

Evaluación y revisión de cuestionarios para uso en inmigrantes Latinos con bajo grado de alfabetización

Karen T. D'Alonzo¹

En la medida en que un número mayor de inmigrantes de lengua española participa de investigaciones, surgen cuestiones relacionadas a la adecuación de instrumentos ya existentes. Cuestionarios, en versiones inglesas, son frecuentemente adaptados y comprobados con americanos de origen hispana con nivel superior. Poco ha sido escrito sobre pruebas y evaluación de esos instrumentos de investigación entre inmigrantes con grados más bajos de escolaridad. El objetivo de este estudio fue evaluar y revisar una batería de cuestionarios para intervención con mujeres inmigrantes hispanas. Se utilizó un proceso de tres etapas para evaluar, adaptar y comprobar las versiones españolas de los instrumentos Self-Efficacy and Exercise Habits Survey, versión abreviada para inmigrantes del Hispanic Stress Inventory y Latina Values Scale. Los instrumentos revisados demostraron validez y confiabilidad aceptables. Las adaptaciones mejoraron la comprensión de los instrumentos, aumentando la tasa de respuestas y reduciendo datos omisos y respuestas extremas. Las limitaciones psicométricas de la adaptación de escalas del tipo Likert también son abordadas.

Descriptores: Psicometría; La Salud Inmigrante; Disparidades de Salud.

¹ Enfermera, Doctor en Enfermería, Profesor Asistente, Rutgers, The State University of New Jersey, College of Nursing, New Jersey, USA. E-mail: kdalonzo@rutgers.edu.

Correspondencia:

Karen T. D'Alonzo
Rutgers, The State University of New Jersey
College of Nursing
180 University Avenue, Rm. 226
Newark, NJ USA 07102
E-mail: kdalonzo@rutgers.edu

Avaliação e revisão de questionários para uso entre imigrantes latinos com baixos níveis de alfabetização

À medida que número maior de imigrantes de língua espanhola participa de pesquisas, surgem questões relacionadas à adequação de instrumentos já existentes. Questionários, em versões inglesas, são frequentemente adaptados e testados com americanos de origem hispânica com nível superior. Pouco tem sido escrito sobre testes e avaliação desses instrumentos de pesquisa entre imigrantes com níveis mais baixos de escolaridade. O objetivo deste estudo foi avaliar e revisar uma bateria de questionários para intervenção com mulheres imigrantes hispânicas. Utilizou-se um processo de três etapas para avaliar, adaptar e testar as versões espanholas dos instrumentos Self-Efficacy and Exercise Habits Survey, versão abreviada para imigrantes do Hispanic Stress Inventory e o Latina Values Scale. Os instrumentos revisados demonstraram validade e confiabilidade aceitáveis. As adaptações melhoraram a compreensão dos instrumentos, aumentando a taxa de respostas e reduzindo dados omissos e respostas extremas. Limitações psicométricas da adaptação de escalas do tipo Likert também são abordadas.

Descritores: Psicométricas; Saúde de Imigrantes; Disparidades de Saúde.

Evaluation and revision of questionnaires for use among low-literacy immigrant Latinos

As more Spanish speaking immigrants participate in and become the focus of research studies, questions arise about the appropriateness of existing research tools. Questionnaires have often been adapted from English language instruments and tested among college-educated Hispanic-Americans. Little has been written regarding the testing and evaluation of research tools among less educated Latino immigrants. The purpose of this study was to evaluate and revise a battery of Spanish-language questionnaires for an intervention among immigrant Hispanic women. A three-step process was used to evaluate, adapt and test Spanish versions of the Self-Efficacy and Exercise Habits Survey, an abbreviated version of the Hispanic Stress Inventory-Immigrant version and the Latina Values Scale. The revised tools demonstrated acceptable validity and reliability. The adaptations improved the readability of the tools, resulting in a higher response rate, less missing data and fewer extreme responses. Psychometric limitations to the adaptation of Likert scales are discussed.

Descriptors: Psychometrics; Immigrant Health; Health Disparities.

Introducción

Los latinos o individuos de descendencia hispana comprenden casi 16% de la población de los Estados Unidos de América (EE.UU) y es la minoría que ha crecido más rápidamente en el país⁽¹⁾. Aproximadamente 40% de más de 45 millones de latinos que viven en los EE.UU son extranjeros y cerca de 75% de la primera generación de inmigrantes latinos hablan español, la mayor parte del tiempo. Además de eso, muchos inmigrantes latinos tienen bajos niveles de educación formal y probablemente

no están familiarizados con la terminología científica. Cuanto más inmigrantes de lengua española participan de investigaciones y se tornan el enfoque de estudios en los EE.UU, más cuestiones son levantadas a respecto de cuan apropiados son los instrumentos ya existentes. Cuestionarios elaborados para ser usados con latinos, viviendo en los EE.UU, han sido traducidos del inglés y comprobados con las segunda y tercera generaciones de americanos de origen hispana con nivel superior. Ese

abordaje asume que las personas de descendencia hispana forman un grupo homogéneo y bien educado. Poco ha sido escrito al respecto de la evaluación y pruebas de instrumentos de investigación entre inmigrantes latinos con niveles educacionales más bajos, particularmente aquellos menos familiarizados con el proceso de investigación.

