

Datos Normativos y Propiedades Psicométricas del SCL-90-R en Estudiantes Universitarios Chilenos

Normative Data and Psychometric Properties of The SCL-90-R in Chilean College Students

René Gempp Fuentealba*

Universidad Alberto Hurtado, Chile

Cecilia Avendaño Bravo

Universidad de Concepción, Chile

(Rec. 11 diciembre de 2007 – Acep: 29 abril 2008)

Resumen

Este trabajo sintetiza la adaptación del Inventario de Síntomas de Derogatis, Revisado (SCL-90-R) en estudiantes universitarios Chilenos. Se presentan las propiedades psicométricas y datos normativos obtenidos en una muestra de 718 estudiantes (47,6% hombres, 52,4% mujeres) de Antofagasta (32,9%), Santiago (32,3%) y Temuco (34,8%). Específicamente, se reportan resultados sobre: las distribuciones de respuesta, la estructura factorial, la capacidad discriminativa de los ítems, la fiabilidad y escalabilidad para cada dimensión primaria de síntomas, la distribución de frecuencias de las dimensiones e índices globales, las correlaciones entre escalas, el efecto del género y la zona geográfica sobre las puntuaciones, y los baremos en Puntuaciones T para los tres índices globales y para las nueve dimensiones primarias de síntomas. Se discuten los resultados y se sugieren guías para el uso de las normas.

Palabras clave: SCL-90-R, inventario, síntomas, malestar psicológico, normas.

Abstract

This paper synthesizes the adaptation of Derogatis' Symptom Checklist, Revised (SCL-90-R) in Chilean nonpatient college students. Psychometric properties and normative data were obtained from a sample of 718 students (47,6% males, 52,4 females) from Antofagasta (32,9%), Santiago (32,3%) and Temuco (34,8%). Specifically, this article reports findings about: response distributions, factor structure, discriminative power of the items, reliability and scalability of the primary dimensions, correlations among scales, gender and city effects on the scores, and T-score norms for the three global indexes and the nine primary dimensions. The results are discussed and practical guidelines are given for the use of the norms.

Keywords: SCL-90-R, checklist, symptoms, distress, norms.

* Correspondencia a: René Gempp, Escuela de Psicología Universidad Alberto Hurtado, Barroso 26, Santiago, Chile. Email: rgempp@uahurtado.cl, rene.gempp@mineduc.cl. Web: www.sigmas.cl/rgempp/

Material complementario a este artículo puede encontrarse en la dirección web <http://www.sigmas.cl/papers/SCL90/>

Los autores desean agradecer a la Doctora Eugenia Vinet (Universidad de la Frontera), quien generosamente suministró material relativo a la versión norteamericana original del SCL-90-R y a la Doctora María Martina Casullo (Universidad de Buenos Aires) que tuvo la amabilidad de proporcionarnos acceso a la adaptación argentina del instrumento elaborada por su equipo de investigación. Además, desean expresar su gratitud al psicólogo Arturo Ayres, que apoyó la recolección de datos en la ciudad de Antofagasta.

Introducción

El *Inventario de Síntomas de Derogatis, Revisado*¹ (Derogatis Symptom Checklist, Revised [SCL-90-R]) es un instrumento de autoinforme desarrollado por Leonard Derogatis y asociados (1975, 1994) para evaluar el grado de malestar psicológico actual [*distress*²] que experimenta una persona. Consiste en un listado de 90 síntomas psiquiátricos de variado nivel de gravedad, frente a los cuales el respondiente debe indicar en qué medida lo han molestado o perturbado cada uno de esos problemas durante el período comprendido entre el día de la evaluación y la semana inmediatamente anterior. Para responder, se ofrece una escala de cinco opciones acotada por los términos “Nada”, “Muy poco”, “Poco”, “Bastante” y “Mucho”, que en la corrección es valorada en un rango de 0 a 4 puntos. Una vez puntuadas las respuestas es posible caracterizar la sintomatología del evaluado en un perfil compuesto por nueve dimensiones primarias de síntomas (*Somatización, Obsesiones, Sensitividad Interpersonal, Depresión, Ansiedad, Hostilidad, Ansiedad Fóbica, Ideación Paranoide y Psicoticismo*) y tres índices globales de psicopatología: el Índice de Gravedad Global (*Global Severity Index, GSI*), el Índice de Malestar Positivo (*Positive Symptom Distress, PSDI*) y el Total de Síntomas Positivos (*Positive Symptom Total, PST*).

El objetivo de este artículo es presentar datos normativos y propiedades psicométricas de una versión del SCL-90-R adaptada para población universitaria chilena. Para ello se revisa primero el origen del instrumento. En un segundo apartado se sintetizan los antecedentes de fiabilidad y validez obtenidos en estudios internacionales. Luego, se describen los antecedentes de la versión aquí presentada, la justificación del estudio y los objetivos específicos que guiaron el trabajo. Posteriormente, se explica la metodología

utilizada y se presentan los resultados obtenidos. Se concluye con una discusión de los resultados, en que se abordan los puntos pendientes en el proceso de validación y se entregan algunas recomendaciones prácticas a los potenciales usuarios.

Origen y desarrollo del SCL-90-R

El SCL-90-R es la culminación de una prolongada tradición de instrumentos de autoinforme clínico que comenzó con el *Personal Data Sheet* (Woodworth, 1918), y que más adelante condujo al desarrollo del *Cornell Medical Index* [CMI] (Wider, 1948). Algunos años más tarde, Parloff, Kelman y Frank (1954) utilizaron varios de los síntomas incluidos en el CMI como base para desarrollar su *Discomfort Scale* que sirvió, junto con los instrumentos anteriores, como esqueleto del *Hopkins Symptom Checklist* [HSCL] (Derogatis, Lipman, Rickels, Uhlenhuth & Covi, 1974), compuesto por 51 ítems agrupados en cinco dimensiones de síntomas. Aunque el inventario mostró un funcionamiento aceptable (McNair, 1974; Waskow & Parloff, 1975), sus creadores detectaron una serie de problemas que restringían su utilidad, lo que motivó un exhaustivo programa de investigación que produjo sucesivas variaciones del HSCL durante la década de 1970, al cabo de las cuales se obtuvo el SCL-90 y, luego de algunas modificaciones menores, la versión “revisada” conocida como SCL-90-R.

Una consecuencia poco feliz de la estrategia seguida para desarrollar el SCL-90-R es que en el presente coexisten numerosas variantes del instrumento, lo que acarrea cierta confusión entre los usuarios. Por ejemplo, para el SCL existen versiones de 35, 58 y 65 ítems todavía en uso (Bech et al., 1993), además de la versión abreviada (53 ítems) desarrollada por el propio autor (Derogatis, 1993; Derogatis & Melisaratos, 1983), denominada *Brief Symptom Inventory* [BSI]. Por otro lado, como el HSCL guarda estrecha consistencia con las cinco primeras dimensiones del SCL-90-R y, por extensión, del BSI, algunos usuarios lo consideran una aproximación breve al inventario completo. Además, existe otro par de instrumentos asociados al SCL-90-R y al BSI: el *Derogatis Psychiatric Rating Scale* [DPRS] y el *SCL-90 Analogue Scale* (Derogatis, 1994). El primero es una versión del SCL-90-R diseñada para ser respondida por un observador experto (i.e., psiquiatra o psicólogo clínico) y, el segundo, un homólogo del anterior que puede ser completado por profesionales de la salud que no tienen entrenamiento formal en psicopatología (i.e., enfermeras, médicos generales).

Esta diversidad de versiones resulta problemática, además, a causa de un relativo desorden en la cronología de las publicaciones. Por ejemplo, la primera versión documentada del SCL-90 (Derogatis, Lipman & Covi, 1973) es previa a la referencia del HSCL (Derogatis et al., 1974). Con posterioridad, esta versión fue modificada, en

¹ En castellano no existe una traducción inequívoca para la expresión “*Symptom Checklist*”. Aunque su significado literal es “*Lista de Revisión de Síntomas*”, dependiendo del país de origen, distintos autores emplean denominaciones tan variadas como “*Listado de Síntomas*”, “*Listado de Revisión de Síntomas*”, “*Inventario de Síntomas*”, “*Cuestionario de Síntomas*” o “*Escala de Síntomas*”. En este artículo hemos optado por la traducción “*Inventario de Síntomas*” por tres razones. En primer lugar, es consistente con la nomenclatura habitual en nuestro idioma para los instrumentos de autoinforme compuestos por escalas múltiples. Además, la denominación “*inventario*” es fiel a la noción original en inglés: una colección sistemática de síntomas dispuestos ordenadamente para que el respondiente reconozca aquellos que le aquejan. Por último, y luego de una revisión bibliográfica exhaustiva, concluimos que “*Inventario de Síntomas*” parece ser la traducción empleada con mayor frecuencia en el mundo hispano parlante para referirse al SCL-90-R.

² Como en nuestro idioma no existe un equivalente para el término anglosajón “*distress*”, algunos autores utilizan la castellanización “*distrés*”. En nuestra opinión, esto añade una cuota ambigüedad al significado del concepto porque en la lengua española el prefijo “*dis*” suele acompañar la negación de una cualidad, lo que no es el caso. Tomando en cuenta la acepción original en inglés en este trabajo preferimos adoptar la traducción “*malestar psicológico*”, de uso frecuente entre los psicólogos de la salud hispanoparlantes.

1977, para obtener el actual SCL-90-R, aunque la primera referencia de este inventario suele fecharse dos años antes (Derogatis, 1975). Para agregar más confusión, buena parte de la investigación original que respalda al instrumento se desarrolló realmente con el SCL-90 y no con la versión revisada (e.g. Derogatis, Rickels & Rock, 1976). No obstante, en algunas referencias actuales se citan trabajos que utilizaron el SCL-90 como si hubieran empleado el SCL-90-R, justificando esta práctica en la gran similitud que existe entre las dos versiones, que en realidad difieren en sólo nueve ítems y en el formato de respuesta. Intentando zanjar la polémica, Derogatis (1994) ha indicado que el SCL-90 debe considerarse sólo un prototipo incompleto de la versión revisada y que, por tanto, se debe preferir el SCL-90-R para cualquier uso práctico. En la misma línea de argumentación, se advierte que toda la evidencia de validez acumulada desde la primera publicación del SCL-90-R se basa en esta versión y no en la previa. De hecho, a principios de la década pasada los editores del instrumento (National Computer Systems, 1993) publicaron una revisión bibliográfica que compila aproximadamente 750 investigaciones publicadas que utilizaron el SCL-90-R.

Uso, características y propiedades psicométricas del SCL-90-R

La gran popularidad del SCL-90-R entre profesionales e investigadores de la salud mental puede atribuirse a tres características: (1) se trata de un inventario de fácil aplicación y corrección, (2) permite tamizar sintomatología en varios niveles de complejidad, y (3) exhibe buenas propiedades psicométricas.

Con relación al primer punto, el SCL-90-R tiene ventajas evidentes: es un inventario autoadministrado, las instrucciones e ítems utilizan un vocabulario simple y requieren un mínimo de comprensión lectora (sexto grado) para ser respondido correctamente. Además, la consigna es simple y la tarea no reviste complejidad para el respondiente, lo que incrementa la validez de la evaluación y reduce el tiempo de administración (normalmente un máximo de 15 minutos). Puede ser aplicado en un amplio rango de edad (13 a 65 años) y sólo requiere que el evaluado tenga un nivel intelectual normal y no presente ideas delirantes o curse un trastorno psicótico en el momento de la evaluación. Su aplicación requiere de un mínimo de entrenamiento, facilitando su administración colectiva o individual por personal sin capacitación formal en salud mental (e.g., enfermeras). Por último, se trata de un instrumento relativamente breve en comparación con otros inventarios multidimensionales y la relación entre el número de ítems y el número de escalas es altamente eficiente (90 ítems que caracterizan al evaluado en nueve dimensiones primarias y tres índices globales).

Respecto a la segunda característica, el SCL-90-R fue diseñado (Derogatis, 1994) para permitir al usuario analizar la psicopatología del evaluado en tres niveles jerárquicos de complejidad: un nivel *global*, un nivel *dimensional* y un nivel *sintomático*.

