

Validación de la Escala de Gravedad de Fatiga en Población General de la Ciudad de México

Validation of the Fatigue Severity Scale in general population of Mexico City

Eduardo Velasco-Rojano *¹, Rocío E. Duarte-Ayala², Angelica Riveros-Rosas³,
Juan José Sánchez-Sosa¹, L. Isabel Reyes-Lagunes¹

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

2 - Universidad del Valle de México, Campus Lomas Verdes, México.

3 - Universidad Nacional Autónoma de México, Facultad de Contaduría y Administración, México.

Recibido: 07/08/2017 Revisado: 24/08/2017 Aceptado: 04/09/2017

Introducción
Objetivos
Justificación
Método
Resultados
Discusión
Conclusión
Referencias

Resumen

Para conocer la prevalencia y las consecuencias de la fatiga se necesitan formas confiables y válidas para medirla. Una de las escalas más usadas es la Escala de Gravedad de Fatiga (FSS; Krupp, LaRocca, Muir-Nash, & Steinberg, 1989), por lo que el propósito de este trabajo fue validar esta escala psicométricamente de manera culturalmente relevante, identificar si su estructura es la más adecuada a través de un análisis factorial confirmatorio y comprobar la invarianza factorial por sexo. Se trabajó con dos grupos de participantes, 150 habitantes de la Ciudad de México para la validación y 220 habitantes de la Ciudad de México para el análisis factorial confirmatorio y prueba de invarianza. Se encontró una estructura con seis reactivos, adecuada confiabilidad ($\alpha = .93$) correcto ajuste [$\chi^2_{(9)} = 12.56, p = .183$; CF1 = .99; RMSEA = .05] e invarianza por sexo, por lo que se concluye que la Escala de Gravedad de Fatiga es una medida confiable y válida.

Palabras clave: *fatiga, medición, validación psicométrica, invarianza, sexo*

Abstract

Reliable and valid ways to measure fatigue are required in order to determine its prevalence and consequences. One of the methods widely used to measure fatigue is the Fatigue Severity Scale (FSS; Krupp, LaRocca, Muir-Nash, & Steinberg, 1989). Thus, the purpose of this study was to validate this scale psychometrically in a culturally relevant manner, to identify whether its structure is the most appropriate through a confirmatory factor analysis, and to test the invariance by sex. Two groups of participants, inhabitants of Mexico City, were included in the present study, 150 for validation studies, and 220 for confirmatory factor analysis and invariance test studies. A structure with six items, adequate reliability ($\alpha = .93$), goodness of fit [$\chi^2_{(9)} = 12.56, p = .183$; CF1 = .99; RMSEA = .05] and invariance by sex was found. So, it is concluded that FSS is a reliable and valid measure of fatigue.

Key words: *fatigue, assessment, psychometric validation, invariance, gender*

*Correspondencia a: Eduardo Velasco Rojano. Universidad Nacional Autónoma de México, Facultad de Psicología, Edificio D, Cubículo 10 Mezzanine. Avenida Universidad 3004, Coyoacán, Copilco Universidad, 04510 Ciudad de México, CDMX. México. Teléfono: 0155 56 22 23 23. eduardorojanova@gmail.com

Cómo citar este artículo: Velasco-Rojano, E., Duarte-Ayala, R. E., Riveros-Rosas, A., Sánchez-Sosa, J. J., & Reyes-Lagunes, L. I. (2017). Validación de la Escala de Gravedad de Fatiga en población general de la Ciudad de México. *Revista Evaluar*, 17(2), 113-125. Recuperado de <https://revistas.unc.edu.ar/index.php/revaluar>

Nota del autor: Una versión preliminar de la adaptación aplicada a otra muestra se presentó en el XVI Congreso Mexicano de Psicología Social. Para este estudio, la escala se estudió en participantes distintos en dos nuevas muestras, por lo que los aquí presentados son datos diferentes y no hay conflicto de interés.

Introducción

La *fatiga* puede entenderse como la sensación de cansancio físico y falta de energía que ocurre como consecuencia de la realización de esfuerzos físicos o mentales o de la existencia de enfermedad (Gottschalk et al., 2005; Viner & Christie, 2005). La fatiga prolongada puede tener consecuencias en las relaciones sociales; en la salud mental, al asociarse con estrés y depresión; en la memoria, y en la motricidad fina (Malta-Purim, Bittencourt-Gumarães, Kapp-Titski, & Leite, 2016).

