudiosos de la comunidad LGBT en Puerto Rico, con el objetivo de identificar si los ítems del instrumento son pertinentes al constructo que se desea medir (método de Lawshe). Para rechazar o retener los ítems se utilizó la razón de validez de contenido (CVR) y para interpretar los resultados, los valores crí-

ticos recalculados por [Wilson, Pan y Schumsky \(2012\)](#). De acuerdo con estos autores, el valor mínimo requerido para 10 jueces, según la tabla de Schipper, es de .62 para la aceptación de un ítem como esencial. Solo 25 ítems obtuvieron valores aceptables, por lo que permanecieron en el instrumento preliminar. La escala de respuesta del instrumento es tipo Likert de cuatro puntos: 1 (*Totalmente de acuerdo*), 2 (*Parcialmente de acuerdo*), 3 (*Parcialmente en desacuerdo*) y 4 (*Totalmente en desacuerdo*). Contiene ítems que son de corrección inversa. A mayor puntuación, mayores actitudes negativas (prejuicios) hacia personas transgénero.

Escala de Actitudes de Psicoterapeutas hacia la Comunidad Transexual. Este instrumento fue diseñado por los investigadores con el propósito de desarrollar una justa medida que examine las actitudes de estudiantes y profesionales de la psicología hacia la comunidad transexual. También se desarrollaron 43 ítems que recogen los mismos aspectos que el instrumento anterior, pero dirigidos a personas transexuales. Los ítems fueron sometidos a la opinión de 10 jueces conocedores y estudiosos de la comunidad LGBT en Puerto Rico (método de Lawshe). Para rechazar o retener los ítems se utilizaron la razón de validez de contenido (CVR) y los valores críticos recalculados por [Wilson et al. \(2012\)](#). En esta ocasión, 30 ítems obtuvieron valores aceptables, por lo que permanecieron en el instrumento preliminar. La escala de respuesta del instrumento es tipo Likert de cuatro puntos: 1 (*Totalmente de acuerdo*), 2 (*Parcialmente de acuerdo*), 3 (*Parcialmente en desacuerdo*) y 4 (*Totalmente en desacuerdo*). Igualmente, contiene ítems que son de corrección inversa. A mayor puntuación, mayores actitudes negativas (prejuicios) hacia personas transexuales.

Procedimientos

La investigación fue aprobada por el Institutional Review Board (IRB) de la Universidad Carlos Albizu de Puerto Rico (*Fall-14-09*). Luego se procedió a recopilar la información de manera presencial y en línea, a través del portal [surveymonkey.com](#). Se utilizó una hoja de consentimiento informado, en la que no era necesario identificarse con nombre, para notificar a los participantes del propósito del estudio, su naturaleza voluntaria, los posibles riesgos y beneficios, así como de su derecho a retirarse del estudio en cualquier momento. Además, se informó sobre la duración de su participación y su derecho a conocer los resultados del estudio. Los acercamientos presenciales se realizaron mediante visitas a centros de trabajo de profesionales de la psicología, así como visitas a centros universitarios donde cursan estudiantes graduados de psicología en Puerto Rico.

Análisis de datos

Una vez recopilados los datos, los analizamos utilizando el sistema para análisis estadísticos IBM SPSS v. 24 ([IBM Corp., 2016](#)). En este programa se realizaron análisis descriptivos de la muestra, análisis de ítems para conocer su índice de discriminación, análisis de factores y análisis de confiabilidad para calcular el coeficiente alfa de Cronbach y división en mitades de Spearman-Brown. Además, se realizó un análisis de correlación r de Pearson entre las dos escalas. Valores inferiores a .35 se consideran correlaciones débiles o bajas, valores entre .36 y .67 se consideran correlaciones moderadas, valores entre .68 y .89 se consideran correlaciones altas y, por último, valores de .90 en adelante se consideran correlaciones muy altas ([Taylor, 1990](#)). El tamaño

del efecto de la correlación se calculó mediante el coeficiente de determinación r^2 . Por último, se efectuó un análisis de comparación de medias prueba t con el propósito de evaluar si existen diferencias estadísticamente significativas entre el estudiantado y los profesionales de la psicología en términos de las actitudes hacia la comunidad trans. El tamaño del efecto de la prueba t se calculó con la d de Cohen.