En el nivel *global*, el SCL-90-R caracteriza al evaluado a partir de tres índices generales. El *Índice de Gravedad Global* [GSI] es un indicador directo del nivel de gravedad de los síntomas, mientras el *Índice de Malestar Positivo* [PSDI] es una medida de la intensidad de los síntomas percibidos por el respondiente. En la práctica, el PSDI se

Tabla 1
Descripción de las dimensiones primarias de síntomas del SCL-90-R

Dimensión	Descripción
Somatización	Malestares que surgen de la percepción de disfunciones corporales cardiovasculares, gastrointestinales, respiratorias y de otros sistemas con fuerte mediación del sistema nervioso autónomo. También incluye dolores de cabeza, musculares y otras manifestaciones somáticas de ansiedad.
Obsesiones	Incluye los síntomas característicos del síndrome obsesivo compulsivo: diversos pensamientos, acciones e impulsos que son vivenciados como involuntarios, indeseados, irresistibles, egodistónicos o imposibles de controlar.
Sensitividad interpersonal	Sentimientos de inferioridad e inadecuación personal, especialmente en comparación con otras personas; auto depreciación e incomodidad durante las interacciones personales. Las personas con altas puntuaciones en esta dimensión también reportan expectativas negativas sobre las interacciones interpersonales y elevados niveles de autoconciencia pública.
Depresión	Los ítems son una muestra representativa de las principales manifestaciones clínicas del trastorno depresivo: estado de ánimo disfórico, falta de motivación, pérdida de interés en las actividades habituales, poca energía vital y sentimiento de desesperanza. También se incluyen ideaciones suicidas y otros correlatos somáticos y cognitivos de la depresión.
Ansiedad	Síntomas y comportamientos asociados clínicamente con elevados grados de ansiedad: inquietud, nerviosismo, tensión y ataques de pánico.

Hostilidad	Incluye pensamientos, sentimientos y acciones característicos del estado afectivo negativo caracterizado como rabia o ira. La selección de ítems incluye la expresión de agresión, irritabilidad, rabia y resentimiento. Por ejemplo, algunos ítems refieren sentimientos de fastidio, necesidad de romper cosas o temperamento explosivo e incontrolable.
<i>Ansiedad fóbica</i>	Caracteriza una respuesta persistente de miedo (a personas, lugares, objetos o situaciones específicas) que es irracional y desproporcionada al estímulo que la provoca, ocasionando que la persona desarrolle una conducta de evitación o escape del estímulo atemorizante. Los síntomas de esta dimensión se centran en las manifestaciones más disruptivas de la conducta fóbica, varios de los cuales son típicos de la agorafobia (miedo a los espacios abiertos, multitudes, lugares públicos, por ejemplo).
<i>Ideación paranoide</i>	Representa fundamentalmente las distorsiones cognitivas típicas del comportamiento paranoide: pensamiento proyectivo, suspicacia, grandiosidad, centralidad, temor a la pérdida de autonomía y delirios.
<i>Psicoticismo</i>	Esta dimensión fue diseñada para representar el constructo en tanto un continuo de la experiencia humana. Por ello incluye un amplio espectro de síntomas referidos a estados de soledad, estilo de vida esquizoide, alucinaciones auditivas, transmisión y control del pensamiento.

emplea como un índice de estilo de respuesta (i.e., si la persona tiende a exagerar o a minimizar los problemas que le aquejan). El tercer índice global, denominado *Total de Síntomas Positivos* [PST], es simplemente el número total de síntomas que presenta la persona, independientemente de su gravedad. Normalmente es interpretado como una medida de la amplitud de la sintomatología presentada por el evaluado.

En el nivel de análisis *dimensional*, el SCL-90-R entrega un perfil compuesto por las nueve dimensiones primarias de síntomas antes mencionadas. Éstas fueron derivadas progresivamente utilizando una combinación de estrategias clínicas, racionales y empíricas con el objetivo de ser clínicamente útiles, para lo cual se impuso el triple requisito de una definición clara y consistente en la literatura, susceptibilidad de medición mediante autoinforme y confirmación psicométrica (Derogatis & Cleary, 1977a, 1977b). Una descripción breve de las nueve dimensiones se presenta en la Tabla 1.

Por último, en el nivel *sintomático*, el SCL-90-R permite al evaluador analizar las respuestas a uno o varios síntomas que le parezcan relevantes de acuerdo con los fines de la aplicación y/o según algún modelo o teoría particular. En este nivel de análisis se incluyen también siete síntomas adicionales, que no aportan puntuación a ninguna de las nueve dimensiones primarias, pero, que se consideran marcadores clínicamente relevantes para su interpretación individual y que son incluidos en el cómputo de los índices globales. Una descripción abreviada de estos ítems puede encontrarse en la primera columna de la Tabla 3, que se presenta más adelante.

La tercera característica que explica la popularidad del SCL-90-R es que sus propiedades psicométricas han resultado apropiadas y estables a través del tiempo, países, grupos de edad y tipos de población evaluada (pacientes versus no pacientes). Sin abundar en detalles, la *fiabilidad* de las nueve dimensiones alcanza valores cercanos o supe-

riores a $r = 0.70$ en la mayoría de los estudios de fiabilidad test-retest y en torno a 0.80 en los análisis de consistencia interna (Derogatis, 1994; Derogatis & Savitz, 2000). Otros estudios han demostrado, además, que la escalabilidad de las dimensiones es substancialmente satisfactoria (Olsen, Mortensen & Bech, 2004; Schmitz et al., 2000) y que muestran un comportamiento psicométrico idóneo cuando se analizan con modelos de Teoría de Respuesta al Ítem (Olsen et al., 2004).

Junto con lo anterior, una multiplicidad de estudios realizados desde 1980 en adelante ha acreditado la *validez* del SCL-90-R, de sus índices globales y de las dimensiones específicas. Por ejemplo, se han documentado los perfiles característicos de varios cuadros psiquiátricos, incluyendo desórdenes ansiosos (e.g. Cameron, Thyer, Nesse & Curtis, 1986), depresión (e.g., Prusoff, Weissman, Klerman & Rounsaville, 1980; Wetzler, Kahn, Cahn, Van Praag & Asnis, 1990), crisis de pánico (e.g., Buller, Maier & Benkert, 1986), disfunciones sexuales (e.g., Derogatis, Meyer & King, 1981), y abuso de sustancias (e.g., Steer, Platt, Ranieri & Metzger, 1989). Además, se ha demostrado (ver Derogatis, 1994 para una revisión) la validez del SCL-90-R como medida del malestar psicológico asociado con trastornos médicos generales y específicos (e.g., trastornos cardiovasculares, renales, dolor crónico). A esto se agrega evidencia que demuestra su efectividad como *screening* genérico de trastornos psicopatológicos (e.g., Derogatis & DellaPietra, 1994; Schmitz, Kruse, Heckrath, Alberti & Tress, 1999; Schmitz, Kruse & Tres, 2001; Waryszak, 1982) y de trastornos específicos de la esfera somatomorfa (e.g., Dickson, et al., 1992; Kellner, Hernández & Pathak, 1992; Kirmayer, Robbins, Dworkind & Yaffe, 1993; Simon & VonKorff, 1991), afectivo-depresiva (e.g., Koeter, 1992; Morgan, Wiederman & Magnus, 1998; Rosenberg, Bech, Mellergard & Ottosson, 1991) y ansiosa (e.g., Morgan et al., 1998; Rief & Fichter, 1992; Vollrath, Koch & Angst, 1990). También existen antecedentes que prueban su utilidad difer-

encial en la detección y evaluación de estrés (e.g., Derogatis, 1994; Derogatis & Coons, 1993), estrés postraumático (e.g., Davidson, Kudler, Saunders & Smith, 1990; Wolfe, Brown & Bucsela, 1992), riesgo suicida (e.g., Bulik, Carpenter, Kupfer & Frank, 1990; Swedo et al., 1991), abuso de alcohol y de sustancias (e.g., Derogatis & Savitz, 2000; Johnson, Brems & Fisher, 1996; Lucht, Jahn, Barnow & Freyberger, 2002), historial de abuso sexual (e.g.; Swett, Surrey & Cohen, 1990; Toomey, Seville, Mann, Abashian & Grant, 1995) y antecedentes de disfunción sexual (e.g., Coffey, Leitenberg, Henning, Turner & Bennett, 1996; Heiman & Rowlands, 1983; Turner, Althof, Levine & Bodner, 1991), entre otros.

Por último, se ha consolidado su empleo como criterio para evaluar cambios sintomáticos inducidos por farmacoterapia (e.g., Florkowski, Stevens, Joyce, Espiner & Donald, 1998; Holland et al., 1991; Pani, Maremmani, Pirastu, Tagliamonte & Gessa, 2000; Primeau, Fontaine & Beauclair, 1990; Strayer et al., 1994) y psicoterapia (e.g., De Jonghe, Kool, van Aalst, Deckker & Peen, 2001; Schauenberg, 1999; Schmitz, Hartkamp & Franke, 2000; van der Sande et al., 1997).

A toda esta evidencia de *validez convergente* se agregan numerosos estudios que demuestran la *validez concurrente* y *predictiva* del inventario y de sus escalas, utilizando como criterios otros instrumentos de evaluación clínica, escalas de *screening*, diagnósticos psiquiátricos, protocolos de evaluación estructurada o indicadores de recidiva, principalmente (Derogatis & Savitz, 2000).

El único componente de la validez del SCL-90-R sobre el que no existe completo acuerdo es la *validez factorial* del instrumento. Aunque los autores del inventario proponen que la estructura de nueve dimensiones es básicamente estable (Derogatis, 1994; Derogatis & Cleary, 1977a, 1977b; Derogatis & Savitz, 2000), los estudios independientes obtienen resultados inconsistentes entre sí. Por ejemplo, análisis factoriales exploratorios de los ítems, realizados en distintas poblaciones encuentran soluciones que van desde un factor general (Bonyngé, 1993), hasta cuatro (Hafkenschheid, 1993; Schwarzwald, Weisenberg & Solomon, 1991), cinco (Hoffman & Overall, 1978; Steer, Clark & Ranieri, 1994), seis (Holi, Sammallahti & Aalberg, 1998), nueve (Holcomb, Adams & Ponder, 1983) o diez factores (Shutty, DeGood & Schwartz, 1986). En contextos más cercanos, un estudio español (Caparrós-Caparrós, Villar-Hoz, Juan-Ferrer & Viñas-Poch, 2007) no logró confirmar la estructura de nueve factores en estudiantes universitarios, mientras en Argentina, Casullo y Fernández-Liporace (2001) consiguieron una solución factorial que ellos consideran concordante con la original. Otros estudios en que se aplicó análisis factorial confirmatorio, también tienden a rechazar el ajuste del modelo de nueve factores (e.g., Carpenter & Hittner, 1995), aunque este resultado es habitual cuando se aplica esta técnica a instrumentos derivados mediante análisis factorial ex-

ploratorio, lo que ha llevado a cuestionar la pertinencia del método (McCrae, Zonderman, Costa, Bond & Paunonen, 1996). Por otro lado, cuando se factorizan las escalas en lugar de los ítems suelen encontrarse soluciones de una o dos dimensiones, aunque a veces es necesario extraer aun más factores para explicar la covarianza entre las escalas primarias de síntomas (Strauman & Wetzler, 1994).

La evidencia factorial acumulada ha generado interpretaciones disímiles. Para algunos autores (e.g., Cyr, McKenna-Foley & Peacock, 1985) estos resultados cuestionan la validez del SCL-90-R y demuestran que se trata de una medida general de malestar psicológico y no de un inventario multidimensional, propiamente tal. Otros autores, en cambio (e.g., Vassend & Skrandal, 1999), sugieren que la inestabilidad factorial del SCL-90-R no es un problema del instrumento sino del tipo de ítems, esto es, los síntomas presentan distribuciones y correlaciones que dependen estrechamente de su prevalencia específica en cada población evaluada. De esta manera, pequeños cambios en la prevalencia de algunos síntomas incluidos en el cuestionario afectan la estructura factorial global. Como éste es un problema inherente a los inventarios de síntomas, más que cuestionar la validez factorial del instrumento habría que discutir la pertinencia de considerar la estructura factorial como criterio de validez para una escala sintomática. Tal tipo de razonamiento nace de la observación realizada por Bollen y de Lennox (1991) a propósito de que el análisis factorial no es un modelo estadístico apropiado para los inventarios de síntomas manifiestos. En esta línea de argumentación, la relativa inestabilidad factorial del SCL-90-R no sería un problema serio si se toma en cuenta el cúmulo de evidencia convergente, concurrente y predictiva que justifica su validez en múltiples ámbitos de aplicación.

En síntesis, la suma de estudios realizados hasta la fecha, utilizando variedad de diseños (experimentales y correlaciones) y técnicas analíticas (e.g., comparación de promedios, modelos de regresión lineal, modelos de regresión logística, análisis ROC, análisis factoriales, modelos de respuesta al ítem), han permitido acumular pruebas empíricas de la *validez para*³ una diversa gama de

³ A propósito de este punto es conveniente recordar que el concepto contemporáneo de validez, cuya definición más autorizada viene dada por la última edición de los *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association & National Council on Measurement in Education, 1999), no la entiende como una propiedad fija e inmutable de los instrumentos, sino como el grado en el cual existe evidencia y teoría que apoye las *interpretaciones* que usuarios específicos hagan de las puntuaciones de un test, en contextos concretos. Esta forma de definir el concepto tiene dos consecuencias cruciales para la discusión sobre la validez del SCL-90-R. Primero, que la validación es un proceso de *acumulación progresiva de evidencia* que acredita empíricamente los usos que se hagan de un instrumento. Por ello la validez es una cuestión de grado y un test nunca puede considerarse definitivamente "validado". Segundo, que la validez concierne a las *interpretaciones* que

usos específicos del SCL-90-R. Una revisión parcial de esos estudios hasta la década recién pasada, puede consultarse en Derogatis (1994).