Cuestiones metodológicas

Algunos investigadores han informado, hace ya algún tiempo, dificultades para conducir investigaciones con poblaciones hispanas⁽²⁻³⁾. Además de cuestiones de traducción, existe un número de cuestiones culturales de naturaleza metodológica que preocupa a los investigadores no hispanos, trabajando con latinos⁽⁴⁾. El sesgo cultural con latinos, asociado al uso de escalas tipo Likert, es frecuentemente citado⁽²⁾. Las escalas Likert son populares en ciencias sociales, pero los inmigrantes latinos con bajo nivel educacional pueden tener poco entendimiento del formato de respuestas tipo escala^(3,5). Tendencias para respuestas específicas incluyen respuestas de conveniencia social⁽⁶⁾, conjunto de respuestas extremas (uso excesivo de los puntos extremos de la escala)⁽⁷⁾ y datos omitidos⁽⁸⁾. Por último, las variables poblacionales no pertinentes pueden afectar las respuestas de latinos en las escalas Likert, incluyendo edad, nivel educacional, aculturación y país de origen^(2,3,7). Los adultos más jóvenes, con más años de educación y puntajes más altos de aculturación, son menos propensos a informar dificultades al responder escalas Likert⁽³⁾. Históricamente, ha habido poca concordancia entre estudiosos en relación a la forma más apropiada para lidiar con esos dilemas metodológicos^(4,9).

El objetivo de este estudio fue evaluar y revisar una batería de cuestionarios en la lengua española para intervención con mujeres inmigrantes hispanas. El estudio consiste de tres fases: fase 1) evaluación de los instrumentos existentes, fase 2) revisión de los instrumentos y fase 3) evaluación preliminar de las propiedades psicométricas de los instrumentos revisados. En las fases 1 y 2, el investigador principal estaba interesado en evaluar cuan apropiado era el formato de respuesta del instrumento, evidencia de cuestiones de conveniencia social, estilos de respuestas extremas, datos omitidos y tiempo necesario

para llenarlo. La fase 3 enfocó la prueba piloto y las revisiones, y también la evaluación de las propiedades psicométricas de los instrumentos adaptados.

Método

Fase 1) evaluación de los instrumentos existentes - una muestra intencional de 13 mujeres fue seleccionada de un grupo de inmigrantes latinos que frecuentaban un programa de entrenamiento para agentes (promotoras) de salud de la comunidad en New Jersey, EE.UU. Las estadísticas descriptivas de los sujetos en las fases 1, 2 y 3 se presentan en la Tabla 1.

Instrumentos para recolección de datos: 1) *Self-Efficacy and Exercise Habits Survey*⁽¹⁰⁾ (SEHS), 2) *Latina Values Scale* (LVS)⁽¹¹⁾ y 3) versión abreviada para inmigrantes del *Hispanic Stress Inventory* (HSI-I)⁽¹²⁾. El SEHS y el HSI-I habían sido previamente traducidos para el español con buena validez y confiabilidad entre poblaciones de lengua española, en los EE.UU. Una versión española del LVS no estuvo disponible durante la fase 1, entonces el autor y tres consultores bilingües de la comunidad tradujeron el instrumento para el español, lo cual fue posteriormente retro-traducido para el inglés. Las versiones inglesa y española de los cuestionarios fueron comparadas y eventuales diferencias en los términos y cuestiones de exactitud conceptual fueron resueltas.

Self-Efficacy and Exercise Habits Survey⁽¹⁰⁾ - El SEHS mide la autoeficacia relacionada al ejercicio físico en relación a dos factores: "encontrando tiempo para ejercitarse" y "resistiendo a recaídas". Los coeficientes alfa de 0,83 y 0,85 para el SEHS han sido informados con confiabilidad de prueba y reprobación de 0,78. Ambas versiones, la inglesa y española, del SEHS tienen una escala Likert de seis puntos; las opciones 1 y 2 corresponden a "estoy segura que no puedo", las opciones 3 y 4 representan "quizás si puedo", en cuanto las opciones 5 y 6 corresponden a "estoy segura que puedo". La versión inglesa desarrollada con estudiantes universitarios contiene 12 ítems, con puntuación posible de 12 a 72. La versión española de 15 ítems fue desarrollada con adultos residentes en la comunidad y tiene puntuación posible de 15 a 90⁽¹³⁾. Puntuaciones más altas son asociadas a un grado más elevado de autoeficacia en relación al ejercicio físico.

Tabla 1 – Características de la muestra

Descripción	Grupo 1 (n=13)	Grupo 2 (n=6)	Grupo 3 (n=81)
Promedio de edad (de)	39,8 (11,76)	31,5 (7,71)	29,8 (7,92)
Promedio de años viviendo en los EE.UU (de)	9,23 (5,10)	8,67 (2,73)	10,36 (6,78)

(continúa...)

Tabla 1 - *continuación*

Descripción	Grupo 1 (n=13)	Grupo 2 (n=6)	Grupo 3 (n=81)
País de origen	México (31%)	México (100%)	México (76%)
	Honduras (25%)		Guatemala (4%)
	Perú (25%)		Otros: Nicaragua, El Salvador, Ecuador, Colombia
	Otros: Puerto Rico, Ecuador, República Dominicana, Colombia		
Promedio de años de educación (de)	10,33 (2,87)	9,17 (2,86)	9,04 (1,86)
Renta familiar			
Menos que US\$21,000/año	31%	50%	26%
US\$21,000-US\$35,000/año	23%	17%	5%
US\$35,000-US\$50,000/año	15%	0%	6%
No sabe	31%	33%	36%

Latina Values Scale (LVS)⁽¹¹⁾ - esta escala Likert, de dos etapas con 31 ítems, fue desarrollada con mujeres hispanas en los EE.UU para medir creencias relacionadas al marianismo (papeles tradicionales atribuidos a las mujeres en la cultura hispana). El LVS mostró confiabilidad entre ítems de 0,87. Una correlación significativa inversa ($r=-0,65$, $p=0,01$) fue observada entre los puntajes del LVS y los informados en una medida de asertividad. Tres subescalas emergieron del análisis factorial exploratorio (AFE): responsabilidad, asertividad y satisfacción. La "principal escala" de la versión original en inglés es una escala Likert, de seis puntos, con respuestas que varían de 1 - "disuerdo totalmente" a 6 - "conuerdo totalmente". Puntajes en la escala principal varían de 31 a 186. Puntajes más altos están asociados a creencias más fuertes en el marianismo. La escala "satisfacción" contiene 31 ítems que corresponden a los ítems de la escala principal. Los sujetos establecen "cuan satisfechos están con su respuesta" para cada pregunta que responden. El sujeto escoge una de las cuatro opciones, que varían de 1 - "estoy muy insatisfecho" a 4 - "estoy muy satisfecho". La puntuación en la escala satisfacción varía de 31 a 124. Puntuaciones más altas indican que el sujeto está bastante comfortable con sus respuestas. Un borrador de la versión española del instrumento fue creado como descrito previamente.