Existe confusión alrededor del término fatiga ya que se presenta ligado con otros como astenia, cansancio, agotamiento, apatía, aburrimiento, psicastenia, pesimismo y vulnerabilidad, como señalan Fuentes-Márquez, Senín-Calderón, Rodríguez-Testal y Carrasco (2015). Esta situación se agrava al considerar que la definición del concepto es imprecisa y no existe un consenso general acerca de la posibilidad de considerarse un proceso adaptativo de las personas sanas, un síntoma físico, un síntoma psicológico o un síndrome (Fuentes-Márquez et al., 2015).

Al considerar esta información, puede entenderse que no sea fácil caracterizar apropiadamente el alcance de las molestias y consecuencias de la fatiga, ni su prevalencia en la población general, lo que se acentúa porque la mayoría de las quejas por fatiga suelen ser desestimadas ya sea por los propios pacientes o por los médicos que las escuchan. Esto se debe a que se trata de un padecimiento con síntomas no específicos, con lo que se considera que alrededor del 70 por ciento de los casos en etapas tempranas permanecen sin detectar y atender (Van't Leven, Zielhuis, Van der Meer, Verbeek, & Bleijenberg, 2009). Así, es necesario realizar estudios para conocer la prevalencia y las consecuencias de la fatiga en la población general y, además, para realizar este tipo de investigación es necesario contar con formas

confiables y válidas para medirla.

La fatiga puede ser evaluada de dos maneras distintas: con aparatos para realizar distintas pruebas físicas (Ali et al., 1994) o con escalas de autoinforme (Gottschalk et al., 2005; Kos et al., 2005), que es la que más se utiliza dado que las mediciones con aparatos son más complejas y costosas (Gottschalk et al., 2005; Kos et al., 2005).

Sin embargo, el desarrollo de escalas de autoinforme es difícil debido a la variabilidad en la experiencia de las molestias informadas (Armutlu et al., 2007; Gottschalk et al., 2005; Kos et al., 2005). Existe un amplio número de instrumentos diseñados para medir la fatiga, pero no existe suficiente investigación sobre sus propiedades psicométricas (Whitehead, 2009). En el caso particular de la población hispanohablante, la falta de estudios sobre los instrumentos hasta ahora adaptados y/o utilizados en experiencias empíricas de diagnóstico y medición también ha llevado a un vacío en la investigación sobre la relación de los puntajes de fatiga con otras variables y sus posibles consecuencias.

En la población hispanohablante se han utilizado diversos instrumentos para medir la fatiga, como el Test de Patrones Subjetivos de Fatiga (PSF; Yoshitake, 1971), el cual se ha empleado para evaluar fatiga en ambientes laborales en México (Orrantia-Daniel & De la Vega-Bustillos, 2003) y Colombia (Neisa, Tangarife, & Godoy, 2014). El PSF consta de 30 reactivos dicotómicos que se responden por *sí* o por *no*, y miden la percepción subjetiva de la fatiga laboral en tres dimensiones: la exigencia mental, los síntomas físicos y los síntomas mixtos.

La subescala de fatiga subjetiva de la Checklist for Individual Strength (CIS; Beurskens et al., 2000), usada en Colombia para evaluar fatiga en ambientes laborales (Useche, Cendales, & Gómez, 2017), mide percepción de fatiga subjetiva relacionada con aspectos conductuales y está

conformada por ocho reactivos con siete opciones de respuesta sobre grado de acuerdo con la situación actual del participante. El coeficiente alfa de Cronbach de la escala fue de $\alpha = .86$. La Need For Recovery Scale (NFR; De Croon, Sluiter, & Frings-Dresen, 2006) también ha sido utilizada para evaluar fatiga en ambientes laborales en Colombia (Useche et al., 2017) y mide los síntomas de fatiga física y mental con 11 reactivos dicotómicos con una confiabilidad de $\alpha = .80$.

La Chronic Fatigue Scale (CFS; Chalder et al., 1993), que mide las dos dimensiones de fatiga física y fatiga mental, con 11 reactivos tipo Likert con cuatro opciones de respuesta y una consistencia interna (α) entre .88 y .90, se ha utilizado para comparar el síndrome de fatiga crónica en población anglosajona e hispanohablante en Estados Unidos (Torres-Harding, Mason-Schutter, & Jason, 2008).