Uso y adaptación del SCL-90-R en Chile

La amplia evidencia a favor de la utilidad, eficiencia y validez del SCL-90-R ha estimulado su adaptación transcultural, contabilizándose versiones en más de 20 idiomas distintos del inglés, incluyendo el español, portugués, alemán, italiano, holandés, francés, suizo, ruso, sueco, japonés, coreano, chino, vietnamita, hebreo, árabe, danés, y noruego (Derogatis, 1994; Derogatis & Savitz, 2000).

En el mundo hispano parlante se han realizado adaptaciones y estudios psicométricos locales del SCL-90-R en varios países, entre ellos España (e.g., Caparrós-Caparrós, et al., 2007; De las Cuevas, et al., 1991; González de Rivera, De las Cuevas, Rodríguez & Rodríguez, 2002, Martínez-Azumendi, Fernández-Gómez & Betia-Fernández, 2001; Robles, Andreu & Peña, 2002), México (Cruz, López, Blas, González & Chávez, 2005; Lara, Espinosa de Santillana, Cárdenas, Fócil & Cavazos, 2005) y Argentina, (e.g., Bonicatto, Dew, Soria & Seghezso, 1997; Casullo, 2004; Casullo & Castro Solano, 1999; Casullo, Cruz, González & Maganto, 2003; Casullo & Fernández-Liporace, 2001).

En el caso de Chile, una revisión de la literatura muestra que el SCL-90-R y/o sus instrumentos asociados (HSLC, BSI), han sido empleados en investigaciones puntuales (e.g., Flores, García, Matus & Sandoval, 2004; Heerlein, et al., 2000), pero no parece existir una versión estandarizada nacional consignada en alguna publicación científica⁴. Esto motivó una primera tentativa de adaptación por parte de nuestro equipo (Gempp, 2003), que fue utilizada en tres investigaciones (Ferrada & Mardones, 2003; García & Seguel, 2003; Silva & Gempp, 2004) en que se observaron buenos niveles de discriminación de los ítems y fiabilidad de las escalas. Con posterioridad, esa versión fue revisada y actualmente se trabaja con una versión depurada. Análisis psicométricos más complejos, algunos de los cuales se pre-

sentan en este artículo, indican un funcionamiento idóneo del inventario. En el presente, varios equipos de investigación universitarios están trabajando con esta versión.

Como parte del proceso de adaptación en curso del SCL-90-R es necesario estudiar sus propiedades psicométricas y producir normas interpretativas en distintas poblaciones, pues los grupos mutuamente exclusivos son preferibles a las muestras heterogéneas al momento de definir criterios para evaluar significancia clínica (Jacobson, Follette & Revenstorff, 1984).

En el caso del SCL-90-R, tradicionalmente se ha considerado necesario (e.g., Derogatis, 1975, 1994) diferenciar las normas por género (hombres versus mujeres), grupo de edad (adolescentes versus adultos), condición diagnóstica (no pacientes versus pacientes) y, en este último caso, según gravedad (pacientes psiquiátricos internados versus pacientes psicológicos ambulatorios). Combinando total o parcialmente estos criterios diversos estudios han producido normas y/o comparado el desempeño del instrumento en subpoblaciones específicas (e.g., Bonyngge, 1993; Carpenter & Hittner, 1995; Hart, Bryer & Martines, 1991; Johnson, El-lison & Heikkinen, 1989; Porter, Wilson & Frisch, 1994).

Por su parte, en el Manual original del instrumento se distinguen cuatro poblaciones normativas, cada una subdividida según género: pacientes adultos en consulta psiquiátrica, adultos no pacientes, pacientes psiquiátricos adultos internados y adolescentes no pacientes. No obstante, varios estudios han advertido la necesidad de examinar con mayor profundidad el comportamiento del instrumento en *muestras universitarias de no pacientes* y de elaborar normas específicas para este tipo de población, debido a que los universitarios típicamente reportan más sintomatología y obtienen puntuaciones más altas en las escalas del SCL-90-R que los adultos no pacientes (Todd, Deane & McKenna, 1997).

Considerando estos antecedentes, y en el marco de los análisis que nuestro equipo desarrolla con el SCL-90-R, el propósito de este trabajo es reportar las propiedades psicométricas del SCL-90-R obtenidas en estudiantes universitarios no pacientes y presentar las normas provisionales derivadas para esta población. Específicamente, se exponen resultados relativos a: (a) la distribución de respuestas a los ítems; (b) la estructura factorial del instrumento; (c) la capacidad discriminativa de los ítems que componen cada dimensión primaria de síntomas; (d) los coeficientes de fiabilidad y escalabilidad para cada dimensión primaria de síntomas; (e) los estadísticos descriptivos y distribución de los ítems adicionales, las nueve dimensiones primarias de síntomas y los tres índices globales; (f) las intercorrelaciones entre las dimensiones primarias de síntomas, (g) el efecto del género y la zona geográfica en el comportamiento de las escalas; (h) los baremos en Puntuaciones T para los tres índices globales y para las nueve dimensiones primarias de síntomas; e (i) algunos indicadores psicométricos

se quieran hacer de los resultados del instrumento y no al instrumento en sí mismo. Ello obliga a reunir evidencia que valide *cada uso específico* que se quiera hacer del test. Por ejemplo, si existe evidencia que demuestre la validez del SCL-90-R para el screening de depresión en adolescentes, ello no garantiza que también sea una medida válida para el screening de depresión en adultos mayores; al menos hasta que se lo demuestre empíricamente. Por estas razones lo correcto es referir cada nueva evidencia de validez como "*validez para* [un uso específico]" y no hablar de la "*validez del test*".

⁴ La consulta informal a varias fuentes permite suponer que actualmente hay algunos grupos de investigación utilizando el SCL-90-R o sus variantes, pero una búsqueda formal en bases de datos no permitió localizar ningún estudio de adaptación, validación, estandarización o análisis psicométrico basado en muestras chilenas. Lo anterior, obviamente, no implica necesariamente que estos estudios no se estén llevando a cabo.

complementarios a estas normas, como el Error Estándar de Medida, la Diferencia Mínima Significativa y los coeficientes de fiabilidad para los puntos de corte. Por razones que se discuten en la sección final del artículo, las normas construidas se consideran provisionales y no definitivas.

Método

Participantes y procedimiento

Siguiendo a Todd et al. (1997), se definió como población objetivo a estudiantes universitarios de ambos géneros que no estuvieran en tratamiento psiquiátrico. Desde esta población se obtuvo una muestra no probabilística organizada en un diseño de cuotas independientes 2x3 con dos controles. Primero, considerando que el instrumento original diferencia sus normas por género (Derogatis, 1994), se intentó balancear proporcionalmente el número de hombres y mujeres de la muestra. Segundo, para aumentar la representatividad de los resultados se decidió dividir la muestra en partes iguales entre participantes provenientes de las zonas norte, centro y sur del país. Para ello se definieron tres centros urbanos representativos de cada zona (Antofagasta, Santiago y Temuco) desde los cuales se extrajeron finalmente los datos.

Aplicando estos criterios, la muestra definitiva quedó conformada por 718 estudiantes universitarios (47,6% de hombres y 52,4% de mujeres), mayoritariamente de primer y segundo año de carrera y con una edad media de $M=21,14$ años ($SD=2,09$). El rango de edades se encuentra entre los 18 y los 27 años, y el cuartil inferior, mediana y cuartil superior de la distribución fueron $C_i=19$, $Mdn=21$ y $C_s=22$ años, respectivamente. Respecto a la ciudad de procedencia, el 32,9% ($n=236$) respondió el inventario en Antofagasta, el 32,3% ($n=232$) en Santiago y el 34,8% ($n=250$) en Temuco, obteniéndose una proporción equivalente de hombres y mujeres en cada ciudad. Los participantes estudiaban carreras de las áreas biológica y de la salud (13,79%), educación (24,51%), matemáticas e ingeniería (31,62%), y ciencias sociales y humanas (30,08%).

Para la aplicación, el SCL-90-R fue incorporado a un cuadernillo más amplio que incluía otros instrumentos e indagaba sobre variables adicionales de interés para los investigadores. En general, la administración se realizó en horario de clases, con autorización previa de los profesores respectivos. Los estudiantes fueron invitados a participar voluntariamente y se les garantizó el anonimato de sus respuestas, toda vez que el cuadernillo no solicitaba identificación personal alguna. Las instrucciones fueron dadas oralmente, en forma colectiva, aunque los investigadores estuvieron presentes para resolver eventuales dudas. Cada sesión de aplicación duró entre 20 y 30 minutos y se presentaron pocas dudas sobre los instrumentos. No se

entregó ningún tipo de remuneración o compensación a los participantes por su colaboración.

Instrumento

Como se ha explicado previamente, el SCL-90-R está compuesto por 90 ítems que intentan ser una muestra representativa de los síntomas psiquiátricos más característicos de los síndromes que el inventario evalúa. Por ejemplo, el ítem 1 corresponde a “Dolores de cabeza”, el ítem 16 es “Escuchar voces que otras personas no oyen”, el ítem 38 señala “Tener que hacer las cosas muy despacio para estar seguro(a) de que están bien hechas” y el ítem 66 corresponde a “Dormir mal o tener un sueño poco reparador”. Del total de ítems, 83 están incluidos en alguna de las nueve dimensiones primarias de síntomas y 7 son síntomas adicionales que no se incorporan a las escalas primarias, pero que se utilizan individualmente como marcadores clínicos relevantes y se agregan al cálculo de los índices globales. En el ejemplo previo, el ítem 1 pertenece a la dimensión de *Somatización*, el ítem 16 integra la dimensión de *Psicoticismo*, el ítem 38 es un síntoma de la dimensión de *Obsesiones* y el ítem 66 forma parte del grupo de ítems adicionales.

En la administración, las instrucciones presentan el inventario como una “lista de problemas que las personas a veces tienen” y solicitan al respondiente señalar en qué medida cada problema lo ha preocupado o molestado durante los últimos siete días, incluyendo el día de hoy, utilizando una escala de cinco opciones (“Nada”, “Muy poco”, “Poco”, “Bastante” y “Mucho”). Según Derogatis (1994), el uso del anclaje temporal de una semana se justifica en que los siete días más recientes en la vida de un individuo usualmente contienen la información más relevante sobre su estado clínico actual (Hamilton, 1970). En cualquier caso, este formato de instrucciones es el más utilizado en los inventarios sintomáticos.

Una vez administrado el inventario, los ítems son puntuados asignando entre 0 a 4 puntos a cada alternativa (“Nada” = 0; “Muy poco”=1; “Poco”=2; “Bastante”=3; “Mucho”=4). Todos los ítems son directos, por lo cual no es necesario invertir la puntuación en ninguno de ellos. De esta manera, una puntuación más alta en un ítem indica que el síntoma respectivo se ha experimentado con mayor frecuencia.

La puntuación bruta para cada dimensión primaria de síntomas se calcula promediando el número de ítems respondidos. Para ello se suman las puntuaciones de los ítems que integran la dimensión respectiva y, luego, se dividen por el número de ítems efectivamente respondidos. Por ejemplo, la dimensión de *Somatización* tiene 12 ítems. Si un evaluado respondió sólo 10 ítems, la puntuación de la escala se calcula sumando las puntuaciones obtenidas en esos 10 ítems y luego dividiendo el resultado por 10. Este procedimiento tiene la ventaja de uniformar el rango de

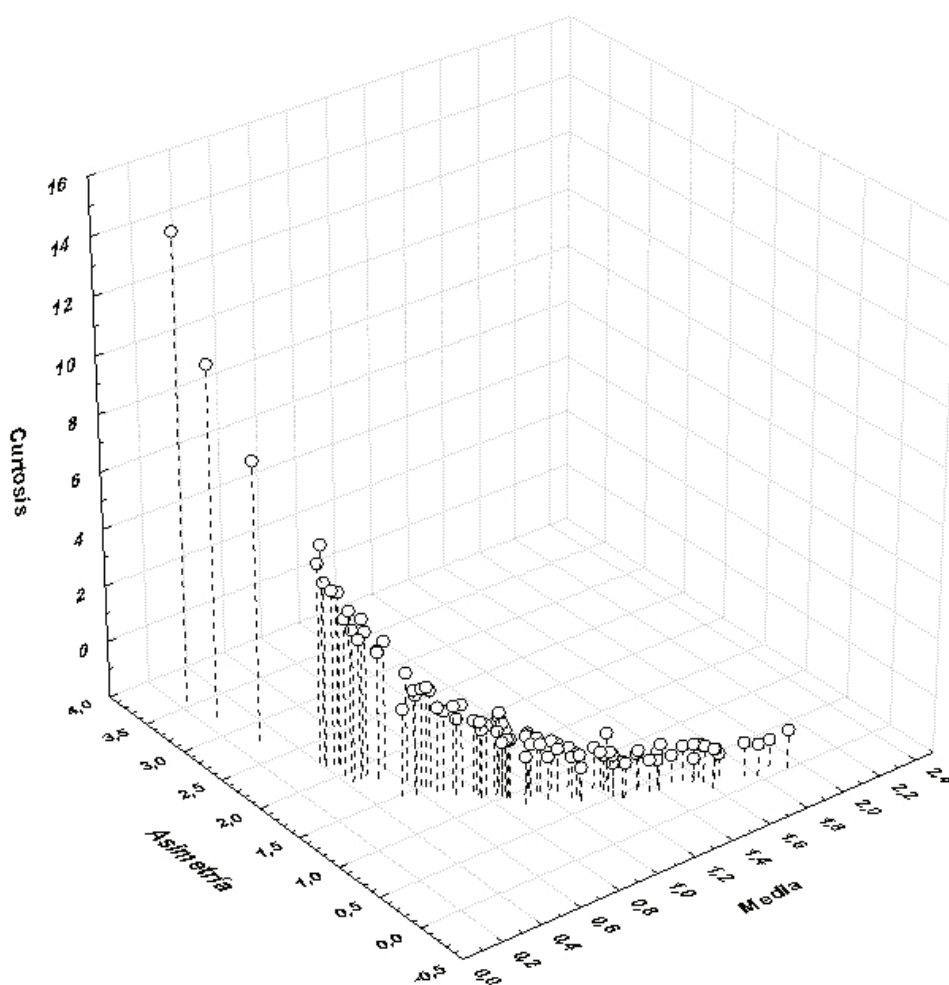
puntuación de cada escala primaria (que queda fijado entre 0 a 4 puntos), compensando el efecto de que cada dimensión tenga un número distinto de ítems. Además, controla el efecto de los ítems no respondidos.