Versión abreviada para inmigrantes del Hispanic Stress Inventory (HSI-I)⁽¹²⁾ - el instrumento de 17 ítems fue desarrollado para inmigrantes latinos. Las cinco subescalas enfocan el estrés ocupacional o económico, estrés parental, estrés conyugal, estrés relacionado a la inmigración y estrés familiar/cultural. La consistencia interna, en todas las subescalas, varió de 0,68 a 0,83. La validez convergente del instrumento revisado corrobora relaciones positivas moderadas, a través de medidas autorrelatadas de niveles de humor relacionados a depresión, ansiedad y rabia. Las entrevistadas respondían "sí" o "no" si ellas hubiesen experimentado o no una

serie de factores de estrés relacionados a la aculturación en los tres meses anteriores. Si la respuesta es "sí", la participante, era entonces, instruida a indicar (en una escala de 1 a 5) cuan preocupada o tensa la situación la hizo sentirse. El puntaje 'uno' significa que la participante "no se sintió extremadamente preocupada/tensa" [*nada preocupada/tensa*], en cuanto el puntaje 'cinco' indica que la participante se sintió "extremadamente preocupada/tensa" [*muy preocupada/tensa*]. La puntuación del HSI-I varía de 17 a 85. Puntajes más altos corresponden a niveles más altos de ansiedad.

Después del término de consentimiento libre e informado, y del esclarecimiento del estudio, cada una de las mujeres recibió un paquete conteniendo los cuestionarios en orden aleatoria. Las participantes fueron instruidas a responder las cuestiones, verificar la necesidad de cambios gramaticales e identificar cualquier pregunta que ellas la sintiesen como ambigua o excesivamente sensible. Tres asistentes de investigación bilingües ayudaron a las participantes que solicitaron explicaciones o la lectura de las preguntas. Los datos fueron procesados y el análisis demográfico y psicométrico fue realizado a través del SPSS, versión 16.0 para Windows (SPSS Inc. Chicago. Ill).

Resultados

Apenas ocho de las 13 mujeres fueron capaces de completar los cuestionarios dentro de una hora. El nivel de alfabetización y los años de educación fueron factores prominentes en el llenado de los cuestionarios. Las cuatro participantes con menos de seis años de educación formal tenían más chances de omitir ítems múltiples. Seis de las mujeres tuvieron dificultad para responder preguntas con dos negativas. Validez y confiabilidad no fueron evaluadas por causa del pequeño tamaño de la muestra y de la preponderancia de datos omitidos.

Todas las mujeres expresaron descontento con el formato Likert. Mitad de los cuestionarios mostró evidencia de la preferencia de las participantes por respuestas extremas. De la misma forma, mitad de las mujeres optó por ítems que no las colocase bajo una perspectiva desfavorable (respuestas socialmente deseables) o simplemente respondieron todas las cuestiones con la misma respuesta (efecto halo). Los cuatro paquetes llenados contenían cantidades considerables de datos omitidos. Apenas cuatro de los cuestionarios SEHS fueron llenados completamente. Todos contenían un gran número de respuestas extremas con tendencia a ser positivas en situaciones en donde ejercitarse puede ser difícil ("estoy segura que si puedo"). De esa forma, la puntuación total del SEHS fue bastante alta (promedio=41,5, de=2,4), a pesar de que ninguna de las mujeres practicase ejercicios regulares. El LVS fue el instrumento más largo y contenía algunas preguntas de naturaleza sexual, que las mujeres consideraron "my personales" para responder. Muchas de las participantes no respondieron la segunda parte de cada pregunta (subescala satisfacción), que era componente crucial del cuestionario. Muchas de aquellas que no respondieron afirmaron más tarde que habían entendido la frase *¿cuán satisfecha está usted con su respuesta?* como *¿cuán segura está usted con su respuesta?* Para no parecer indecisas, la mayoría respondió *estoy muy satisfecha* sin entender completamente la naturaleza de la pregunta. Más de la mitad de las mujeres omitió la segunda parte de cada pregunta en el HSI-I y tres de los cuestionarios contenían respuestas con fuerte efecto halo.

Fase 2) revisión de los instrumentos - el autor principal se encontró con dos consultores bilingües de la comunidad y un investigador latino con experiencia, antes de las revisiones, para discutir las dificultades encontradas, durante el llenado de los cuestionarios. Basado en los resultados de la fase 1, los siguientes cambios fueron sugeridos e implementados: 1) las escalas Likert fueron reducidas a tres puntos: si/no/no-sé o no estoy segura; 2) un conjunto de instrucciones simples y ejemplos de preguntas y respuestas fueron adicionadas a cada cuestionario; 3) todos los instrumentos fueron administrados en la misma orden, comenzando con el más simple y terminando con el más largo y/o con más preguntas sensibles; 4) las mujeres de la comunidad local fueron entrenadas para actuar como asistentes de investigación, de modo que había una razón de 2:1 de sujetos/asistentes; 5) el cuidado de los niños fue providenciado durante la recolección de datos; 6) las escalas fueron escritas en la misma dirección; 7) los instrumentos fueron rescritos usando una fuente mayor con más "espacio en blanco" al rededor de cada pregunta