Para el análisis factorial exploratorio, se utilizó el método de extracción de mínimos cuadrados no ponderados (*unweighted least squares*) con rotación oblicua para identificar las variables latentes que subyacen en los ítems. Se utilizó este procedimiento de ajuste por dos razones principales: (1) el método de extracción de mínimos cuadrados no ponderados es el más recomendado actualmente para trabajar con muestras pequeñas, especialmente si el número de factores a retener es pequeño (Jung, 2013) y evita la aparición de casos Heywood (saturaciones mayores que la unidad y varianzas de error negativas), más frecuentes con otros métodos de estimación (Lloret-Segura, Ferreres-Traver, Hernández-Baeza, & Tomás-Marco, 2014); y (2) la rotación oblicua es más precisa y provee más información que la rotación octogonal (Schmitt, 2011).

Para determinar el número de factores se utilizó la cantidad de varianza explicada mediante el número de factores extraídos. Hatcher (1994) sugiere que los factores a considerar sean aquellos que al menos expliquen un 5% de la varianza. Comúnmente, se considera que un análisis factorial es aceptable si la suma de las varianzas de los factores que se retienen es igual o mayor al 50% (Ferré, 1995; Streiner, 1994). No obstante, en ciertos análisis factoriales exploratorios, pueden ser admisibles soluciones factoriales que expliquen al menos el 30% de la varianza total (Macía-Sepúlveda, 2010). Como criterios de aceptación, consideramos aquellos ítems con una

carga factorial mayor a .30 en un solo factor (Kline, 2000). Para conocer la capacidad de discriminación de los ítems de los factores, se calculó el índice de correlación ítem total (*rbis*). La literatura identifica como índices adecuados de discriminación valores mayores que .30 y menores que .70 (Field, 2013). Para conocer el grado de confiabilidad de la escala, se calcularon el coeficiente alfa de Cronbach, alfa de Cronbach estandarizado y el coeficiente de división en mitades de Spearman-Brown. Field (2013) indica que índices mayores a .70 son aceptables.

Resultados

Validez de construcción lógica

Actitudes hacia la comunidad transgénero. Se realizó un primer análisis factorial exploratorio para evaluar la adecuación de los datos e identificar los factores que explican el 5% o más de la varianza para ser retenidos en la escala. La prueba de Kaiser-Meyer-Olkin apoyó la adecuación de los datos de muestreo para el análisis: KMO = .719. La prueba de esfericidad de Bartlett fue significativa ($\chi^2(300) = 728.939, p < .001$), indicando que las correlaciones entre los ítems fueron significativamente diferentes de cero y es un indicador adicional de la adecuación para el análisis factorial. Considerando los criterios de selección de ítems, se evidenció la retención de un solo factor, ya que los factores 2 en adelante explicaban menos del 5% de la varianza o no contaban con suficientes ítems. Al realizar el análisis restringido a un factor, de los 25 ítems se eliminaron 15 porque no cumplieron con los criterios de selección. Este procedimiento redujo la escala a 10 ítems que explican el 33% de la varianza de los datos originales. En la Tabla 2 se puede apreciar las cargas factoriales de esta versión de la escala.

Actitudes hacia la comunidad transexual. Se realizó un análisis factorial exploratorio para evaluar la adecuación de los datos e identificar los factores que explican el 5% o más de la varianza para ser retenidos en la escala. La prueba de Kaiser-Meyer-Olkin apoyó la adecuación de los datos de muestreo para el análisis: $KMO = .734$. La prueba de esfericidad de Bartlett fue significativa ($\chi^2(435) = 1454.164, p < .001$). Considerando los criterios de selección de ítems, se evidenció la retención de un solo factor. Al realizar el análisis restringido a un factor, de los 30 ítems se eliminaron 20 que no cumplieron con los criterios de selección. Este procedimiento redujo la escala a 10 ítems que explican el 44% de la varianza de los datos originales (véase Tabla 2).