El método para calcular las puntuaciones de los índices globales es el siguiente: el *Índice de Gravedad Global* [GSI] se calcula sumando las puntuaciones a todos los ítems, incluyendo los siete síntomas adicionales, y luego dividiendo el resultado por el total de ítems respondidos (que será 90, si no se dejaron ítems sin responder). Por su parte, el *Total de Síntomas Positivos* [PST] es derivado contando el total de ítems respondidos con una puntuación distinta de cero. Finalmente, el *Índice de Malestar Positivo* [PSDI] es calculado dividiendo la suma de todos los ítems por el PST.

La versión del SCL-90-R utilizada en esta investigación fue obtenida mediante la adaptación directa de la versión

original en inglés, utilizando como criterio comparativo las versiones española, mexicana y argentina. Siguiendo las guías de la *International Test Commission* para la adaptación transcultural de instrumentos psicométricos (e.g., Hambleton, 2001; Tanzer y Sim, 1999; van de Vijver y Hambleton, 1996), la traducción se realizó mediante el *enfoque de comité* y aplicando *descentramiento* (Brislin, 1986) para procurar que la redacción de cada ítem reflejara el significado original de los síntomas, ajustado al contexto chileno y al tipo de población estudiada, más que su traducción literal. El borrador fue revisado por psicólogos clínicos con experiencia cultural pertinente y por traductores profesionales, todos los cuales trabajaron independientemente entre sí. Estas revisiones fueron luego analizadas por los autores y dos psicólogos bilingües, contrastándolas además con las adaptaciones realizadas en España, México y Argen-

Figura 1: Medias, asimetrías y curtosis de los 90 ítems del SCL-90-R



tina. La versión obtenida fue sometida a un par de pilotajes con estudiantes universitarios y de enseñanza media, cuyos resultados se consideraron satisfactorios.

Resultados

Para comenzar, se analizaron las distribuciones de frecuencia de cada uno de los ítems del inventario. Los resultados muestran que en los 90 ítems se presentaron respuestas en todo el rango posible de alternativas (0 a 4) aunque la mayoría de éstas se concentran en las tres primeras categorías de respuesta. En consecuencia, los ítems tuvieron bajos promedios de adhesión. Por ejemplo, los síntomas con promedio más bajo y más alto (i.e., con menor y mayor prevalencia) fueron, respectivamente, el ítem 16 (*"Escuchar voces que otras personas no oyen"*) con un promedio $M=0.24$, y el ítem 2 (*"Nerviosismo"*) con un promedio $M=1.94$. Estos promedios equivalen, aproximadamente, a la primera y tercera alternativa de respuesta (*"Nada"* y *"Poco"*). El promedio global de todos los ítems es $M=0.96$, que corresponde a la segunda categoría de respuesta (*"Muy poco"*).

En general, este resultado es típico de los instrumentos de autoinforme psicopatológico cuyos ítems consisten en síntomas frente a los cuales el respondiente indica la frecuencia en que los ha experimentado. Como es obvio, a mayor gravedad del síntoma, menor frecuencia de aparición, ocasionando alta concentración de casos en las primeras categorías de respuesta. Esto puede ilustrarse con los ítems 16 y 2, recién mencionados. Estadísticamente, este patrón de respuestas se refleja en que los ítems muestran promedios bajos, curtosis elevadas y asimetrías marcadamente positivas, especialmente si el instrumento se aplica a sujetos no pacientes. Al representar estos indicadores en un dispersiograma tridimensional se observa normalmente una curva como la obtenida por los ítems del SCL-90-R, exhibida en Figura 1. Allí puede verse que el patrón de

promedios, asimetrías y curtosis de los ítems es totalmente consistente con lo esperado teóricamente.

El análisis continuó verificando la estructura factorial del instrumento. Como ya se adelantó, este ha sido un punto conflictivo en el desarrollo del SCL-90-R. Por ello se optó por seguir una estrategia *semi-confirmatoria* consistente en someter a los ítems a un análisis factorial exploratorio, pero, definiendo a priori la estructura factorial esperada y forzando la rotación hacia la configuración de cargas factoriales predicha teóricamente, para luego evaluar el ajuste de la solución obtenida. En términos genéricos, se trata de una estrategia clásica de análisis factorial conocida como rotación de Procrusto (Cliff, 1966; Schönemann, 1966). En fecha más reciente este método ha sido recomendado como una alternativa plausible para evaluar la replicabilidad de la estructura factorial de inventarios de personalidad o clínicos en investigación transcultural, dado que evita la inflexibilidad de los métodos confirmatorios (que rara vez confirman las estructuras exploratorias) y, al mismo tiempo, permite prefiar una estructura factorial y evaluar el ajuste de la solución (McCrae et al., 1996; Paunonen, 1997).

De acuerdo con la estrategia indicada, se comenzó estimando la matriz de correlaciones policóricas entre los 83 ítems que conforman las nueve dimensiones primarias. Se optó por factorizar la matriz policórica, debido a que resulta más apropiada que las correlaciones de Pearson para variables ordinales, especialmente cuando se observan asimetrías y curtosis elevadas (Muthén & Kaplan, 1985), como sucede en el presente caso y en los inventarios clínicos en general. La matriz resultante mostró índices adecuados de factorizabilidad (prueba de Bartlett significativa; $KMO=0.89$), así que se procedió a extraer nueve factores mediante el método de Mínimos Cuadrados Ordinarios y, posteriormente, se rotó la solución hacia una configuración a priori en que cada ítem saturaba únicamente en su respectiva escala. Finalmente, se evaluó la congruencia entre la matriz

Tabla 2
Propiedades psicométricas de las nueve dimensiones primarias de síntomas

Escala	Discriminación de los ítems					
	N. Ítems	Mínima	Mdn	Máxima	H	
Somatización	12	.22	.49	.62	.82	.31
Obsesiones	10	.23	.48	.65	.80	.31
Sensitividad interpersonal	9	.26	.42	.56	.76	.29
Depresión	13	.23	.45	.66	.82	.29
Ansiedad	10	.24	.46	.52	.76	.27
Hostilidad	6	.21	.44	.51	.66	.28
Ansiedad fóbica	7	.24	.35	.51	.64	.23
Ideación paranoide	6	.37	.51	.57	.74	.35
Psicoticismo	10	.20	.40	.52	.73	.25

Tabla 3
Numeración, descripción, media, desviación estándar y porcentaje de respuesta a cada opción para los 7 ítems adicionales

Ítem	Opciones de respuesta						
	<i>M</i>	<i>SD</i>	0	1	2	3	4
19. Poco apetito	0.78	1.05	55.7	21.3	14.1	7.0	1.9
44. Problemas para dormir	1.25	1.27	40.0	19.4	21.9	13.0	5.7
59. Pensamientos sobre la muerte o morirse	0.47	0.92	73.8	13.7	6.7	3.9	1.9
60. Comer en exceso	1.42	1.35	36.5	19.1	18.5	17.8	8.1
64. Despertarse muy temprano	0.97	1.20	50.0	22.0	12.7	11.3	4.0
66. Sueño intranquilo o poco reparador	1.23	1.16	35.5	24.8	25.3	10.3	4.1
89. Sentimientos de culpa	0.93	1.05	46.1	24.9	21.2	5.3	2.5

factorial obtenida y la matriz a priori mediante Índices de Similaridad Euclídeana con escalamiento doble (Barret, 2005). Estos son coeficientes de similaridad factorial cuyo valor mínimo es cero (ausencia de congruencia) y máximo es 1 (congruencia factorial perfecta). Para la solución global se obtuvo un coeficiente de 0.87 y para las nueve dimensiones evaluadas se obtuvieron valores entre 0.84 y

0.90, con una mediana de 0.87. En general, estos resultados muestran una congruencia aceptable. Así, puede concluirse que la estructura de nueve factores predichos teóricamente es capaz de explicar la varianza de los ítems en forma razonable, aunque otras soluciones serían más apropiadas para lograr un ajuste óptimo a los datos. En general, un examen cuidadoso de las matrices factoriales sugiere que

Tabla 4

Media, desviación estándar, asimetría, curtosis, puntuaciones mínima y máxima de las nueve dimensiones primarias de síntomas y de los tres índices globales.

Escala	Media	SD	Asimetría	Curtosis	Mínimo	Máximo
Somatización	1.06	0.69	0.96	0.90	0.00	3.83
Obsesiones	1.32	0.68	0.47	-0.32	0.10	3.30
Sensitividad interpersonal	0.94	0.61	0.70	-0.10	0.00	2.89
Depresión	1.09	0.63	0.65	-0.07	0.00	3.08
Ansiedad	1.01	0.59	0.96	0.97	0.00	3.60
Hostilidad	0.84	0.62	1.04	0.72	0.00	3.17
Ansiedad fóbica	0.48	0.49	1.66	4.50	0.00	4.00
Ideación paranoide	0.87	0.71	1.06	0.99	0.00	3.33
Psicoticismo	0.74	0.56	0.85	0.17	0.00	3.00
GSI	0.96	0.48	0.74	0.36	0.09	2.89
PSDI	1.82	0.40	0.42	0.34	1.00	3.53
PST	45.68	16.52	0.16	-0.46	8	85

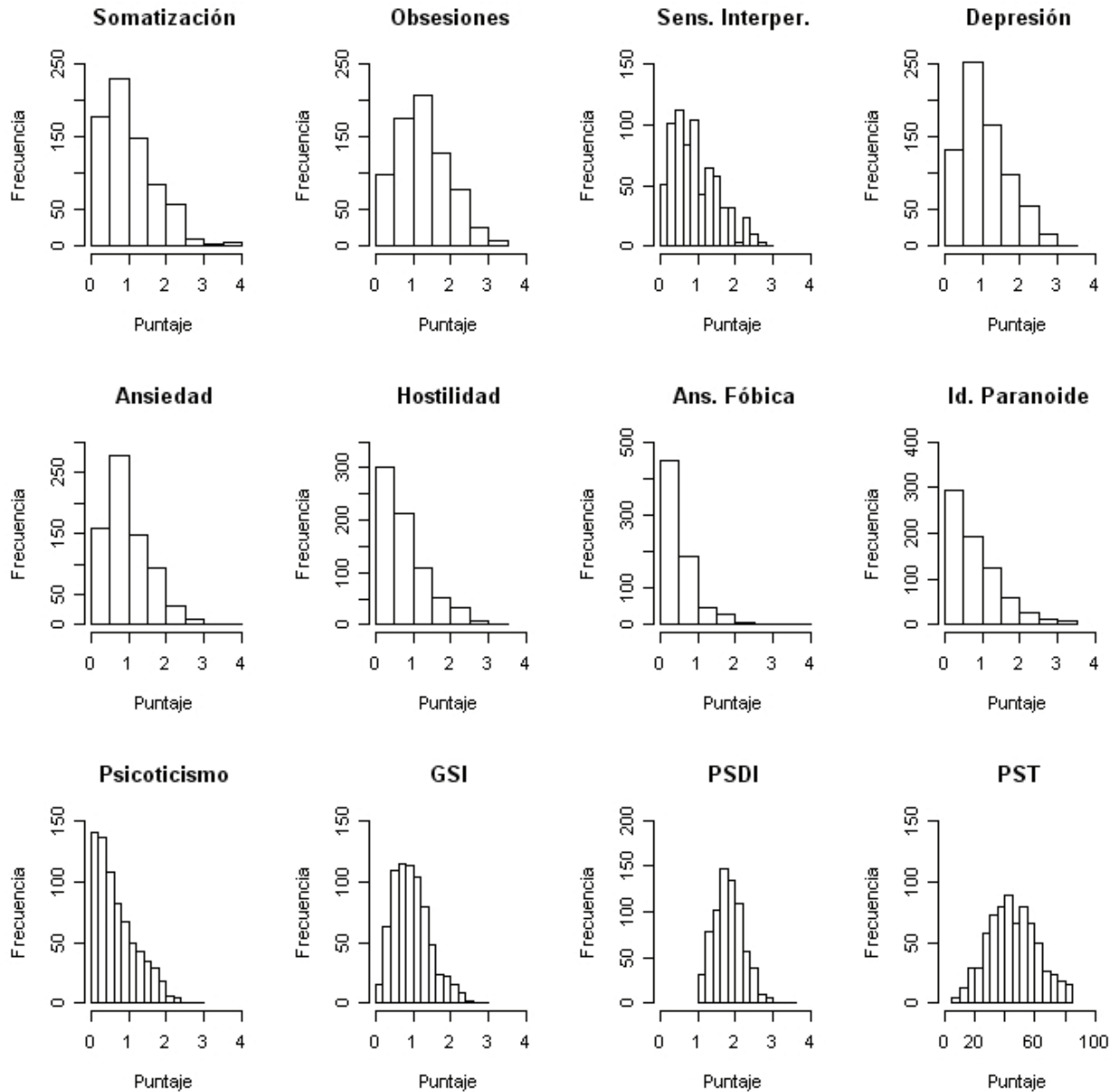
esto podría explicarse por la alta superposición entre los ítems de algunas escalas⁵.