y 8) el LVS fue substituido por el *Latina Values Scale* revisado (LVS-R)⁽¹⁴⁾, que es una versión española con 28 ítems, desarrollado para latinas jóvenes y de media edad. La escala satisfacción fue substituida en el LVS-R por la escala conflicto y la segunda parte de cada pregunta fue rescrita como: *¿Su respuesta a esta pregunta ha causado problemas o conflictos en su vida?* Las participantes del estudio sintieron que esa frase reflejaba mejor el objetivo del cuestionario. El Alfa de Cronbach para el LVS-R fue 0,94 y 0,95 para la escala conflicto. Además de la escala conflicto, el análisis factorial exploratorio del LVS-R⁽¹⁴⁾ reveló otros seis factores: autosacrificio, asertividad, culpa, autocensura, colocando los otros en primer lugar y responsabilidad. La presencia de siete factores en el LVS-R, al contrario de apenas tres factores en el LVS, muestra la complejidad del concepto del marianismo⁽¹⁴⁾. Todos los instrumentos revisados fueron adaptados para asegurar un nivel de lectura en español equivalente a la sexta serie. Una vez que las autorizaciones fueron obtenidas de los autores, la validez aparente de los instrumentos revisados y 'cuan culturalmente apropiados estos eran' fueron evaluados por consultores de la comunidad antes que los mismos fuesen administrados a seis agentes comunitarias (promotoras). Las características de esa muestra están descritas en la Tabla 1.

Resultados: todas - seis participantes - completaron la batería total de cuestionarios. El tiempo promedio para completar los cuestionarios fue de 35 minutos y todas las participantes respondieron cerca de 100% de las preguntas. La prueba psicométrica no fue realizada por causa del pequeño número de sujetos.

Fase 3 - propiedades psicométricas de los instrumentos revisados - los instrumentos revisados fueron ministrados a 81 mujeres inmigrantes hispanas, con edad entre 18 y 55 años, y que participaron del estudio *Physical Activity Intervention for Latinas* [Intervención de Actividad Física para Latinas] (manuscrito en andamio). Las *promotoras* auxiliaron los sujetos leyendo y explicando el significado de las preguntas seleccionadas y comprobando si todas las preguntas habían sido respondidas. El tiempo promedio necesario para completar los instrumentos fue de 40 minutos, con menos de 2% de información omitida para todas las mujeres.

Psicometría: la confiabilidad de los instrumentos revisados fue evaluada a través del ítem corregido - total de correlaciones y consistencia interna de cada instrumento y subescalas adaptadas. La validez del constructo fue evaluada a través de análisis factorial confirmatorio (AFC), conducido para cada instrumento revisado a través del eje principal de extracción con rotación Varimax⁽¹⁵⁾. Antes de realizar la AFC, los presupuestos del análisis

factorial fueron comprobados y verificados. La prueba de esfericidad de Bartlett, que mide la fuerza de la relación entre las variables, y la medida Kaiser-Meyer-Olkin (KMO), para verificar la adecuación de la muestra, fueron usados para acceder a la factorabilidad de la matriz de correlación⁽¹⁵⁾. En la teoría, *a priori*, los instrumentos y autovalores originales guiaron la extracción de factores (factores con autovalores por abajo de 1,0 fueron

excluidos). Los ítems con carga factorial mayores de 0,32 fueron mantenidos⁽¹⁵⁾. La diferencia mínima de 0,20 entre la carga de un ítem en factores teóricamente alineados y opuestos fue recomendada. Gráficos de declive fueron usados para subsidiar decisiones relacionadas al factor de extracción. Los resultados de los análisis psicométricos son presentados en las Tablas 2, 3 y 4.

Tabla 2 – Cargas factoriales para análisis factorial confirmatorio con rotación Varimax del *Self-Efficacy for Exercise Habits Survey* (n=81)

Ítems	Encontrando tiempo para ejercitarse	Resistiendo a las recaídas	Comunalidad
Llamo a mis vecinos o amigos para caminar conmigo con regularidad [pedirle a sus vecinos o amigos ir a caminar con Ud., regularmente]	0,69	-0,50	0,85
Llamo a mis amigos para ejercitarse conmigo [pedirle a sus amigos que hagan ejercicios con Ud.]	0,48	0,27	0,63
Me ejercito con mi familia [hacer ejercicios con su familia]	0,58	0,24	0,67
Mantengo un programa de ejercicios inclusive cuando surgen emergencias en el trabajo [seguir con su programa de ejercicios aunque tenga demasiadas exigencias en el trabajo]	0,65	0,06	0,82
Mantengo un programa de ejercicios inclusive cuando obligaciones sociales toman mucho tiempo [seguir con su programa de ejercicios aunque tenga obligaciones sociales que le toman mucho tiempo]	0,71	0,25	0,71
Comienzo un programa de ejercicios después de un intervalo (como, por ejemplo, Navidad o cuando llueve) [empezar su programa de ejercicios de después de una interrupción (como la Navidad o cuando llueve)]	0,58	0,33	0,78
Reservo tiempo para ejercitarme por 30 minutos tres veces por semana [apartar tiempo para un programa de ejercicio, por lo menos 30 minutos tres veces por semana]	0,61	0,24	0,78
Cambio mi horario de alimentación para que mi ejercicio sea más conveniente [cambiar su horario de comer para hacer más conveniente el ejercicio]	0,18	0,50	0,52
Me levanto temprano para ejercitarme [levantarse más temprano para hacer ejercicios]	0,13	0,55	0,65
Reservo tiempo para actividad física [planear el ejercicio dentro su horario]	0,18	0,57	0,47
Camino en lugar de manejar en distancias cortas [caminar en vez de manejar distancias cortas]	0,34	0,42	0,71
Voy a una fiesta apenas después de ejercitarme [ir a una fiesta sólo después de haber hecho sus ejercicios]	0,31	0,57	0,33
Me ejercito durante mi horario de alimentación [hacer ejercicios durante la hora de comida]	-0,08	0,60	0,63
Mantengo un programa de ejercicio cuando el trabajo me exige mucho y estoy cansada [seguir con su programa de ejercicios después de un día de trabajo largo y cansado]	0,43	0,62	0,71
Me ejercito cuando me siento deprimida [hacer sus ejercicios aunque se sienta deprimido]	0,34	0,52	0,41
Autovalor	4,6	1,7	
Porcentaje de variancia explicada	30,7	11,5	