Análisis de distribución de ítems

Con los datos obtenidos se llevaron a cabo análisis de ítems para determinar los índices de discriminación de cada instrumento. La Escala de Actitudes de Psicoterapeutas hacia la Comunidad Transgénero mostró índices de discriminación que fluctuaban entre .36 y .55. Por su parte, la Escala de Actitudes de Psicoterapeutas hacia la Comunidad Transexual evidenció índices de discriminación que fluctuaban entre .44 y .66. Todos los valores están dentro de los puntos de corte recomendados por la literatura. En la Tabla 2 se presentan los índices de discriminación por ítem de cada escala.

Análisis de confiabilidad y correlación

Los ítems de cada escala fueron sometidos a un análisis de consistencia interna para determinar el índice de confiabilidad de los instrumentos. Para esto se utilizaron dos métodos: el coefi-

ciente alfa de Cronbach y la división en mitades de Spearman-Brown. La Escala de Actitudes de Psicoterapeutas hacia la Comunidad Transgénero mostró un coeficiente alfa de .75 y de .76 para la prueba Spearman-Brown. Por su parte, la Escala de Actitudes de Psicoterapeutas hacia la Comunidad Transexual evidenció un coeficiente alfa de .85 y de .81 para la prueba Spearman-Brown. El análisis de correlación r de Pearson evidenció una asociación alta, positiva y estadísticamente significativa entre ambas escalas ($r = .70, p < .001, r^2 = .49$)

Diferencias entre profesionales y estudiantes de psicología

Se realizó un análisis de comparación de medias mediante prueba t con el propósito de evaluar si existen diferencias estadísticamente significativas entre las medias de actitudes hacia la comunidad trans entre profesionales de la psicología y estudiantes graduados de psicología (véase Tabla 4). Los análisis reflejaron que no existen diferencias estadísticamente significativas en las actitudes hacia la comunidad transgénero y transexual entre profesionales de la psicología y estudiantes graduados de psicología.

Discusión

Lamentablemente, existe suficiente evidencia como para afirmar categóricamente que la comunidad trans es blanco de un continuo rechazo, prejuicio y discrimen (Francia-Martínez, Esteban, & Lespier, 2017). El desconocimiento, la confusión y la invisibilización de la comunidad puede provocar que los profesionales de la psicología asuman actitudes negativas hacia personas transexuales y transgénero. No obstante, ni Puer-

Tabla 2

Índice de discriminación, varianza explicada de los ítems y cargas factoriales.

Escala de Actitudes de Psicoterapeutas hacia la Comunidad Transgénero				Escala de Actitudes de Psicoterapeutas hacia la Comunidad Transexual			
Ítem	<i>rbis</i>	R ²	Factor 1	Ítem	<i>rbis</i>	R ²	Factor 1
2	.55	.38	.64	30	.55	.37	.60
4	.36	.24	.44	38	.44	.25	.48
12	.38	.27	.44	41	.59	.46	.64
15	.42	.28	.51	44	.62	.47	.67
17	.40	.23	.45	47	.47	.26	.51
19	.49	.38	.59	50	.59	.45	.64
20	.44	.26	.50	52	.50	.39	.56
23	.36	.28	.41	53	.62	.48	.68
24	.42	.57	.46	54	.61	.87	.66
25	.51	.61	.58	55	.66	.88	.71

Nota. *rbis* = índice de discriminación del ítem; R² = varianza explicada de los ítems; *Factor 1* = cargas del análisis de factores exploratorio. (n = 131).

Tabla 3

Medias, desviaciones estándar, alfas y correlaciones entre las escalas.