Una vez examinada la estructura factorial, se procedió a analizar las propiedades métricas de cada una de las nueve dimensiones primarias de síntomas, de los índices globales y de los siete ítems adicionales. Para las dimensiones primarias, la Tabla 2 presenta un resumen de las capacidades

discriminativas de los ítems (correlaciones ítem-total corregidas) y de las fiabilidades de cada escala, estimadas mediante el coeficiente (Cronbach, 1951). Además, se reportan los niveles de escalabilidad de cada dimensión, estimados a través del coeficiente *H* de Loevinger (1948). Como puede observarse en la Tabla 2, la capacidad discriminativa de los ítems se encuentra en un rango apropiado. La fiabilidad también resulta aceptable aunque decrece en *Hostilidad* y *Ansiedad Fóbica*, dado que son escalas con un número reducido de ítems. Los coeficientes

⁵ Por limitaciones de espacio no es posible presentar aquí la matriz completa de saturaciones factoriales.

Figura 2: Histogramas de las nueve dimensiones primarias de síntomas y de los tres índices globales del SCL-90-R



H se encuentran en el límite de los valores considerados apropiados (0.30), cuestión que puede atribuirse a la distribución de las opciones de respuesta. Respecto a los índices globales, las fiabilidades del GSI, PST y PSDI fueron de = .96, =.94 y =.94, respectivamente.

En la Tabla 3, se presentan las medias, desviaciones típicas y porcentajes de elección de cada alternativa para los siete ítems adicionales.

Luego de revisar las propiedades métricas se calcularon las puntuaciones totales de cada dimensión primaria de

síntomas y de los tres índices globales, de acuerdo al procedimiento descrito previamente en la sección Método. En la Tabla 4 se presentan los estadísticos descriptivos obtenidos, mientras en la Figura 2 se pueden revisar los respectivos histogramas. Examinando los resultados se observa que predominan las asimetrías positivas (baja prevalencia de síntomas), que resulta consistente con lo esperado para una población de estudiantes no pacientes.

Posteriormente, se analizó el efecto del género sobre las escalas, ya que existen numerosos estudios que han

Tabla 5
Correlaciones entre las nueve dimensiones primarias de síntomas para los hombres (triángulo superior) y para las mujeres (triángulo inferior)

	Somatiza.	Obsesiones	Sens. Interp.	Depresión	Ansiedad	Hostilidad	Ans. Fób.	I. Paranoid.	Psicoticismo
Somatiza.	-	0.53	0.41	0.49	0.69	0.35	0.48	0.40	0.45
Obsesiones	0.50	-	0.70	0.73	0.68	0.43	0.54	0.65	0.59
Sens. Interp.	0.27	0.64	-	0.67	0.60	0.36	0.62	0.72	0.65
Depresión	0.50	0.74	0.67	-	0.71	0.42	0.59	0.66	0.66
Ansiedad	0.70	0.66	0.50	0.67	-	0.41	0.66	0.61	0.69
Hostilidad	0.48	0.42	0.38	0.43	0.51	-	0.29	0.41	0.36
Ans. Fób.	0.41	0.52	0.52	0.54	0.64	0.37	-	0.66	0.58
I. Paranoid.	0.31	0.64	0.71	0.62	0.56	0.45	0.60	-	0.67
Psicoticismo	0.41	0.67	0.68	0.68	0.67	0.45	0.59	0.73	-

mostrado prevalencias de sintomatología diferenciales según género. Para ello se calcularon Pruebas t de Student entre los promedios obtenidos por hombres y mujeres en cada una de las nueve dimensiones y en los tres índices globales. La única diferencia significativa ($t=2.31$; $p=0.02$) se encontró en *Hostilidad*, donde las mujeres obtienen una media levemente inferior ($M=0.79$, $SD=0.58$) a los hombres ($M=0.89$, $SD=0.64$). En el resto de las escalas se observan significancias entre $p=0.11$ a $p=0.91$, con una mediana igual a $p=0.56$. Para evaluar la magnitud de la diferencia encontrada en *Hostilidad* se calculó la diferencia media estandarizada, obteniéndose un valor $d=0.16$, que indica un tamaño de efecto despreciable (Cohen, 1988).

Para profundizar en el examen de posibles diferencias entre géneros se compararon las matrices de correlaciones

entre las nueve dimensiones primarias obtenidas para los hombres y para las mujeres. Los resultados son presentados en la Tabla 5. Una revisión minuciosa de cada par de correlaciones indica que los resultados son esencialmente equivalentes para ambos géneros.

Para confirmar esta presunción se sometió a prueba la hipótesis de que ambas matrices de correlaciones eran iguales. En este caso, el objetivo del análisis es determinar si el patrón completo de correlaciones es similar entre géneros. Para ello se utilizó una técnica basada en modelos de ecuaciones estructurales mediante el software LISREL 8.8. Básicamente, el procedimiento consiste en una versión modificada del análisis factorial confirmatorio multigrupo, en que se contrasta la invarianza de dos matrices de correlaciones y de sus respectivas matrices residuales⁶.

Tabla 6
Diferencias medias observadas entre Antofagasta (n=236), Santiago (n=232) y Temuco (n=250) en las escalas del SCL-90-R

Escala	Antofagasta		Santiago		Temuco		F	p	
	Media	SD	Media	SD	Media	SD			
Somatización	0.98	0.63	1.06	0.69	1.15	0.73	3.95	.020	.011
Obsesiones	1.18	0.65	1.26	0.59	1.51	0.73	16.33	< .001	.044
Sens. Interp.	0.87	0.60	0.91	0.54	1.03	0.68	4.22	.015	.012
Depresión	0.98	0.62	1.06	0.54	1.22	0.70	9.46	< .001	.026
Ansiedad	0.95	0.56	0.94	0.54	1.13	0.65	8.78	< .001	.024
Hostilidad	0.75	0.54	0.94	0.64	0.83	0.64	6.01	.003	.017
Ans. Fóbica	0.49	0.44	0.37	0.42	0.57	0.57	10.97	< .001	.030
I. Paranoide	0.81	0.67	0.88	0.65	0.92	0.79	1.40	.246	.004
Psicoticismo	0.70	0.51	0.69	0.51	0.82	0.63	3.83	.022	.011
GSI	0.88	0.47	0.94	0.41	1.06	0.54	9.09	< .001	.025
PSDI	1.69	0.38	1.87	0.37	1.90	0.40	21.41	< .001	.057
PST	44.73	15.79	43.84	14.54	48.29	18.55	5.00	.007	.014

⁶ La sintaxis utilizada puede consultarse en el Anexo a este artículo, disponible en la dirección web: <http://www.sigmas.cl/papers/SCL90/>

Tabla 7
 Normas en Puntuaciones T para las nueve dimensiones primarias de síntomas y para los índices GSI y PSDI

P. Bruta	Somat.	Obses.	S. Interp.	Depresión	Ansiedad	Hostilidad	A. Fóbica	I. Paran.	Psicot.	GSI	PSDI
0.00	24	23	29	27	24	33	42	37	35	12	-
0.05	26	24	32	29	26	35	43	39	37	18	-
0.10	29	26	34	30	28	37	45	40	38	22	-
0.15	32	28	36	32	31	39	46	41	40	26	-
0.20	34	30	38	33	33	40	48	42	41	29	-
0.25	36	32	40	35	35	42	49	43	43	32	-
0.30	38	33	41	37	37	43	50	44	44	34	-
0.35	40	35	42	38	38	45	51	46	46	36	-
0.40	41	36	44	39	40	46	52	47	47	38	-
0.45	43	37	45	41	41	47	54	48	48	39	-
0.50	44	39	46	42	43	48	55	48	49	41	-
0.55	45	40	47	43	44	49	55	49	50	42	-
0.60	46	41	47	44	45	50	56	50	51	43	-
0.65	47	42	48	45	47	51	57	51	52	45	-
0.70	48	43	49	46	48	52	58	52	53	46	-
0.75	49	44	50	47	49	53	59	52	53	47	-
0.80	49	45	50	48	50	53	59	53	54	48	10
0.85	50	45	51	49	50	54	60	53	55	49	14
0.90	51	46	52	50	51	55	61	54	55	50	18
0.95	51	47	53	50	52	55	61	54	56	51	21
1.00	52	47	53	51	53	56	62	55	56	52	24
1.05	52	48	54	52	53	56	62	55	57	53	27
1.10	53	49	55	52	54	57	63	56	57	54	29
1.15	53	49	55	53	55	58	63	56	58	55	31
1.20	54	50	56	53	55	58	64	57	59	56	33
1.25	54	50	57	54	56	59	64	57	59	57	35
1.30	55	51	57	55	57	59	65	58	60	58	37
1.35	55	52	58	55	57	60	66	58	60	59	38
1.40	56	52	58	56	58	60	66	59	61	60	40
1.45	56	53	59	56	58	61	67	59	62	60	41
1.50	57	54	60	57	59	61	67	60	62	61	42
1.55	57	54	60	58	60	62	68	60	63	62	44
1.60	58	55	61	58	60	62	69	61	64	62	45
1.65	59	56	62	59	61	63	69	61	65	63	46
1.70	59	56	62	60	61	63	70	62	65	64	47
1.75	60	57	63	60	62	64	71	63	66	64	49
1.80	61	58	63	61	63	64	71	63	67	65	50
1.85	61	58	64	62	63	65	72	64	68	66	51
1.90	62	59	64	63	64	65	73	64	69	66	52
1.95	62	60	65	63	65	66	74	65	70	67	54
2.00	63	60	66	64	65	67	75	65	71	68	55
2.05	64	61	66	65	66	67	76	66	71	69	56
2.10	64	62	67	65	67	68	76	66	72	69	57

2.15	65	62	67	66	67	68	77	67	73	70	59
2.20	66	63	68	66	68	69	78	67	74	71	60
2.25	66	64	69	67	69	69	79	68	75	72	61
2.30	67	64	69	67	69	70	81	68	76	74	62
2.35	68	65	70	68	70	71	82	68	76	75	63
2.40	68	65	71	68	71	71	83	68	77	76	64
2.45	69	66	72	69	71	72	84	69	78	78	66
2.50	69	66	73	69	72	73	85	69	79	79	67
2.55	70	67	74	70	73	73	86	69	79	80	68
2.60	70	68	75	70	73	74	88	69	80	82	69
2.65	71	68	77	71	74	75	89	69	81	83	70
2.70	71	69	79	72	74	76	90	69	82	84	71
2.75	71	69	80	72	75	76	-	70	82	85	72
2.80	71	70	82	73	75	77	-	70	83	86	72
2.85	72	71	85	75	76	78	-	70	84	87	73
2.90	72	71	87	76	76	79	-	71	85	87	74
2.95	72	72	90	78	77	80	-	71	86	90	75
3.00	72	74	-	80	77	81	-	72	87	-	76
3.05	72	75	-	83	77	83	-	73	88	-	77
3.10	73	76	-	87	78	84	-	74	90	-	77
3.15	73	78	-	90	78	86	-	76	-	-	78
3.20	73	80	-	-	79	88	-	78	-	-	79
3.25	73	83	-	-	79	90	-	81	-	-	80
3.30	73	86	-	-	80	-	-	84	-	-	81
3.35	74	90	-	-	80	-	-	88	-	-	82
3.40	74	-	-	-	81	-	-	92	-	-	83
3.45	75	-	-	-	82	-	-	-	-	-	84
3.50	75	-	-	-	83	-	-	-	-	-	86
3.55	76	-	-	-	84	-	-	-	-	-	87
3.60	77	-	-	-	85	-	-	-	-	-	89
3.65	78	-	-	-	87	-	-	-	-	-	91
3.70	80	-	-	-	88	-	-	-	-	-	-
3.75	82	-	-	-	90	-	-	-	-	-	-
3.80	84	-	-	-	-	-	-	-	-	-	-
3.85	86	-	-	-	-	-	-	-	-	-	-