Tabla 3 – Factor de carga para análisis factorial confirmatorio con rotación Varimax de la *Latina Values Scale Revised* (LVSR) (n=81)

Ítems	Conflicto	Auto sacrificio	Asertividad	Culpa	Auto censura	Poniendo los otros en primer lugar	Responsabilidad	Comunalidad
Me veo haciendo cosas para los otros que preferiría no hacer [me encuentro a mi misma haciendo cosas para otros que preferiría no hacer]	0,73	0,13	0,01	0,15	0,01	0,16	0,03	0,59
Me siento orgullosa cuando los otros me elogian por los sacrificios que he hecho [me siento orgullosa cuándo otros me halagan por los sacrificios que he hecho]	0,52	0,26	0,41	0,32	0,23	0,08	-0,12	0,69
Frecuentemente, asumo responsabilidades con mi familia, que yo preferiría no hacerlo, porque eso me hace sentir una persona mejor [a menudo tomo responsabilidades familiares que preferiría no tomar, porque me hace sentir una mejor persona]	0,66	0,09	0,18	0,02	0,34	-0,09	-0,05	0,60
Yo, frecuentemente, me veo haciendo cosas que harán a mi familia feliz mismo cuando sé que no es lo que me gustaría hacer [yo a menudo me encuentro haciendo cosas que hacen feliz a mi familia, aún cuando se no es lo que quiero hacer]	0,75	0,03	-0,01	0,01	0,09	-0,03	0,16	0,60
Ser visto por los otros como una "buena persona" es muy importante para mí [ser visto como una "buena persona" por otros es muy importante para mí]	0,03	0,89	-0,02	0,02	0,06	0,04	0,01	0,80
Creo que el sacrificio que hago por los otros será un día recompensado [creo que mi sacrificio podría ser recompensado algún día]	0,13	0,83	0,09	0,11	0,16	0,08	0,25	0,82
Considero a mi familia una fuente importante de apoyo [considero mi familia una gran fuente de apoyo]	0,10	0,07	0,83	0,11	0,08	0,12	-0,01	0,74
La familia es importante para mí [la familia es muy importante para mí]	0,23	0,01	0,79	0,12	0,01	0,04	0,05	0,72
Yo frecuentemente asumo responsabilidades que son relacionadas a mi familia [a menudo tomo responsabilidades familiares]	0,34	0,24	0,17	0,46	-0,31	0,26	-0,47	0,81
Yo tengo dificultad para expresar mi rabia [tengo dificultad expresando mi enojo]	0,09	0,05	0,10	0,70	0,07	0,02	-0,18	0,56
Creo que sacrificarme por los otros me hace una persona mejor [creo que sacrificarse por otros te hace una mejor persona]	0,04	0,23	0,46	0,50	-0,03	0,26	0,39	0,70

(continúa...)

Tabla 3 - *continuación*

Ítems	Conflicto	Auto sacrificio	Asertividad	Culpa	Auto censura	Poniendo los otros en primer lugar	Responsabilidad	Comunalidad
Me siento una persona mala cuando sé que alguien está enojado o desilusionado conmigo [yo me siento como una persona terrible cuando sé que alguien está molesto o desilusionado conmigo]	0,03	0,03	0,11	0,74	0,02	0,03	0,26	0,64
Encuentro difícil decir "no" a las personas mismo cuando "no" es lo que quiero decir [encuentro difícil decir "no" a la gente, aún cuando se que decir "no" es lo que debería]	0,17	0,05	0,18	0,13	0,61	0,28	0,18	0,56
Yo tengo dificultades en afirmarme delante de figuras de autoridad [tengo dificultades para hacerme valer ante personas de autoridad]	0,10	0,25	0,29	0,24	0,58	0,04	0,09	0,56
Me siento culpado cuando le pido a los otros que hagan cosas por mi [me siento culpable cuando le pido a otros que hagan cosas por mi]	0,34	0,02	-0,15	0,01	0,66	0,07	-0,23	0,64
Frecuentemente me siento inferior en relación a figuras de autoridad [a menudo me siento menos ante las personas de autoridad]	0,65	0,33	-0,21	0,49	0,53	0,13	0,09	0,70
Trato de hacer a los otros felices a cualquier costo [trato de hacer feliz a los otros a toda costa]	0,11	0,02	-0,19	0,12	0,23	0,75	-0,05	0,68
Me veo colocando las necesidades de los otros al frente de las mías [considero primero las necesidades y los otras personas antes de las mías]	0,14	0,02	0,15	0,05	0,05	0,66	-0,02	0,51
Me veo colocando las necesidades de mi familia delante de las mías [considero primero las necesidades de mi familia antes de las mías]	0,01	0,37	0,01	0,03	-0,10	0,54	0,39	0,60
Hacer a mi compañero feliz me hace sentir mejor respecto a mi misma [hacer feliz a mi pareja me hace sentir bien conmigo misma]	0,17	0,25	0,04	0,14	0,03	0,03	0,82	0,78
Autovalor	2,1	2,1	2,0	1,8	1,6	1,6	1,5	
Porcentaje de variancia explicada	11,1	10,4	10,3	9,0	8,2	7,3		