Otro instrumento de los más utilizados es la Escala de Gravedad de la Fatiga (FSS; Krupp et al., 1989), la cual mide gravedad de fatiga de manera unidimensional con nueve reactivos tipo Likert con siete opciones de respuesta en términos de acuerdo y desacuerdo. Esta escala se ha utilizado y validado psicométricamente en población general y en pacientes con distintos tipos de enfermedades crónicas, como esclerosis múltiple, lupus, hepatitis C y pacientes con VIH (Lerdal et al., 2011). Además, la escala FSS se ha adaptado a diversos idiomas, como el turco (Gök, Cengiz, Erol, & Özgöçmen, 2016), el italiano (Ottonello, Pellicciari, Giordano, & Foti, 2016) y el español colombiano (Bernal-Vargas, Riveros-Munévar, Vinaccia-Alpi, & Quiceno-Sierra, 2017), y en todos los casos se encontraron adecuadas propiedades psicométricas en términos de discriminación, confiabilidad y validez (Whitehead, 2009).

La ventaja principal de esta escala es su breve extensión. Los nueve reactivos originales miden si la fatiga afecta la motivación, el ejer-

cicio y el funcionamiento físico, las actividades de la vida cotidiana y social. Existen versiones con siete reactivos desarrolladas en pacientes con VIH y esclerosis múltiple angloparlantes que no incluyen las consecuencias de la fatiga en la motivación o el ejercicio físico (Johansson, Kottorp, Lee, Gay, & Lerdal, 2014; Lerdal et al., 2011). Por otra parte, existe una versión con ocho reactivos evaluada en pacientes de esclerosis múltiple italo-parlantes que no incluye consecuencias sobre la motivación (Ottonello et al., 2016).

La FSS ha sido utilizada previamente en México para evaluar fatiga en programas de educación médica (Duarte-Ayala, 2017), pero no existen estudios de validación en población general de la Ciudad de México.

Objetivos

El presente estudio tuvo como objetivos: validar psicométricamente una versión adaptada culturalmente relevante de la Escala de Gravedad de Fatiga en población general de la Ciudad de México, comprobar cuál es la estructura factorial más adecuada con análisis factorial confirmatorio, y realizar una prueba de invarianza por sexo para determinar si su funcionamiento difiere en hombres y mujeres.

Justificación

Con el presente estudio se validará el instrumento de fatiga FSS a fin de determinar su prevalencia en la población general de la Ciudad de México. Además, nuestros resultados pueden ser válidos para el diagnóstico de la fatiga en etapas tempranas, lo que hasta el momento es una necesidad identificada por la literatura del campo (Van't Leven et al., 2009).

Método

Tipo de estudio

Se llevó a cabo un estudio instrumental psicométrico (Montero & León, 2007).

Participantes

Se trabajó con dos grupos de participantes. El grupo de validación psicométrica estuvo conformado por 180 habitantes de la Ciudad de México, 95 mujeres y 85 hombres, entre 18 y 60 años, con una edad promedio de 29.84 años y una desviación estándar de 10.75. Con respecto a la escolaridad de los participantes, 10 habían finalizado la escuela primaria, 14 la escuela secundaria, 35 la preparatoria, 100 la licenciatura y 21 habían culminado estudios de posgrado. En cuanto al estado civil, 132 eran solteros y 48 eran casados. Para determinar el tamaño de la muestra se decidió exceder 10 veces el número de reactivos, de acuerdo con la propuesta de Nunnally y Bernstein (1994), lo que se corroboró con un adecuado valor ($> .60$) de la medida de adecuación muestral Kaiser-Meyer-Olkin.

El grupo de análisis factorial confirmatorio y prueba de invarianza estuvo conformado por 220 habitantes de la Ciudad de México, 110 mujeres y 110 hombres, entre 18 y 60 años, con un promedio de 32.33 años y una desviación estándar de 11.60. Respecto a la escolaridad de los participantes, 7 habían finalizado la primaria, 22 la secundaria, 65 la preparatoria, 107 la licenciatura y 19 un posgrado, 110 eran solteros y 110 casados. El tamaño de la muestra se consideró con base en un análisis de potencia estadística calculado para GFI (MacCallum & Hong, 1997), en el que se consideró tener nueve reactivos, 45 grados de libertad, un coeficiente alfa de Cronbach de .05 y una potencia deseada de .80, lo que resultó en

157 participantes. Se decidió exceder 20 veces el número de reactivos, de acuerdo con la propuesta de Jackson (2003).