Tabla 7
Normas en Puntuaciones T para el índice Total de Síntomas Positivos (PST)

P. Bruta	PST	P. Bruta	PST	P. Bruta	PST	P. Bruta	PST
0	18	13	28	26	39	39	47
1	19	14	29	27	39	40	48
2	20	15	30	28	40	41	49
3	20	16	31	29	41	42	49
4	21	17	32	30	41	43	50
5	22	18	32	31	42	44	50
6	23	19	33	32	43	45	51

7	24	20	34	33	44	46	51
8	24	21	35	34	44	47	52
9	25	22	35	35	45	48	52
10	26	23	36	36	45	49	53
11	27	24	37	37	46	50	53
12	28	25	38	38	47	51	54
52	54	65	61	78	70		
53	55	66	61	79	71		
54	55	67	62	80	72		
55	56	68	62	81	73		
56	56	69	63	82	75		
57	57	70	63	83	76		
58	57	71	64	84	78		
59	58	72	65	85	79		
60	58	73	66	86	81		
61	59	74	66	87	83		
62	59	75	67	88	85		
63	60	76	68	89	87		
64	60	77	69	90	89		

Aunque se trata de una hipótesis altamente restrictiva, fue plenamente confirmada por los resultados, observándose una prueba de Chi Cuadrado no significativa ($=50.08$, $p=0.28$) y un índice $RMSEA=0.017$. Adicionalmente, en la muestra de mujeres se obtienen índices de ajuste $RMR=0.025$ y $GFI=0.99$, mientras en la muestra de hombres se observan valores $RMR=0.028$ y $GFI=0.98$.

Una vez establecida la ausencia de diferencias entre géneros, y con miras a la elaboración de normas interpretativas, se exploraron posibles diferencias de promedios entre las zonas geográficas desde las que se obtuvo la muestra. Para ello, se compararon los promedios en las nueve dimensiones primarias y en los tres índices globales alcanzados por los participantes de Antofagasta, Santiago y Temuco a través de Análisis de Varianza unidireccionales (ANOVA). Los resultados, sintetizados en la Tabla 6, muestran diferencias significativas para la casi totalidad de las variables, con la excepción de *Ideación Paranoide*. Las respectivas pruebas post hoc sugieren que no existe un patrón consistente de diferencias (i.e., que una ciudad exhiba sistemáticamente promedios más altos o más bajos), sino que éstas dependen de la dimensión evaluada. Si se consideran los tamaños de efecto mediante el coeficiente, indicados en la última columna de la Tabla 6, puede concluirse que las diferencias, aunque significativas, son de muy baja magnitud y no representan una variación importante entre ciudades.

Una consecuencia práctica de los resultados anteriores (diferencias no significativas entre géneros y de baja magnitud entre ciudades) es que no resulta necesario construir normas diferenciadas por género o por ciudad.

A partir de la conclusión anterior, la última etapa de análisis consistió en elaborar normas interpretativas utili-

zando la muestra total. Aunque la asimetría de las escalas recomienda construir baremos percentiles, se optó por trabajar con puntuaciones T porque son la métrica utilizada convencionalmente con el SCL-90-R y el tipo de escala que emplea el Manual del instrumento, aun cuando las distribuciones allí presentadas también tengan una obvia asimetría positiva (ver Derogatis, 1994). En este sentido, hemos creído preferible respetar el tipo de puntuación más conocida por los usuarios del instrumento, aunque previamente normalizamos las distribuciones de cada dimensión para lograr que las escalas resultantes tengan intervalos iguales y que exista equivalencia entre las puntuaciones T y los respectivos percentiles. Los baremos así construidos se presentan en las Tablas 6 y 7.

Adicionalmente, en la Tabla 8 se presentan algunos datos métricos útiles para complementar el uso de estos datos normativos. En primer lugar, el Error Estándar de Medida (EEM) para las puntuaciones brutas, cuyo uso puede consultarse en Gempp (2006). En segundo lugar, la diferencia mínima significativa (DMS) para las puntuaciones brutas de cada escala. Este índice fue desarrollado por Gullkisen (1950) e indica el puntaje mínimo a partir del cual la diferencia entre dos puntuaciones puede considerarse significativa (i.e., no es producto del error de medida). En el contexto de la evaluación psicológica, la DMS de cada escala puede utilizarse para decidir cuántos puntos de incremento o decremento en la puntuación son necesarios antes de considerar significativo un cambio. Por ejemplo, si se aplica el SCL-90-R antes y después de un proceso psicoterapéutico, los resultados de la Tabla 8 indican que es necesario observar una diferencia mayor a 11 puntos en el Total de Síntomas Positivos (PST) antes de concluir que se ha producido un cambio significativo. En la escala *Ideación*

Tabla 8

Error Estándar de Medida (EEM), Diferencia Mínima Significativa (DMS), Puntuación Bruta equivalente a T=70 y Coeficiente de Livingston para las nueve dimensiones primarias de síntomas y los tres índices globales.

Escala	EEM	DMS	PB=T70	
Somatización	0,29	0,81	2.55	0,97
Obsesiones	0,30	0,84	2.80	0,97
Sensitividad interpersonal	0,30	0,83	2.35	0,96
Depresión	0,27	0,74	2.55	0,97
Ansiedad	0,29	0,80	2.35	0,96
Hostilidad	0,36	1,00	2.30	0,95
Ansiedad fóbica	0,29	0,81	1.70	0,95
Ideación paranoide	0,36	1,00	2.75	0,97
Psicoticismo	0,29	0,81	1.95	0,95
GSI	0,10	0,27	2.15	0,99
PSDI	0,10	0,27	2.65	0,99
PST	4,05	11,22	78	0,99

Paranoide, por otro lado, es necesario un cambio de sólo 1 punto. Cabe recordar, en todo caso, que en la mayoría de las escalas la puntuación total es el promedio y no la suma de las puntuaciones de cada ítem, tal como se explicó en la sección Método.

Otro indicador incluido en la Tabla 8 es la puntuación bruta equivalente a T=70 para cada escala, es decir, el punto a partir del cual un puntaje debería considerarse indicador de riesgo.

Finalmente, para el punto de corte anterior, se indica el valor del coeficiente de Livingston (1972), un tipo especial de coeficiente de consistencia interna que indica la fiabilidad de una decisión dicotómica. En este caso, la fiabilidad de clasificar a un paciente en la categoría “en riesgo” a partir del valor anterior.

Discusión

Los resultados presentados indican que el SCL-90-R exhibe un desempeño psicométrico aceptable, y dentro de los cánones habituales para un autoinforme psicopatológico cuando es aplicado a población universitaria no consultante. Dado que los resultados de este estudio son básicamente descriptivos, dedicaremos este apartado a comentar brevemente los hallazgos que, en nuestra opinión, merecen mayor investigación, y sugerir algunas guías para el uso de estos datos normativos.

Por simplicidad, los aspectos que requieren la atención de futuras investigaciones pueden agruparse en tres áreas: validez, características métricas y desempeño en otras poblaciones.

Respecto a la validez, los resultados aquí presentados aportan evidencia indirecta y parcial sobre la *validez de constructo* del instrumento, tal como ésta es definida en la última versión de los *Standards for Educational and*

Psychological Testing (AERA, APA & NCME; 1999). Por ejemplo, los promedios de adhesión de los ítems (Ver Figura 1) son consistentes con el patrón esperado para síntomas psicológicos medidos en población general. Al mismo tiempo, la estructura factorial (i.e., *validez factorial*), evaluada con una estrategia semi-confirmatoria, resulta congruente con la hipotetizada, mientras que el nivel de escalabilidad de las respuestas también es equivalente a la encontrada en otros estudios (Olsen, et al., 2004; Schmitz et al., 2000). Además, y tal como se esperaba teóricamente, no se observan diferencias en los niveles de sintomatología entre ciudades.

Sin embargo, hay un resultado que resulta discrepante con la evidencia y teoría previa: la ausencia de diferencias entre varones y mujeres. Es un hecho ampliamente demostrado que la prevalencia de malestar psicológico y síntomas psicopatológicos es mayor en las mujeres (McCaffrey, Bauer, O'Bryant & Palav, 2006) y, en el caso del SCL-90-R, toda la evidencia tiende a confirmar este patrón (e.g. Derogatis, 1994). Por ello el presente resultado debe interpretarse con cautela, toda vez que podría indicar un desempeño idiosincrásico de la población en estudio y/o un problema de sensibilidad clínica de los ítems o de su adaptación chilena, lo que comprometería directamente la validez de esta versión. En consecuencia, es recomendable que nuevas investigaciones diluciden este punto, por ejemplo, a través de un estudio de validación independiente con una muestra comparable y/o realizando estudios de *Funcionamiento Diferencial de los Ítems* (DIF) en función del género.

Al mismo tiempo, la validez de un inventario no puede considerarse satisfactoria hasta que se evalúen los diferentes grados de equivalencia transcultural que pueden existir entre la versión adaptada y la versión original del instrumento (Hambleton, Merenda & Spielberger, 2005). Sólo una vez garantizada una *validez transcultural* aceptable se pueden

extrapolar al nuevo contexto las conclusiones obtenidas en la cultura de origen, a partir de la versión original. En el presente caso, para que el cúmulo de antecedentes que respalda la validez del SCL-90-R original pueda aplicarse a la versión adaptada presentada en este trabajo, primero es necesario esclarecer qué tan equivalente resulta la versión chilena con la norteamericana. Si la equivalencia es elevada (alta *validez transcultural*), los hallazgos sobre validez obtenidos con la versión original pueden extenderse directamente a la versión local, sin necesidad de replicar los estudios o hacer nuevas investigaciones. Si, por el contrario, la equivalencia es limitada (baja *validez transcultural*), los resultados originales no serán directamente aplicables a la versión adaptada, obligando a: (a) realizar estudios de validación en el contexto local, y (b) precisar si la falta de equivalencia obedece a diferencias transculturales de la población evaluada o a fallas técnicas en el proceso de adaptación del instrumento. Todo lo anterior advierte sobre la necesidad de evaluar la *validez transcultural* de esta versión, por ejemplo, comparando los indicadores psicométricos de la adaptación y de la versión original, en muestras teóricamente comparables (e.g., muestras universitarias evaluadas con el SCL-90-R original o muestras universitarias evaluadas con otras adaptaciones al castellano).

En suma, los resultados de este estudio son promisorios en cuanto a la validez del SCL-90-R adaptado, pero aún es necesario explorar cuando menos dos líneas de investigación antes de emitir una conclusión al respecto. La primera es explicar la causa de la falta de diferencias en los niveles de malestar psicológico reportados por hombres y mujeres. La segunda es garantizar la validez transcultural de esta versión. Este último punto es crítico porque permitiría extrapolar la evidencia de validez obtenida en estudios internacionales.

Además de la validez, otra área que podría beneficiarse con una exploración exhaustiva en futuras investigaciones atañe a las características métricas del instrumento; específicamente, las referidas a la escala de respuesta. Aunque el formato graduado es ampliamente utilizado en psicología, en general, y en las escalas clínicas, en particular, se ha demostrado que la puntuación total calculada por sumatoria simple de los ítems no siempre representa apropiadamente el nivel de rasgo latente en el caso de los ítems ordinales (van der Ark, 2005). Este resultado es sensible a la asimetría de las distribuciones de respuesta y podría ocasionar distorsiones en el desempeño psicométrico del instrumento (ver, por ejemplo, Gelin & Zumbo, 2003). En los presentes resultados, los valores del coeficiente de escalabilidad (H de Loevinger) inducen a sospechar que no todas las alternativas de respuesta discriminan en forma óptima, especialmente en el caso de las dimensiones de *Ansiedad Fóbica* y *Psicoticismo*, cuyos ítems corresponden a síntomas de baja prevalencia en población no clínica. Por ello es deseable explorar el efecto de utilizar puntuaciones dicotómicas

(ausencia/presencia del síntoma) sobre la calidad métrica de los ítems y la validez del instrumento.

Un tercer punto que resulta necesario abordar en próximas investigaciones es la evaluación del desempeño psicométrico de esta versión en muestras de distintos grupos de edad (e.g., adolescentes y adultos) y en muestras equivalentes de estudiantes universitarios, para establecer el grado de generalización y estabilidad de los resultados aquí presentados.

Para terminar, creemos necesario entregar algunas recomendaciones prácticas para el uso e interpretación de los datos normativos reportados en este estudio. La primera es insistir en que, hasta no acumular más evidencia sobre la validez de esta versión, estas normas sólo pueden considerarse preliminares y no deberían emplearse para formular conclusiones diagnósticas individuales en caso que el SCL-90-R haya sido el único instrumento aplicado. Por el mismo motivo, es prudente sugerir que cuando el instrumento se utilice para fines diagnósticos sea siempre aplicado en conjunto con otro procedimiento de evaluación clínica.