Tabla 4 – Carga factorial para análisis factorial confirmatorio con rotación Varimax de la versión para inmigrantes del *Hispanic Stress Inventory* (HSI-I) (n=81)

Ítem	Estresores	Estresores	Comunalidad
	intrafamiliares	extrafamiliares	
La falta del inglés torna la socialización difícil [por no saber suficiente inglés ha sido difícil para mí socializar con otros]	0,61	0,03	0,38
Las personas esperan que yo trabajé más pesado porque soy latina [debido a que soy latina se ha esperado que trabaje más rápido o duro]	0,53	0,39	0,40
No creo que gano dinero suficiente para sustentar a mi familia o a mi misma [mis ingresos no han sido suficientes para mantener a mi familia o a mi misma]	0,64	0,16	0,43
Encontrar un trabajo que me gustaría ha sido difícil porque soy latina [debido a que soy latina he tenido dificultad para encontrar el tipo de trabajo que quiero]	0,63	0,52	0,67
A veces siento que si los otros esperan que yo sea perezosa [he tenido que ser cuidadosa con la calidad de mi trabajo para que otros no piensen que soy una perezosa]	0,70	0,31	0,59
Promociones o aumentos de salarios son difíciles de conseguir porque soy latina [debido a que soy latina, ha sido difícil obtener ascensos o aumentos de salario]	0,67	0,35	0,57
Tengo que aceptar empleos que pagan mal [me he visto forzada a aceptar trabajos con salarios bajos]	0,75	0,11	0,58
Me he sentido presionada a aprender inglés [me he sentido presionada (o forzada) para aprender inglés]	0,67	-0,02	0,45
Pienso que si fuese a una agencia social o gubernamental sería deportada [he pensado que si voy a una agencia social o del gobierno sería deportada]	0,57	-0,05	0,32
Mi compañero y yo discordamos en cómo educar nuestro hijos [mi esposo y yo hemos tenido desacuerdos acerca de como criar a nuestros hijos]	0,05	0,75	0,56
Las personas me tratan mal por causa de mi inglés deficiente [debido a mi mal inglés, la gente me ha tratado mala]	0,39	0,59	0,48
Mis hijos no respetan mi autoridad de la forma que deberían [mis hijos(as) no respetan mi autoridad en la forma que deberían]	-0,02	0,77	0,59
Siento que las ideas de mis hijos sobre sexualidad son muy liberales [he sentido que las ideas de mis hijos(as) acerca de la sexualidad son demasiado liberales]	0,14	0,12	0,03
Han habido conflictos (o desentendimientos) entre los miembros de la mi familia [ha habido conflictos o malos desentendidos entre miembros de mi familia]	0,15	0,49	0,26
Ha habido violencia física entre los miembros de mi familia [ha habido violencia física entre miembros de mi familia]	0,09	0,55	0,31
Mis hijos han mencionado salir de casa [mis hijos(as) han hablado acerca de irse de la casa]	-0,20	0,49	0,28
Mis hijos ha recibido boletines escolares malos (o notas malas) [mis hijos(as) han recibido malas calificaciones o informes en la escuela]	0,20	0,41	0,21
Yo tuve pelias serias con miembros de mi familia [he tenido serios problemas con miembros de mi familia]	0,17	0,35	0,15
Autovalor	5,2	2,0	
Porcentaje de variancia explicada	28,9	11,4	

Self-Efficacy and Exercise Habits Surve - la escala Likert fue reducida a tres frases de referencia: 1) "estoy segura que no puedo"; 2) "tal vez yo pueda"; 3) "estoy segura que puedo". La versión española revisada del SEHS contenía 15 ítems, con puntajes entre 15 y 45. El puntaje promedio obtenido en el cuestionario fue 37,5 (de=4,81), variando de 28 a 45. El histograma de los datos reveló asimetría de -0,22 y curtosis de -1,0. A pesar de que la estadística descriptiva no fue calculada en una muestra anterior, parece que hubo una tendencia menor de las mujeres para escoger respuestas positivas extremas con el instrumento revisado. Todos los coeficientes de correlación ítem para subescala total variaron de 0,41 a 0,72. El Alfa de Cronbach para el SEHS fue de 0,81. El KMO fue de 0,64 y la prueba de esfericidad de Barlett fue estadísticamente

significativa ($p < 0,001$), ambos aceptables⁽¹⁵⁾. Los resultados de el AFC son presentados en la Tabla 2. Dado que apenas dos factores tuvieron autovalores por arriba de 1,0 y el gráfico de declive subsidió una solución de dos componentes, el AFC reveló esa misma solución de dos factores en la versión inglesa⁽¹⁰⁾. La consistencia interna para cada una de las dos subescalas "encontrando tiempo para ejercitarse" y "resistiendo las recaídas" fue satisfactoria 0,75. Esos dos factores fueron responsables por 42,2% de la variancia total.

Latina Values Scale Revised - veinte y tres de los 28 ítems del LVSR fueron mantenidos y comprobados. Los cinco ítems omitidos fueron de naturaleza sexual y no relevantes para el estudio o fueron considerados redundantes. La escala Likert de cinco puntos, usada en el LVRS original,