Instrumento

Para la validación psicométrica se utilizó la versión traducida y adaptada de la Escala de Gravedad de Fatiga (FSS; Krupp et al., 1989), producto de un estudio previo (Duarte, Velasco, & Reyes-Lagunes, 2016), la cual se compone de nueve afirmaciones con siete opciones de respuesta de tipo Likert pictórico (Reyes-Lagunes, 1993), rotuladas *De acuerdo* y *En desacuerdo* en los extremos.

Para el análisis factorial confirmatorio y prueba de invarianza se empleó la versión validada de la Escala de Gravedad de Fatiga (FSS; Krupp et al., 1989) producto de la fase anterior, con nueve reactivos y un coeficiente alfa de Cronbach de .93.

También se utilizó la Escala de Impacto de Fatiga (FIA; Duarte-Ayala, Velasco-Rojano, Sánchez-Sosa, & Reyes-Lagunes, 2017), validada en población mexicana con un coeficiente alfa de Cronbach de .93, con el fin de comprobar validez convergente.

Procedimiento

En la validación psicométrica, los participantes fueron convocados a lugares públicos, como jardines, parques, plazas y mercados; se obtuvo su consentimiento, se les administró el instrumento, leyendo cada reactivo y las opciones de respuesta en los casos en que esto fuera necesario, como con participantes invidentes o que no supieran leer y escribir, y se les agradeció su participación. Una vez obtenidos los datos, se

Tabla 1

Análisis de Distribución de los Reactivos de la Escala de Gravedad de Fatiga FSS adaptada.

| Reactivo | Análisis de distribución de frecuencia de los Reactivos | | | | | | | | Prueba de Normalidad | | | |
|----------|---|-------|-------|-------|------------|-------|-------|------|----------------------|----------|----------------|--------------------------------|
| | Frecuencia en porcentaje por opción de respuesta | | | | | | | | Asimetría | Curtosis | Z | p con corrección de Lilliefors |
| | En Desacuerdo | | | | De acuerdo | | | | | | | |
| FSS1 | 14.44 | 11.67 | 13.33 | 14.44 | 17.22 | 17.22 | 11.67 | -.14 | -1.19 | .14 | $p < .01^{**}$ | |
| FSS2 | 20 | 21.67 | 13.89 | 17.78 | 15.56 | 7.22 | 3.89 | .35 | -.92 | .18 | $p < .01^{**}$ | |
| FSS3 | 21.11 | 17.22 | 15.56 | 20.56 | 10.56 | 8.89 | 6.11 | .36 | -.87 | .15 | $p < .01^{**}$ | |
| FSS4 | 15.56 | 17.22 | 21.67 | 16.11 | 13.33 | 8.33 | 7.78 | .35 | -.83 | .15 | $p < .01^{**}$ | |
| FSS5 | 21.67 | 20.56 | 16.67 | 17.78 | 12.78 | 5.56 | 5 | .47 | -.69 | .17 | $p < .01^{**}$ | |
| FSS6 | 19.44 | 19.44 | 17.22 | 16.67 | 13.89 | 8.33 | 5 | .36 | -.88 | .16 | $p < .01^{**}$ | |
| FSS7 | 21.67 | 20.56 | 16.11 | 12.78 | 11.11 | 11.67 | 6.11 | .42 | -1.03 | .18 | $p < .01^{**}$ | |
| FSS8 | 21.67 | 20.56 | 17.78 | 12.78 | 11.11 | 7.78 | 8.33 | .51 | -.86 | .17 | $p < .01^{**}$ | |
| FSS9 | 24.44 | 20.56 | 13.33 | 9.44 | 11.67 | 10.56 | 10 | .44 | -1.16 | .20 | $p < .01^{**}$ | |

Nota. $^{**}p \leq .01$

realizó la validación psicométrica culturalmente relevante propuesta por Reyes-Lagunes y García y Barragán (2008), que incluye análisis de distribución, discriminación, direccionalidad, confiabilidad y estructura factorial de los reactivos de la escala.