Otra recomendación, dirigida especialmente a usuarios aplicados, es tener presente que las normas se produjeron utilizando puntuaciones T normalizadas, de manera que es fácil encontrar su equivalencia a percentiles en caso de resultar necesario. Sin embargo, un efecto indeseable de este procedimiento es que las normas se basan en una distribución normal idealizada y no en la distribución observada de las escalas.

En línea con lo anterior, aquellos investigadores que deseen utilizar esta versión en estudios correlacionales (e.g., correlación entre alguna dimensión y alguna variable clínicamente relevante) deberían emplear las puntuaciones brutas y no las puntuaciones transformadas en métrica T. Lo mismo vale para aquellos investigadores que pretendan trabajar en estudios experimentales o donde deban comparar promedios de distintos grupos. A estos últimos se les recomienda, además, emplear índices de tamaño de efecto, además de las clásicas pruebas estadísticas.

Finalmente, en caso que este instrumento o sus escalas sean utilizados para evaluar cambio clínico, se recomienda encarecidamente tomar en cuenta los valores del EEM y de la DSM en el análisis e interpretación de los resultados. Utilizados en conjunto, estos índices pueden contribuir a un empleo más apropiado del instrumento.

Referencias

- American Educational Research Association, American Psychological Association & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Bech, P., Malt, U. F., Dencker, S. J., Ahlfors, U. G., Elgen, K., Lewander, T., Lundell, A., Simpson, G. M. & Lingjaerde, O. (1993). Scales for assessment of diagnosis and severity of mental disorders. *Acta Psychi-*

- atrica Scandinavica*, 87, (Suppl. 372), 45-46.
- Barrett, P. T. (2005) *Person-Target Profiling*. En André Beauducel, Bernhard Biehl, Michael Bosnjak, Wolfgang Conrad, Gisela Schönberger, and Dietrich Wagener (Eds.) *Multivariate Research Strategies: a Festschrift for Werner Wittman*. Chapter 4, pp 63-118. Aachen: Shaker-Verlag.
- Bollen, K. A. & Lennox, R. (1991). Conventional wisdom of measurement: A structural equating perspective. *Psychological Bulletin*, 110, 305-314.
- Bonicatto, S., Dew, M. A., Soria, J. J. & Seghezze, M. E. (1997). Validity and reliability of Symptom Checklist '90 (SCL90) in an Argentine population sample. *Social Psychiatry and Psychiatric Epidemiology*, 32, 332-338.
- Bonyng, E. R. (1993). Unidimensionality of SCL-90-R scales in adult and adolescent crisis samples. *Journal of Clinical Psychology*, 49, 212-215.
- Brislin, R. W. (1986). The wording and translation of research instruments. En W. J. Lonner & J. W. Berry (eds.). *Field methods in cross-cultural psychology* (pp. 137-164). Newbury Park, CA: Sage Publications.
- Bulik, C. M., Carpenter, L. L., Kupfer, D. J. & Frank, E. (1990). Features associated with suicide attempts in recurrent major depression. *Journal of Affective Disorders*, 18, 29-37.
- Buller, R., Maier, W. & Benkert, O. (1986). Clinical subtypes in panic disorder: Their descriptive and prospective validity. *Journal of Affective Disorders*, 11, 105-114.
- Cameron, O. G., Thyer, B. A., Nesse, R. M. & Curtis, G. C. (1986). Symptom profiles of patients with DMS-III anxiety disorders. *American Journal of Psychiatry*, 143 (9), 1132-1137.
- Caparrós-Caparrós, B., Villar-Hoz, E., Juan-Ferrer, J. & Viñas-Poch, F. (2007). Symptom Check-List-90-R: fiabilidad, datos normativos y estructura factorial en estudiantes universitarios. *International Journal of Clinical and Health Psychology*, 7 (3), 781-794.
- Carpenter, K. M. & Hittner, J. B. (1995). Dimensional characteristics of the SCL-90-R: evaluation of gender differences in dually diagnosed inpatients. *Journal of Clinical Psychology*, 51, 383-390.
- Casullo, M. M. (2004). Síntomas psicopatológicos en adultos urbanos. *Psicología y Ciencia Social*, 6 (1), 49-57.
- Casullo, M. M. & Castro Solano, A. (1999). Síntomas psicopatológicos en estudiantes adolescentes argentinos. *Anuario de Investigaciones*, 7, 147-157.
- Casullo, M. M., Cruz, M. S., González, R. & Maganto, C. (2003). Síntomas psicopatológicos en adolescentes. Estudio comparativo. *Revista Iberoamericana de Diagnóstico y Evaluación Psicológica*, 16 (2), 135-149.
- Casullo, M. M. & Fernández Liporace, M. (2001). Malestares psicológicos en estudiantes adolescentes. *Psyche*, 10 (1), 155-162.
- Cliff, N. (1966). Orthogonal rotation to congruence. *Psychometrika*, 31, 33-42.
- Coffey, P., Leitenberg, H., Henning, K., Turner, T. & Bennett, R. T. (1996). The relation between methods of coping during adulthood with a history of childhood sexual abuse and current psychological adjustment. *Journal of Consulting and Clinical Psychology*, 64, 1090-1093.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New Jersey: Lawrence Erlbaum.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16 (3), 297-334.
- Cruz, C. S., López, L., Blas, C., González, L. & Chávez, R. A. (2005). Datos sobre la validez y confiabilidad de las Symptom Check List 90 (SCL 90) en una muestra de sujetos mexicanos. *Salud Mental*, 28 (1), 72-81.
- Cyr, J. J., McKenna-Foley, J. M. & Peacock, E. (1985). Factor structure of the SCL-90-R: is there one? *Journal of Personality Assessment*, 49, 571-578.
- Davidson, J. R. T., Kudler, H. S., Saunders, W. B. & Smith, R. D. (1990). Symptom and comorbidity patterns in World War II and Vietnam veterans with posttraumatic stress disorder. *Comprehensive Psychiatry*, 31 (2), 162-170.
- De Jonghe, F., Kool, S., van Aalst, G., Dekker, J. & Peen, J. (2001). Combining psychotherapy and antidepressants in the treatment of depression. *Journal of Affective Disorders*, 64, 217-229.
- De las Cuevas, C., González de Rivera, J. L., Henry-Benítez, M., Monterrey, A., Rodríguez, F. & Gracia, R. (1991). Análisis factorial de la versión española del SCL-90-R en la población general. *Anales de Psiquiatría*, 7, 27-30.
- Derogatis, L. R. (1975). *The SCL-90-R*. Baltimore: Clinical Psychometric Research.
- Derogatis, L. R. (1993). *The Brief Symptom Inventory (BSI). Administration, Scoring and Procedures Manual* (Third Edition). Minneapolis: National Computer Systems.
- Derogatis, L. R. (1994). *SCL-90-R. Administration, Scoring and Procedures Manual* (Third Edition). Minneapolis: National Computer Systems.
- Derogatis, L. R. & Cleary, P. (1977a). Confirmation of dimensional structure of the SCL-90: A study in construct validation. *Journal of Clinical Psychology*, 33(4), 981-989.
- Derogatis, L. R. & Cleary, P. (1977b). Factorial invariance across gender for the primary symptom dimension of the SCL-90. *British Journal of Social & Clinical Psychology*, 16, 347-356.
- Derogatis, L. R. & Coons, H. L. (1993). Self-report measures of stress. En L. Goldberg & S. Breznitz (Eds.). *Handbook of stress* (2nd ed.). New York: The Free Press.
- Derogatis, L. R. & DellaPietra, L. (1994). Psychological tests in screening for psychiatric disorder. In M. Maruish (Ed.). *The use of psychological testing for treatment planning and outcome assessment*. Hillsdale: Lawrence Erlbaum.
- Derogatis, L. R., Lipman, R. S. & Covi, L. (1973). SCL-90: An outpatient psychiatric rating scale—Preliminary report. *Psychopharmacology Bulletin*, 9 (1), 13-27.
- Derogatis, L. R., Lipman, R. S., Rickels, K., Uhlenhuth, E. H. & Covi, L. (1974). The Hopkins Symptom Checklist (HSCL): A self-report symptom inventory. *Behavioral Science*, 19, 1-15.
- Derogatis, L. R. & Melisaratos, N. (1983). The Brief Symptom Inventory: An introductory report. *Psychological Medicine*, 13 (3), 595-605.
- Derogatis, L. R., Meyer, J. K. & King, K. M. (1981). Psychopathology in individuals with sexual dysfunction. *American Journal of Psychiatry*, 138 (6), 757-763.
- Derogatis, L. R., Rickels, K. & Rock, A. (1976). The SCL-90 and the MMPI: A step in the validation of a new self-report scale. *British Journal of Psychiatry*, 128, 280-289.
- Derogatis, L. R. & Savitz, K. L. (2000). The SCL-90-R and Brief Symptom Inventory (BSI) in primary care. In M. E. Maruish (Ed.), *Handbook of psychological assessment in primary care setting* (pp. 297-334). Mahwah: Lawrence Erlbaum.
- Dickson, L. R., Hays, L. R., Kaplan, C., Shelr, E., Abbott, S. & Schmitt, F. (1992). Psychological profile of somatizing patients attending the integrative clinic. *International Journal of Psychiatry in Medicine*, 22 (2), 141-153.
- Flores, L., García, M., Matus, T. & Sandoval, M. (2004). *Figuras estructurales de la violencia escolar. Hacia una recuperación de la "subjetividad" educativa*. (Proyecto Fondecyt 1040694). Santiago: Fondo Nacional de Ciencia y Tecnología.
- Florkowski, C. M., Stevens, I., Joyce, P., Espiner, E. A. & Donald, R. A. (1998). Growth hormone replacement does not improve psychological well-being in adult hypopituitarism: a randomized crossover trial. *Psychoneuroendocrinology*, 23, 57-63.
- Ferrada, L. & Mardones, M. I. (2003). *Alexitimia y su relación con síntomas somáticos y depresivos: una exploración en población universitaria de las comunas de Antofagasta, Santiago y Temuco*. Tesis de Licenciatura en Psicología no publicada, Universidad Mayor, Temuco, Chile.
- García, R. & Seguel, L. (2003). *Síntomas psicopatológicos en adolescentes de Temuco: un estudio exploratorio*. Tesis de Licenciatura en Psicología no publicada, Universidad Mayor, Temuco, Chile.
- Gempp, R. (2003). *Inventario de Síntomas de Derogatis, Revisado (SCL-90-R): Versión de investigación para adolescentes y jóvenes chilenos*. Manuscrito no publicado, Facultad de Psicología, Universidad Mayor.