fue reducida a tres frases de referencia: 1) "No estoy de acuerdo"; 2) "No estoy seguro"; 3) "Estoy de acuerdo". Las cinco frases usadas en la escala conflicto del LVSR original fueron reducidas a tres: 1) "nunca"; 2) "algunas veces"; 3) "siempre". El puntaje posible total variaba de 23 a 69. El puntaje promedio del LVSR fue 45 (de=7,23), variando de 33 a 69. El histograma reveló ligera inclinación negativa -0,28 y curtosis de -0,78, sugiriendo que las participantes no tendieron a adherir al estilo de respuestas extremas. El alfa de Cronbach para el LVSR fue aceptable a 0,74. Dos ítems tuvieron correlaciones ítem para subescala abajo del nivel aceptable de 0,40: "Frecuentemente me siento inferior en relación a los hombres" [*a menudo me siento inferior en comparación a los hombres*] y "Pienso que la crítica es consecuencia de mis fallas" [*creo que los conflictos y problemas son mi culpa*]. Con base en discusión informal con un subgrupo de mujeres, la primera pregunta debería haber sido omitida de este estudio y era relacionada a otros ítems que tenían temas relacionados a la sexualidad. Con relación a la segunda pregunta, las mujeres sintieron que la redacción de la versión española de la pregunta era confusa y evocaba un tipo diferente de respuesta de la que era pretendida por la versión inglesa. Como resultado, ambos ítems fueron excluidos del análisis de este estudio. El KMO fue 0,50, indicando tamaño de muestra mínimamente aceptable⁽¹⁶⁾, en cuanto la prueba de esfericidad de Barlett fue estadísticamente significativa ($p < 0,001$). El AFC subsidio el modelo de siete factores, responsable por 66,5% de la variancia. La consistencia interna para cada una de las subescalas fue $> 0,70$. Los resultados del AFC aparecen en la Tabla 3.

Hispanic Stress Inventory - la primera mitad de cada pregunta conteniendo dos partes, enfocando fuentes de estrés, relacionadas a la inmigración, no fue alterada. La escala Likert de cinco puntos, usada en la segunda mitad de cada pregunta, fue reducida a tres: 1) "ni un poco preocupada"; 2) "un poco preocupada" y 3) "muy preocupada". El puntaje promedio de la escala principal fue 4,06 (de=3,95), indicando que la mayoría de las entrevistadas informó pocos momentos de estrés relacionados a la inmigración. El intervalo posible de puntajes para la segunda mitad del instrumento revisado era de 0 a 54. El puntaje promedio para la segunda mitad de las cuestiones fue 7,79 (de= 9,41), variando de 0 a 36. Menos de 5% de valores omitidos fue observado en la versión revisada. Los coeficientes de correlación ítem para subescala fueron aceptables, variando de 0,41 a 0,77. El Alfa de Cronbach para todo el instrumento fue 0,74, el KMO fue de 0,67 y la prueba de esfericidad de Barlett $p < 0,001$. Dado que apenas dos factores tenían

autovalores por arriba de 1,0, AFC subsidio o modelo de dos factores⁽¹³⁾ de los factores de estrés intrafamiliar y extrafamiliar. La consistencia interna para las dos subescalas fue 0,72 y 0,76, respectivamente. Esos dos factores fueron responsables por 40,3% de la variancia total. El gráfico de declive confirmó la solución de dos factores. La carga del ítem "yo creo que las ideas de mis hijos sobre sexualidad son muy liberales" no fue muy alta en ninguna de las escalas, probablemente por causa de la poca edad de los padres. Los resultados son presentados en la Tabla 4. El ítem adicionado al instrumento adaptado "yo creo que si fuese al servicio social o a una agencia gubernamental, sería deportada" mostró coeficiente de correlación ítem para subescala de 0,42. La consistencia de la subescala extrafamiliar se mantuvo la misma, no importando si el ítem fue o no incluido, lo que sugiere un buen ajuste de la subescala.

Discusión

Los resultados corroboran estudios anteriores que indican que el uso de escalas Likert entre inmigrantes latinos es generalmente problemático. Eso es verdad mismo cuando los instrumentos son previamente traducidos para el español y presentan validez y confiabilidad adecuadas entre individuos de lengua española, viviendo en los EE.UU. Este estudio demostró que las versiones españolas de los instrumentos adaptados pueden ser usadas con suceso entre grupos de inmigrantes con bajo nivel de alfabetización y que no están familiarizados con procesos de investigación. Las adaptaciones parecen mejorar la comprensión de los instrumentos, resultando en una tasa de respuesta mayor, menos datos omitidos y tendencia menor para escoger respuestas extremas.

A pesar de que la reducción del número de puntos en la escala Likert de cinco para tres parece haber resuelto algunas de las ambigüedades de las participantes, también redujo la variabilidad de las respuestas. En el proceso de AFC, los factores aislados explicaron, apenas, 40 a 65% de la variancia total. A pesar de que es deseable la identificación de factores suficientes responsables por 80 a 90% de la variancia, ese criterio puede ser tan bajo como 50%, cuando el objetivo es explicar la variancia con el menor número de factores posible⁽¹⁷⁾. Dado que la AFC de cada uno de los tres instrumentos confirmaron la estructura factorial original, esto presta subsidios para el formato de los instrumentos adaptados. La pregunta de un número óptimo de puntos en una escala Likert es una pregunta controvertida; cinco a siete puntos son generalmente considerados ideales para asegurar