Factor	M	DE	α	α_{est}	Spearman-Brown	1 (<i>r</i> ²)	2 (<i>r</i>)
1. Transgénero	13.25	3.96	.75	.77	.76	-	.70
2. Transexual	14.94	5.38	.85	.86	.81	.49	-

Nota. M = media; DE = desviación estándar; α = alfa de Cronbach; α_{est} = alfa de Cronbach estandarizado. La correlación sobre la diagonal representa la correlación *r* de Pearson, mientras que la correlación bajo la diagonal representa el coeficiente de determinación *r*². Todas las correlaciones fueron significativas a *p* < .001. (n = 131).

Tabla 4Resultados del análisis de comparación de medias prueba *t* y tamaño del efecto.

	N	M	DE	<i>t</i>	<i>gl</i>	<i>p</i>	<i>d</i>
<i>Actitudes hacia Transgénero</i>							
Profesionales	87	13.22	4.14	.107	122	.92	.02
Estudiantes	37	13.14	3.45				
<i>Actitudes hacia Transexual</i>							
Profesionales	89	15.10	5.68	.500	125	.62	.10
Estudiantes	38	14.58	4.66				

Nota. N = participantes; M = medias; *gl* = grados de libertad; DE = desviación estándar; *t* = valor de la prueba *t*; *p* = significancia; *d* = tamaño del efecto *d* de Cohen. (N = 131).

to Rico ni el Caribe contaban con un instrumento diseñado para medir las actitudes de los psicoterapeutas hacia la comunidad trans. Es por ello que el presente estudio tuvo como propósito desarrollar, validar y examinar las propiedades psicométricas de dos instrumentos en una muestra de profesionales y estudiantes graduados de psicología: (a) Escala de Actitudes de Psicoterapeutas hacia la Comunidad Transgénero; y (b) Escala de Actitudes de Psicoterapeutas hacia la Comunidad Transexual. Los resultados obtenidos confirman preliminarmente que ambos instrumentos poseen propiedades psicométricas adecuadas para medir actitudes favorables o desfavorables hacia las personas trans.

En términos teóricos, los análisis factoriales exploratorios revelaron la existencia de un solo factor latente en ambos instrumentos. Esto evidencia cómo el prejuicio hacia la comunidad trans, tanto hacia transexuales como hacia transgénero, es consistente en sus expresiones, formas y manifestaciones. Sorprendentemente, un 20% de los participantes evidenció puntuaciones correspondientes a un nivel alto de prejuicio hacia la comunidad trans. Esto sugiere que educarse formalmente en psicología puede colaborar en el mantenimiento de un discurso académico de inclusión, pero no asegura que en el ámbito personal –y, en ocasiones, en el laboral– los profesionales de la psicología mantengan conductas coherentes con la inclusión y la aceptación.

Por su parte, los análisis comparativos revelaron que no existen diferencias estadísticamente significativas en las actitudes hacia la comunidad trans entre estudiantes y profesionales de la psicología. Esto puede implicar, entre otras cosas, que el prejuicio hacia las personas trans no se debe a un asunto generacional ni de experiencia profesional. Este hallazgo nos da a entender que los años de experiencia en psicoterapia por sí solos no mejoran las actitudes de prejuicio hacia la co-

munidad trans en los y las psicoterapeutas. Además, la falta de capacitación en el tema en el país aporta a la falta de introspección, conocimiento y competencias clínicas de los y las psicoterapeutas para trabajar con las identidades de género (Esteban-Reyes, Ortíz-Mendoza, Rivera-Morales, Purcell-Baerga, & Ruiz-Mojica, 2016).

En cuanto a la confiabilidad de las escalas, se obtuvieron índices superiores al mínimo recomendado por la literatura científica (Field, 2013; Kline, 2000). Esto sugiere preliminarmente que tanto la Escala de Actitudes de Psicoterapeutas hacia la Comunidad Transgénero como la Escala de Actitudes de Psicoterapeutas hacia la Comunidad Transexual son instrumentos estables, reproducibles y consistentes en la medida de prejuicio hacia las personas trans. Asimismo, las correlaciones de cada ítem con la puntuación total manifiestan una consistencia interna notable. Esto sugiere que los ítems de ambos instrumentos discriminan adecuadamente y son capaces de diferenciar personas con distintos niveles de prejuicio hacia la comunidad trans.