- Gempp, F. (2006). El error estándar de medida y la puntuación verdadera de los tests psicológicos: Algunas recomendaciones prácticas. *Terapia Psicológica*, 24 (2), 117-130.
- Gelin, M. N. & Zumbo, B. D. (2003). Differential Item Functioning results may change depending on how an item is scored: An illustration with the Center for Epidemiologic Studies Depression Scale. *Educational and Psychological Measurement*, 63 (1), 65-74.
- González de Rivera, J. L., De las Cuevas, C., Rodríguez, M. & Rodríguez F. (2002). *Cuestionario de 90 síntomas SCL-90-R de Derogatis, L.* Adaptación española. Madrid: TEA Ediciones.
- Gulliksen, H. (1950). *Theory of mental tests*. Hillsdale, NJ: LEA.
- Hafkenscheid, A. (1993). Psychometric evaluation of the Symptom Checklist in psychiatric inpatients. *Personality and Individual Differences*, 14, 751-756.
- Hambleton, R. K. (2001). The next generation of the ITC translation and application guidelines. *European Journal of Psychological Assessment*, 17 (3), 164-172.
- Hambleton, R. K., Merenda, P. F., & Spielberger, C. D. (2005). *Adapting educational and psychological tests for cross-cultural assessment*. New Jersey: Lawrence Erlbaum.
- Hamilton, M. (1970). The Hamilton Depression Scale. En W. Guy & R. Bonato (Eds.), *ECDFU Assessment Manual*. Rockville, MD: National Institute of Mental Health.
- Hart, K. M., Bryer, J. B. & Martines, K.A. (1991). SCL-90-R symptom pattern for adolescents psychiatric inpatients. *Journal of Personality Assessment*, 56 (1), 168-178.
- Heerlein, A., Gabler, G., Chaparro, C., Kraus, A., Richter, P. & Berkau, C. (2000). Comparación psicométrica transcultural de la depresión mayor entre Chile y Alemania. *Revista Médica de Chile*, 128 (6), 613-618.
- Heiman, J. R. & Rowlands, D. L. (1983). Affective and physiological sexual response patterns: The effects of instructions on sexually functional and dysfunctional men. *Journal of Psychosocial Research*, 27 (2), 105-116.
- Hoffman, N. G. & Overall, P. B. (1978). Factor structure of the SCL-90 in a psychiatric population. *Journal of Consulting and Clinical Psychology*, 6, 1187-1191.
- Holcomb, W. R., Adams, N. A. & Ponder, H. M. (1983). Factor structure of the symptom checklist-90 with acute psychiatric inpatients. *Journal of Consulting and Clinical Psychology*, 51, 535-538.
- Holi, M. M., Sammallahti, P. R. & Aalberg, V. A. (1998). A Finnish validation study of the SCL-90. *Acta Psychiatrica Scandinavica*, 97, 42-46.
- Holland, J. C., Morrow, G. R., Schmale, A., Derogatis, L., Stefanek, M., Berenson, S., Carpenter, P. J., Breitbart, W. & Feldstein, M. A. (1991). A randomized clinical trial of alprazolam versus progressive muscle relaxation in cancer patients with anxiety and depressive symptoms. *Journal of Clinical Oncology*, 9, 1004-1011.
- Jacobson, N. S., Follette, W. C. & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy*, 15, 336-352.
- Johnson, M. E., Brems, C. & Fisher, D. G. (1996). Self-reported levels of psychopathology and drug abusers not currently in treatment. *Journal of Psychopathology and Behavioral Assessment*, 18, 21-34.
- Johnson, R. W., Ellison, R. A. & Heikkinen, C. A. (1989). Psychological symptoms of counseling center clients. *Journal of Counseling Psychology*, 36, 110-114.
- Kirmayer, L. J., Robbins, J. M., Dworkind, M. & Yaffe, M. J. (1993). Somatization and the recognition of depression and anxiety in primary care. *American Journal of Psychiatry*, 150 (5), 734-741.
- Kellner, R., Hernández, J. & Pathak, D. (1992). Hypochondriacal fears and beliefs, anxiety and somatization. *British Journal of Psychiatry*, 160, 525-532.
- Koeter, M. W. J. (1992). Validity of the GHQ and SCL anxiety and depression scales: a comparative study. *Journal of Affective Disorders*, 24, 271-280.
- Lara, C., Espinosa de Santillana, I., Cárdenas, M. L., Fócil, M. & Cavazos, J. (2005). Confiabilidad y validez de la SCL-90 en la evaluación de psicopatología en mujeres. *Salud Mental*, 28 (3), 42-50.
- Livingston, S.A. (1972). A criterion-referenced application of classical test theory. *Journal of Educational Measurement*, 9, 13-26.
- Loevinger, J. (1948). The technique of homogeneous tests compared with some aspect of 'scale analysis' and factor analysis. *Psychological Bulletin*, 45, 507-530.
- Lucht, M., Jahn, U., Barnow, S. & Freyberger, H. J. (2002). The use of a Symptom Checklist (SCL-90-R) as an easy method to estimate the relapse risk after alcoholism detoxification. *European Addiction Research*, 8, 190-194.
- Martínez-Azumendi, O., Fernández-Gómez, C. & Betia-Fernández, M. (2001). Variabilidad factorial del SCL-90-R en una muestra psiquiátrica ambulatoria. *Actas Españolas de Psiquiatría*, 29 (2), 95-102.
- McCaffrey, R. J., Bauer, L., O'Bryant, S. E. & Palav, A. A. (2006). *Practitioner's Guide to Symptom Base Rates in the General Population*. New York: Springer.
- McCrae, R. R., Zonderman, A. B., Costa, P. T., Bond, M. H., & Paunonen, S. V. (1996). Evaluating replicability of factors in the Revised NEO Personality Inventory: confirmatory factor analysis versus Procrustes rotation. *Journal of Personality and Social Psychology*, 70, 552-566.
- McNair, D. (1974). Self-evaluations of anti-depressants. *Psychopharmacology*, 37, 281-302.
- Morgan, C. D., Wiederman, M. W. & Magnus, R. D. (1998). Discriminant validity of the SCL-90 dimensions of anxiety and depression. *Assessment*, 5, 197-201.
- Muthén, B. & Kaplan, D. (1985). A comparison of some methodologies for the factor analysis of non-normal Likert variables. *British Journal of Mathematical and Statistical Psychology*, 38, 171-189.
- National Computer Systems. (1993). *SCL-90-R bibliography*. Minneapolis: National Computer Systems.
- Olsen, L. R., Mortensen, E. L. & Bech, P. (2004). The SCL-90 and SCL-90-R versions validated by item response models in a Danish community sample. *Acta Psychiatrica Scandinavica*, 110 (3), 225-229.
- Pani, P. P., Maremmanni, I., Pirastu, R., Tagliamonte, A. & Gessa, G. L. (2000). Buprenorphine: a controlled clinical trial in the treatment of opioid dependence. *Drug & Alcohol Dependence*, 60, 39-50.
- Parloff, M. B., Kelman, H. C. & Frank, J. D. (1954). Comfort, effectiveness and self-awareness as criteria of improvement in psychotherapy. *American Journal of Psychiatry*, 111 (5), 343-352.
- Paunonen, S. V. (1997). On chance and factor congruence following orthogonal Procrustes rotation. *Educational and Psychological Measurement*, 57, 33-59.
- Porter, J., Wilson, C. & Frisch, G. R. (1994). Psychotherapy and symptom change. *Psychological Reports*, 75 (3), 1313-1314.
- Primeau, F., Fontaine, R. & Beauclair, L. (1990). Valproic acid and panic disorder. *Canadian Journal of Psychiatry - Revue Canadienne de Psychiatrie*, 35, 248-250.
- Prusoff, B. A., Weissman, M. M., Klerman, G. L. & Rousanville, B. J. (1980). Research diagnostic criteria subtypes of depression: Their role as predictors of differential response to psychotherapy and drug treatment. *Archives of General Psychiatry*, 37, 796-801.
- Robles, J., Andreu, J. & Peña, M. (2002). SCL-90: Aplicación y análisis de sus propiedades psicométricas en una muestra de sujetos clínicos españoles. *Sicopatología Clínica, Legal y Forense*, 2, 1-19.
- Rosenberg, R., Bech, P., Mellegard, M. & Ottosson, J. O. (1991). Secondary depression in panic disorder: An indicator of severity with a weak effect on outcome in alprazolam and imipramine treatment. *Acta Psychiatrica Scandinavica*, 83 (Suppl. 365), 39-45.
- Rief, W. & Fichter, M. (1992). The Symptom Check List SCL-90-R and its ability to discriminate between dysthymia, anxiety disorders and anorexia nervosa. *Psychopathology*, 25, 128-138.
- Schauenberg, H. (1999). Measuring psychotherapeutic change with the Symptom Check List SCL-90-R. *Psychotherapy and Psychosomatics*, 68, 199-206.
- Schmitz, N., Hartkamp, N. & Franke, G. H. (2000). Assessing clinically significant change: application to the SCL-90-R. *Psychological Reports*, 86, 263-274.
- Schmitz, N., Hartkamp, N., Kruse, J., Franke, G. H., Reister, G. & Tress, W. (2000). The Symptom Check-List-90-R (SCL-90-R): A German validation study. *Quality of Life Research*, 9, 185-193.

- Schmitz, N., Kruse, J. & Tress, W. (2001). Improving screening for mental disorders in the primary care setting by combining GHQ-12 and the SCL-90-R subscales. *Comprehensive Psychiatry*, 42, 166-173.
- Schmitz, N., Kruse, J., Heckrath, C., Alberti, L. & Tress, W. (1999). Diagnosing mental disorders in primary care: The General Health Questionnaire (GHQ) and the Symptom Check List (SCL-90-R) as screening instruments. *Social Psychiatry and Psychiatric Epidemiology*, 34, 360-366.
- Schönemann, P. H. (1966). A generalized solution of the orthogonal Procrustes problem. *Psychometrika*, 31, 1-10.
- Schwarzwald, J., Weisenberg, M. & Solomon, Z. (1991). Factor invariance of SCL-90-R: The case of combat stress reaction. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 3, 385-390.
- Shutty, M. S. Jr., DeGood, D. E. & Schwartz, D. P. (1986). Psychological dimensions of distress in chronic pain patients: a factor analytic study of Symptom Checklist-90 responses. *Journal of Consulting and Clinical Psychology*, 54, 836-842.
- Silva, J. & Gempp, R. (2004). Diferenciación emocional y psicopatología: La hipótesis de la acentuación somática. *Revista Electrónica de Motivación y Emoción*, 6 (16). [<http://reme.uji.es/articulos/avilij3023006105/texto.html>]
- Simon, G. E. & VonKorff, M. (1991). Somatization and psychiatric disorder in the NIMH epidemiologic catchment area study. *American Journal of Psychiatry*, 148 (11), 1491-1500.
- Steer, R. A., Clark, D. A. & Ranieri, W. F. (1994). Symptom dimensions of the SCL-90-R: A test of the tripartite model of anxiety and depression. *Journal of Personality Assessment*, 62, 525-536.
- Steer, R. A., Platt, J. J., Ranieri, W. F. & Metzger, D. S. (1989). Relationships of SCL-90 profiles to methadone patients' psychosocial characteristics and treatment response. *Multivariate Experimental Clinical Research*, 9 (2), 45-54.
- Strauman, T. J. & Wetzler, S. (1994). The factor structure of SCL-90 and MCMI scale scores: within measure and interbattery analyses. *Multivariate Behavioral Research*, 27, 1-20.
- Strayer, D. R., Carter, W. A., Brodsky, I., Cheney, P., Peterson, D., Salvato, P., Thompson, C., Loveless, M., Shapiro, D. E. & Elsasser, W. (1994). A controlled clinical trial with a specifically configured RNA drug, Poly(I)-Poly(C)[12]U, in chronic fatigue syndrome. *Clinical Infectious Diseases*, 18 (Suppl.1), S88-95.
- Swedo, S. E., Rettew, D. C., Kuppenheimer, M., Lum, D., Dolan, S. & Goldberger, E. (1991). Can adolescent suicide attempters be distinguished from at-risk adolescents? *Pediatrics*, 88 (3), 620-629.
- Swett, C., Surrey, J. & Cohen, C. (1990). Sexual and physical abuse histories and psychiatric symptoms among male psychiatric outpatients. *American Journal of Psychiatry*, 147 (5), 632-636.
- Tanzer, N. K., & Sim, C.Q.E. (1999). Adapting instruments for use in multiple languages and cultures: A reviews of the ITC guidelines for tests adaptations. *European Journal of Psychological Assessment*, 15(3), 258-269.
- Tood, D. M., Deane, F. P. & Mc.Kenna, P. A. (1997). Appropriateness of SCL-90-R adolescent and adult norms for outpatient and nonpatient college students. *Journal of Counseling Psychology*, 44 (3), 294-301.
- Toomey, T. C., Seville, J. L., Mann, J. D., Abashian, S. W. & Grant, J. R. (1995). Relationship of sexual and physical abuse to pain description, psychological distress and health care utilization in a chronic pain sample. *Clinical Journal of Pain*, 11, 307-315.
- Turner, L. A., Althof, S. E., Levine, S. B. & Bodner, D. (1991). External vacuum devices in the treatment of erectile dysfunction: A one-year study of sexual and psychosocial impact. *Journal of Sex & Marital Therapy*, 17 (2), 81-93.
- van der Ark, A. (2005). Stochastic ordering of the latent trait by the sum score under various polytomous IRT models. *Psychometrika*, 70(2), 283-304.
- van de Vijver, F., & Hambleton, R. K. (1996). Translating tests: Some practical guidelines. *European Psychologist*, 1 (2), 89-99.
- van der Sande, R., van Rooijen, L., Buskens, E., Allart, E., Hawton, K., van der Graaf, Y. & van Engeland, H. (1997). Intensive in-patient community intervention versus routine care after attempted suicide. A randomized controlled intervention study. *British Journal of Psychiatry*, 171, 35-41.
- Vassend, O. & Skrandal, A. (1999). The problem of structural indeterminacy in multidimensional symptom report instruments. The case of SCL-90-R. *Behavior Research and Therapy*, 37, 685-701.
- Vollrath, M., Koch, R. & Angst, J. (1990). The Zurich Study: IX. Panic disorders and sporadic panic: Symptoms, diagnosis, prevalence and overlap with depression. *European Archives of Psychiatric & Neurological Science*, 239 (4), 221-230.
- Waryszak, Z. (1982). Symptomatology and social adjustment of psychiatric patients before and after hospitalization. *Social Psychiatry*, 17, 149-154.
- Waskow, I. & Parloff, M. (1975). *Psychotherapy change measures*. Rockville: National Institute of Mental Health.
- Wetzler, S., Kahn, S., Cahn, W., Van Praag, H.M. & Asnis, G.M. (1990). Psychological test characteristics of depressed and panic patients. *Psychiatry Research*, 31 (2), 179-192.
- Wider, A. (1948). *The Cornell Medical Index*. San Antonio: Psychological Corp.
- Wolfe, J., Brown, P. J. & Bucsele, M. L. (1992). Symptom responses of female Vietnam veterans to Operation Desert Storm. *The American Journal of Psychiatry*, 149 (5), 676-679.
- Woodworth, R. S. (1918). *The personal data sheet*. Chicago: Stoelting.