la completa representación de las respuestas, pero tres puntos pueden ser suficientes si se enfatiza los datos del grupo en lugar de los datos individuales. En este estudio, el uso de más puntos del que los sujetos eran capaces de comprender podría haber resultado en mayor variabilidad, pero no necesariamente en mayor validez o confiabilidad. Otra pregunta potencial es la pérdida de la variabilidad cuando una escala Likert de tres puntos es usada en un formato de pre-prueba/pos-prueba. Cuando existe la necesidad de capturar cambios, un número mayor de puntos y sujetos puede ser necesarios para obtener una información completa antes de iniciar la recolección de datos. Es digno de nota que las participantes hubiesen tenido poca dificultad para llenar otros instrumentos en el idioma español (como, por ejemplo, el *Center for Epidemiologic Studies Depression Scale* (CES-D)), usando escalas de múltiple elección cuando estas eran más cuantificables. Ese mismo formato podría ser implementado en la elaboración de nuevos instrumentos para medir conceptos de actividad física. En lugar de preguntar cuan seguro el sujeto está en relación a despertar temprano para ejercitarse, la pregunta podría ser redactada como "¿Cuántas veces en la semana usted siente que podría despertar temprano para ejercitarse?". En ese caso, las respuestas podrían basarse en el número de días en la semana (por ejemplo 0, 1-2, 3-4, 5-6, 7). Ese abordaje mejoraría la claridad y posiblemente preservaría la variabilidad de las respuestas. En cuanto las mujeres de este estudio tenían, en promedio, nueve años de educación, 21% frecuentó apenas hasta el sexto grado. Dato que muestra que la capacidad de alguien conversar en una lengua es más probable de que tener familiaridad con terminología científica; no fue una sorpresa que muchas de las participantes tuvieron dificultades para completar los instrumentos en la fase 1. Los investigadores que trabajan con poblaciones de inmigrantes deberían considerar que los sujetos probablemente no son familiarizados con el lenguaje de investigación. De esa forma, tiempo y atención adicional son necesarios para ofrecer instrucciones sobre la recolección de datos. Una limitación de la fase 3 del estudio fue el tamaño pequeño de la muestra. Valores de KMO para dos de los tres instrumentos sugieren que el tamaño de la muestra era probablemente adecuado. Si hay cuatro o más variables con cargas por arriba de 0,60, el patrón puede ser interpretado no importando el tamaño de la muestra usada, los tamaños de las muestras serán adecuados cuando las comunalidades son altas (>60) mismo que estos estén bien abajo de 100. Es recomendable repetir el AFC de los instrumentos con tamaños de muestras mayores.

Conclusión

Adaptar instrumentos de investigación ya existentes continúa siendo un desafío significativo para instrumentar a los investigadores que trabajan con inmigrantes latinos. Dada la creciente diversidad de la población de los EE.UU, los investigadores en el área de enfermería precisan considerar la variedad de métodos culturalmente apropiados para incentivar la participación de la población de inmigrantes, con bajo nivel de alfabetización, en estudios de investigación. La colaboración transnacional entre investigadores del área comprometidos con la salud de inmigrantes latinoamericanos puede ser una forma de identificar métodos culturalmente apropiados.

Referencias

1. US Census Bureau, Population Division. Annual Estimates of the Resident Population by Sex, Race, and Hispanic Origin for the United States: April 1, 2000 to July 1, 2009; 2010. Retrieved June 15, 2010, from <http://www.census.gov/popest/national/asrh/NC-EST2009-srh.html>.
2. Flakerud, JH. Is the Likert scale format culturally biased? *Nurs Res.* 1988;37(3):185-6.
3. Bernal H, Wooley S, Schensul, JJ. Methodology corner: the challenge of using Likert-type scales with low-income ethnic populations. *Nurs Res.* 1997;46(3):170-81.
4. Lange JW. Methodological concerns for non-Hispanic investigators conducting research with Hispanic Americans. *Res Nurs Health.* 2002;25:411-9.
5. McQuiston C, Larson K, Parrado, EA, Flakerud, JH. AIDS knowledge and measurement considerations with unacculturated Latinos. *West J Nurs Res.* 2002;24(4):354-72.
6. Marin G, Triandis H C, Betacourt H, Kashima, Y. Ethnic affirmation versus social desirability: Explaining discrepancies in bilinguals' responses to a questionnaire. *J Cross Cult Psychol.* 1983; 14:173-86.
7. Marin G, Gamba RJ, Marin BV. Extreme response style and acquiescence among Hispanic Americans. *J Cross Cult Psychol.* 1992;23:498-509.
8. Aday LA, Chiu GY, Andersen, R. Methodological issues in health care surveys of the Spanish heritage population. *Am J Publ Health.* 1980;70:367-74.
9. Cavazos-Rehg P, Zayas L, Spitznagel E. Legal status, emotional well-being and subjective health status of Latino immigrants. *J Natl Med Assoc.* 2007;99(10):1126-31.
10. Sallis JF, Pinski PB, Grossman RM, Patterson TI, Nader PR. The development of self-efficacy scales for health-related diet and exercise behaviors. *Health Educ Res.* 1988;3:283-92.

11. Marano MR. The creation of the Latina Values Scale: An analysis of marianismo's effects on Latina women attending college. Dissertation Abstracts International: Section A. Humanities and Social Sciences; 2004.
12. Cavazos-Rehg PA, Zayas LH, Walker MS, Fisher EB. Evaluating an abbreviated version of the Hispanic Stress Inventory for immigrants. *Hispanic J Behav Sci.* 2006;28(4):498-515.
13. McArthur LH, Holbert D, Pena, M. Development and application of rapid assessment diet and physical activity indexes, which suggest high consumption of energy-dense foods and inadequate exercise among adolescents from six Latin American cities: a pilot study. *Nutr Res.* 2008;28(9):590-9.
14. Melendez FA. (2004). The Latina Value Scale Revised: a cross-sectional analysis of marianismo in Latino women. Dissertation Abstracts International: Section A. Humanities and Social Sciences; 2008.
15. Tabachnick BG, Fidell LS. Using multivariate statistics. 5th ed. Boston, MA: Allyn and Bacon; 2007.
16. Kaiser HF, Rice J. Little Jiffy, Mark IV. Educational and Psychological Measurement. *Educ Psychol Meas.* 1974;34:111-7.
17. Stevens JP. Applied multivariate statistics for the social sciences. 4th ed. Mahwah, NJ: Lawrence Erlbaum; 2002.

Recibido: 18.1.2011

Aceptado: 15.8.2011

Como citar este artículo:

D'Alonzo KT. Evaluación y revisión de cuestionarios para uso en inmigrantes Latinos con bajo grado de alfabetización. *Rev. Latino-Am. Enfermagem* [Internet]. sep.-oct. 2011 [acceso: ____/____/____];19(5):[12 pantallas]. Disponible en:

URL

____/____/____
 día año
 mes abreviado con punto