En cuanto a las implicaciones prácticas, ambas escalas son de las primeras que se construyen en español en Latinoamérica y las primeras en Centro América y el Caribe que miden actitudes hacia la población trans. Además, son las únicas escalas en español que miden actitudes hacia esta población en el proceso de psicoterapia. Instrumentos como estos son sumamente importantes para el desarrollo de la psicología latinoamericana y el avance hacia una psicoterapia auténticamente inclusiva. Las escalas no solo miden los niveles de actitudes de los terapeutas hacia la comunidad trans, sino que también demuestran la necesidad de formación en el tema y las repercusiones en psicoterapia sobre la transignorancia (Francia-Martínez, Esteban, & Lespier, 2017). Es responsabilidad de todo terapeuta formarse en el tema y deuda de las instituciones educativas,

integrarlo en sus currículos académicos (Coleman et al., 2012). La comunidad trans no puede continuar siendo un sector marginado, juzgado ni discriminado, especialmente tratándose de profesionales de salud mental como los profesionales de la psicología.

En términos de corrección, las puntuaciones de las escalas se calculan mediante la sumatoria de los 10 ítems para obtener un índice general. En los apéndices A y B se encuentran las escalas finales con sus respectivas normas de corrección, ya que varios ítems se corrigen de forma inversa. Al finalizar la sumatoria, se clasifican las puntuaciones de acuerdo a los cuartiles. Las puntuaciones de la Escala de Actitudes de Psicoterapeutas hacia la Comunidad Transgénero se categorizan de la siguiente manera: *ningún prejuicio* (puntuación de 10), *prejuicio bajo* (de 11 a 15) y *prejuicio alto* (de 16 a 40). Las puntuaciones de la Escala de Actitudes de Psicoterapeutas hacia la Comunidad Transexual se categorizan de la siguiente forma: *ningún prejuicio* (puntuación de 10), *prejuicio bajo* (de 11 a 17) y *prejuicio alto* (de 16 a 40).

Limitaciones y recomendaciones

El estudio tuvo la limitación metodológica de ser realizado con una muestra por disponibilidad relativamente pequeña. No obstante, se tomó en consideración el tamaño de la muestra y el nivel de significancia al momento de tomar decisiones sobre los ítems, así como un método de extracción adecuado a dicha limitación (Jung, 2013). Otra limitación fue que no se pudo establecer la confiabilidad del instrumento a través del tiempo, solo se pudo hacer a través de sus elementos. Se recomienda para futuros estudios llevar a cabo una muestra probabilística para evitar cualquier sesgo en los participantes. También,

se recomiendan análisis estadísticos adicionales para mejorar los diversos tipos de confianza y validez de la escala. Además, se recomienda ampliar la muestra a otros países y otras disciplinas que ofrecen psicoterapia, terapia o consejería. Es necesario crear escalas para medir actitudes en otros profesionales de la salud mental y de la salud en general.

Conclusión

Tanto la Escala de Actitudes de Psicoterapeutas hacia la Comunidad Transgénero como la Escala de Actitudes de Psicoterapeutas hacia la Comunidad Transexual mostraron preliminarmente indicadores adecuados de validez y confiabilidad, lo que sugiere que presentan el potencial para avanzar en la investigación con miras a comprender con mayor amplitud el fenómeno del prejuicio hacia la comunidad trans entre los profesionales de la psicología. Se espera que dichos instrumentos sean beneficiosos en el ámbito de la investigación y faciliten nuevos conocimientos sobre este fenómeno en Puerto Rico y América Latina.