Para el análisis factorial confirmatorio y prueba de invarianza se convocó a los participantes como ha sido descrito anteriormente. Una vez obtenidos los datos se analizó cuál era la estructura factorial más adecuada y se realizó una prueba de invarianza por sexo siguiendo el procedimiento propuesto por Milsap y Olivera-Aguilar (2012), que consiste en evaluar cuatro modelos anidados, cada uno más restrictivo que el anterior. Primeramente, se probó el ajuste de la configuración teniendo en cuenta que la estructura factorial fuera la misma para los grupos. Posteriormente, se consideró la invarianza métrica o débil (en la que las cargas factoriales se restringen a igualdad), luego se evaluó el modelo con invarianza fuerte (en el que las cargas factoriales y los interceptos se restringen a igualdad), y finalmente un

modelo de invarianza estricta (en el que las cargas factoriales, los interceptos y las varianzas únicas de los reactivos se restringieron a igualdad) y se comparó el ajuste de los modelos. Si no se obtiene una diferencia significativa en chi cuadrada ni cambios mayores de .01 en el CFI, se considera que existe invarianza (Kline, 2016). Si se obtiene una invarianza estricta, los cambios resultan de la variable latente y no de un sesgo en la medición (DeShon, 2004).

Resultados

Validación Psicométrica

Se realizó un análisis de frecuencias de cada uno de los reactivos y se encontró que todas las opciones de respuesta tuvieron frecuencia. Posteriormente, se calcularon el sesgo, la curtosis y la prueba Z de Kolmogorov-Smirnov. Los resultados muestran que todos tuvieron una distribución sesgada, lo que fue considerado en análisis subsiguientes. La Tabla 1 incluye los porcentajes

Tabla 2
Análisis de Discriminación de los Reactivos de FSS versión adaptada.

| Reactivo | Comparación de grupos extremos | | | Correlación reactivo total | Correlación reactivo total sin el reactivo |
|----------|--------------------------------|-----------|----------------|----------------------------|--|
| | U Mann Whitney | <i>gl</i> | <i>p</i> | | |
| FSS1 | 85.00 | 45 | $p < .01^{**}$ | .68** | .59** |
| FSS2 | 169.00 | 45 | $p < .01^{**}$ | .70** | .62** |
| FSS3 | 12.00 | 45 | $p < .01^{**}$ | .83** | .78** |
| FSS4 | 10.00 | 45 | $p < .01^{**}$ | .84** | .79** |
| FSS5 | 17.50 | 45 | $p < .01^{**}$ | .85** | .80** |
| FSS6 | 35.50 | 45 | $p < .01^{**}$ | .84** | .80** |
| FSS7 | 0 | 45 | $p < .01^{**}$ | .87** | .83** |
| FSS8 | 10.50 | 45 | $p < .01^{**}$ | .84** | .79** |
| FSS9 | 1.00 | 45 | $p < .01^{**}$ | .85** | .79** |

Nota. ** $p \leq .01$.

por opción de respuesta, el coeficiente de asimetría, la curtosis y los resultados de la prueba de normalidad.

Posteriormente se estudió el poder discriminativo de los reactivos con dos análisis: una comparación de grupos extremos respecto al total de la prueba (puntajes inferiores al primer cuartil y puntajes superiores al tercer cuartil), mediante la prueba U de Mann Whitney, y una correlación entre el total de la prueba y cada uno de los reactivos. También se obtuvo/analizó una versión corregida de la misma en la cual el total de la prueba no incluía el reactivo con el que se iba a relacionar. Se eliminaron los reactivos con correlaciones bajas ($r < .30$) y todos los reactivos discriminaron significativamente como se muestra en la Tabla 2 (Magnusson, 1998).

Al analizar las correlaciones reactivo-total se verificó que en caso de eliminar cualquiera de los reactivos la consistencia de la escala disminuiría, debido a lo cual se consideró conveniente conservar la totalidad de los reactivos para mantener una buena consistencia. En la Tabla 2 se muestra

el valor de la comparación entre grupos extremos y las correlaciones entre reactivo y total.