Referencias

- Ali, N., Fleisher, W., & Erickson, J. (2016). Psychiatrists' and psychiatry residents' attitudes toward transgender people. *Academic Psychiatry, 40*(2), 268-273. doi: 10.1007/s40596-015-0308-y
- Asociación de Psicología de Puerto Rico. (2014). *Estándares para el trabajo e intervención en comunidades lesbianas, gay, bisexuales e identidades trans (LGBT)*. San Juan, Puerto Rico: Autor.
- Ato, M., López, J. J., & Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en

- psicología. *Anales de Psicología*, 29(3), 1038-1059. doi: 10.6018/analesps.29.3.178511
- Bockting, W. O., Miner, M. H., Swinburne-Romine, R. E., Hamilton, A., & Coleman, E. (2013). Stigma, mental health, and resilience in an online sample of the US transgender population. *American Journal of Public Health*, 103(5), 943-951. doi: 10.2105/AJPH.2013.301241
- Carrera-Fernández, M. V., Lameiras-Fernández, M., Rodríguez-Castro, Y., & Vallejo-Medina, P. (2014). Spanish adolescents' attitudes toward transpeople: Proposal and validation of a short form of the Genderism and Transphobia Scale. *The Journal of Sex Research*, 51(6), 654-666. doi: 10.1080/00224499.2013.773577
- Coleman, E., Bockting, W., Botzer, M., Cohen-Kettenis, P., DeCuypere, G., Feldman, J., ... Zucker, K. (2012). Standards of care for the health of transsexual, transgender and gender-nonconforming people. *International Journal of Transgenderism*, 13(4), 165-232. doi: 10.1080/15532739.2011.700873
- Esteban-Reyes, C., Ortiz-Mendoza, C. M., Rivera-Morales, N., Purcell-Baerga, P., & Ruiz-Mojica, G. (2016). La educación del género en peligro de extinción: Preparación de psicoterapeutas clínicos en Puerto Rico. *Revista Puertorriqueña de Psicología*, 27(1), 80-93. Recuperado de <http://www.ojs.repsasppr.net/index.php/reps/index>
- Ferré, L. (1995). Selection of components in principal component analysis: A comparison of methods. *Computational Statistics & Data Analysis*, 19(6), 669-682. doi: 10.1016/0167-9473(94)00020-j
- Field, A. P. (2013). *Discovering statistics using IBM SPSS statistics* (4ª ed.). London, England: SAGE.
- Francia-Martínez, M., Esteban, C., & Lesprier, Z. (2017). Actitudes, conocimiento y distancia social de psicoterapeutas con la comunidad transgénero y transexual. *Revista Puertorriqueña de Psicología*, 28(1), 98-113. Recuperado de <http://www.ojs.repsasppr.net/index.php/reps>
- Grant, J. M., Mottet, L. A., Tanis, J., Harrison, J. L., Herman, J., & Keisling, M. (2011). *Injustice at every turn: A report of the National Transgender Discrimination Survey*. Recuperado de <https://transequality.org>
- Hatcher, L. (1994). *A step-by-step approach to using the SAS system for factor analysis and structural equation modeling*. Cary, NC: SAS Institute.
- Hill, D. B., & Willoughby, B. L. B. (2005). The development and validation of the Genderism and Transphobia Scale. *Sex Roles*, 53(7-8), 531-544. doi: 10.1007/s11199-005-7140-x
- IBM Corp. (2016). *IBM SPSS Statistics for Windows* (version 24.0). [software de cómputo] Armonk, NY: Autor.
- Jorge, J. C. (2016). El nombrar clínico: La disforia de género. *Boletín Diversidad*, 7(1), 6-9. Recuperado de <https://www.boletindiversidad.org>
- Jung, S. (2013). Exploratory factor analysis with small sample sizes: A comparison of three approaches. *Behavioural Processes*, 97(1), 90-95. doi: 10.1016/j.beproc.2012.11.016
- Kline, P. (2000). *The Handbook of Psychological Testing* (2ª ed.). New York, NY: Routledge.
- LGBT Advisory Committee & San Francisco Human Rights Commission. (2011). *Bisexual invisibility: Impacts and recommendations*. San Francisco, CA: San Francisco Human Rights Commission. Recuperado de <https://sf-hrc.org>
- Lilienfeld, S. O., Lynn, S. J., Namy, L. L., & Woolf, N. J. (2011). *Psicología: Una introducción*. Madrid, España: Pearson Education.
- Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., & Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: Una guía práctica, revisada y actualizada. *Anales de Psicología*, 30(3), 1151-1169. doi: 10.6018/analesps.30.3.199361
- Macía-Sepúlveda, F. (2010). Validez de los tests y el análisis factorial: Nociones generales. *Ciencia & Trabajo*, 12(35), 276-280. Recuperado de <https://dialnet.unirioja.es/servlet/articulo?codigo=3218921>
- McCann, E., & Sharek, D. (2014). Survey of lesbian, gay, bisexual and transgender people's experiences of

- mental health services in Ireland. *International Journal of Mental Health Nursing*, 23(2), 118-127. doi: 10.1111/inm.12018
- Montero, I., & León, O. G. (2007). A guide for naming research studies in psychology. *International Journal of Clinical and Health Psychology*, 7(3), 847-862. Recuperado de <https://aepec.es/ijchp/busca.php>
- Norton, A. T., & Herek, G. M. (2013). Heterosexuals' attitudes toward transgender people: Findings from a national probability sample of U.S. adults. *Sex Roles*, 68, 738-753. doi: 10.1007/s11199-011-0110-6
- Páez, J., Hevia, G., Pesci, F., & Rabbia, H. H. (2015). Construcción y validación de una escala de actitudes negativas hacia personas trans. *Revista de Psicología*, 33(1), 151-188. Recuperado de <http://revistas.pucp.edu.pe/index.php/psicologia>
- Ramos-Pibernus, A. G. (2016). Hombres trans en Puerto Rico: Una comunidad escondida. *Boletín Diversidad*, 7(1), 10-11. Recuperado de <http://www.boletin-diversidad.org>
- Ramos-Pibernus, A. G., Rodríguez-Madera, S. L., Padilla, M., Varas-Díaz, N., & Vargas-Molina, R. (2016). Intersections and evolution of 'Butch-trans' categories in Puerto Rico: Needs and barriers of an invisible population. *Global Public Health*, 11(7-8), 966-980. doi: 10.1080/17441692.2016.1180703
- Riggs, D. W., & Sion, R. (2016). Gender differences in cisgender psychologists' and trainees' attitudes toward transgender people. *Psychology of Men & Masculinities*, 18(2), 187-190. doi: 10.1037/men0000047
- Rodríguez-Madera, S. L. (2012). TRANS-acciones de la carne: Criminalización de mujeres trans puertorriqueñas que ejercen el trabajo sexual. En S. M. Serrano-Rivera (Ed.), *Registros Criminológicos Contemporáneos* (pp. 107-150). San Juan, Puerto Rico: Situm.
- Rodríguez-Madera, S., Ramos-Pibernus, A., Padilla, M., & Varas-Díaz, N. (2015). Radiografías de las comunidades trans en Puerto Rico: Visibilizando femineidades y masculinidades alternas. En M. Vázquez-Rivera, A. Martínez-Taboas, M. Francia-Martínez & J. Toro-Alfonso (Eds.), *LGBT 101: Una mirada introductoria al colectivo* (pp. 315-342). Hato Rey, Puerto Rico: Publicaciones Puertorriqueñas.
- Rodríguez-Madera, S., & Toro-Alfonso, J. (2002). Ser o no ser: La transgresión del género como objeto de estudio de la psicología. *Avances en Psicología Clínica Latinoamericana*, 22, 63-78.
- Rodríguez-Madera, S., & Toro-Alfonso, J. (2003). La comunidad de la cual no hablamos: Vulnerabilidad social, conductas de riesgo y VIH/SIDA en la comunidad de transgénero en Puerto Rico. *Revista Puertorriqueña de Psicología*, 14(1), 7-40. Recuperado de <http://www.ojs.repsasppr.net/index.php/reps/index>
- Schmitt, T. A. (2011). Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment*, 29(4), 304-321. doi: 10.1177/0734282911406653
- Stangor, C. (2012). Social Psychology Principles. Recuperado de <https://opentextbc.ca/socialpsychology/#-main>
- Streiner, D. L. (1994). Figuring out factors: The use and misuse of factor analysis. *The Canadian Journal of Psychiatry*, 39(3), 135-140. doi: 10.1177/070674379403900303
- Taylor, R. (1990). Interpretation of the correlation coefficient: A basic review. *Journal of Diagnostic Medical Sonography*, 6(1), 35-39. doi: 10.1177/875647939000600106
- Walzer, A. (2015). *School psychologists' knowledge of and attitudes towards transgender students* (Tesis de maestría). Recuperado de <https://rdw.rowan.edu/etd>
- Wilson, F. R., Pan, W., & Schumsky, D. A. (2012). Recalculation of the critical values for Lawshe's content validity ratio. *Measurement and Evaluation in Counseling and Development*, 45(3), 197-210. doi: 10.1177/0748175612440286

Apéndice A

Escala de Actitudes de Psicoterapeutas hacia la Comunidad Transgénero

	Totalmente de acuerdo	Parcialmente de acuerdo	Parcialmente en desacuerdo	Totalmente en desacuerdo
1. Si puedo escoger, evito tener que atender a clientes transgénero.				
2. Se me hace difícil hablar de las relaciones sexuales o de pareja con un/a cliente transgénero.				
3. Creo inadecuado que una persona transgénero asista a una actividad donde haya niños/as.				
4. Reconozco que la transgresión del género reta mis valores personales.				
5. Me sentiría tranquilo/a si mis hijos/as tuvieran un maestro/a transgénero.				
6. Entiendo que darles derechos y reconocimiento a las personas transgénero trastoca los valores de nuestra sociedad.				
7. Considero que las personas transgénero que deciden salir a la calle vestida/os del otro género, son ejemplo de superación y valentía.				
8. Considero que las personas transgénero son inestables.				
9. Compartiría mi oficina con un/a terapeuta transgénero.				
10. Estoy de acuerdo que las personas transgéneros, puedan cambiar todos sus documentos, con su identidad de género nueva.				

Método de corrección sumaria

Puntuación regular: los ítems se puntúan de 1 (*Totalmente de acuerdo*) a 4 (*Totalmente en desacuerdo*).

Puntuación inversa: los ítems 1, 2, 3, 4, 6 y 8 se puntúan de forma inversa, de 4 (*Totalmente de acuerdo*) a 1 (*Totalmente en desacuerdo*).

Apéndice B

Escala de Actitudes de Psicoterapeutas hacia la Comunidad Transexual

	Totalmente de acuerdo	Parcialmente De acuerdo	Parcialmente en desacuerdo	Totalmente en desacuerdo
1. Se me hace difícil hablar de las relaciones sexuales o de pareja con un/a cliente transexual.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Me sentiría tranquilo/a si mis hijos/as tuvieran un/a maestro/a transexual.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Reconozco que la transexualidad reta mis valores personales.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Me sentiría tranquilo/a si alguno de mis hijos/as fuera transexual.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Entiendo que darles derechos y reconocimiento a las personas transexuales trastoca los valores de nuestra sociedad.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Me sentiría tranquilo/a si alguno de mis familiares fuera transexual.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Compartiría mi oficina con un/a terapeuta transexual.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. Apoyo el que las personas transexuales puedan cambiar todos sus documentos con su identidad de género nueva.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. Para mí una mujer transexual (hombre a mujer) es una mujer como cualquier otra.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. Para mí un hombre transexual (mujer a hombre) es un hombre como cualquier otro.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Método de corrección sumaria

Puntuación regular: los ítems se puntúan de 1 (*Totalmente de acuerdo*) a 4 (*Totalmente en desacuerdo*).

Puntuación inversa: los ítems 1, 3 y 5 se puntúan de forma inversa, de 4 (*Totalmente de acuerdo*) a 1 (*Totalmente en desacuerdo*).