A fin de conocer la estructura de la escala se realizó un análisis factorial exploratorio de ejes principales con rotación oblicua, adecuado para identificar las variables latentes que subyacen a los reactivos (Fabrigar, MacCallum, Wegener, & Strahan, 1999; Schmitt, 2011). Se eligió este método de extracción debido a que no se basa en el supuesto de normalidad, y se eligió el tipo de rotación debido a que este es más preciso y provee más información que la rotación ortogonal. Se encontró una solución unidimensional con los nueve reactivos que explica el 62% de la varianza.

En la Tabla 3 se muestran las cargas factoriales de cada uno de los reactivos, el autovalor y el porcentaje de varianza explicada.

Para verificar la pertinencia del análisis, se calculó la medida de adecuación muestral Kaiser-Meyer-Olkin ($KMO = .91$) y la prueba de esfericidad de Bartlett ($\chi^2_{(36)} = 1323.33, p < .01$) y se pudo concluir que la matriz de correlaciones de la que partía era adecuada. Posteriormente, se

calculó la confiabilidad total de la escala alfa de Cronbach y se obtuvo un valor adecuado de consistencia interna ($\alpha = .93$).

Tabla 3

Resultados del análisis factorial exploratorio de ejes principales con rotación oblicua del FSS adaptado.

| Reactivo | Factor 1. Fatiga |
|-------------------------|------------------|
| FSS7 | .87 |
| FSS5 | .84 |
| FSS6 | .84 |
| FSS8 | .83 |
| FSS9 | .83 |
| FSS4 | .82 |
| FSS3 | .81 |
| FSS2 | .64 |
| FSS1 | .60 |
| % Varianza Explicada | .62 |
| Autovalor o valor Eigen | 5.66 |

Análisis factorial confirmatorio y prueba de invarianza

A continuación se llevó a cabo un análisis factorial confirmatorio de máxima verosimilitud robusta, método seleccionado por la falta de normalidad en los datos (West, Taylor, & Wu, 2012)

y se probaron tres modelos comparativos, todos con una dimensión pero con diferente número de reactivos: uno con los nueve reactivos, otro con ocho y uno con siete reactivos. La identificación de los modelos se basó en establecer la carga factorial del primer reactivo como 1 para definir la métrica de la variable latente (Kenny & Milan, 2012) con el software R y el paquete Lavaan (Rosseel, 2012).

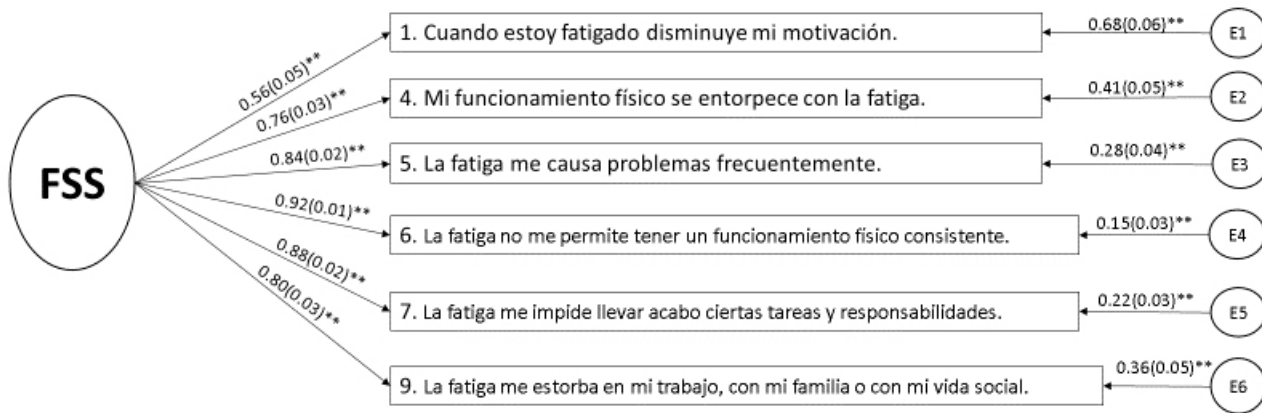
Para evaluar el ajuste de los modelos se utilizaron los siguientes índices de bondad de ajuste: Chi cuadrada normada, dividida por los grados de libertad (debería ser menor a 3.00 con una probabilidad no significativa; entre menor cociente mejor ajuste). Índice de bondad de ajuste Lisrel (GFI), índice de ajuste comparativo de Bentler (CFI; para que exista un buen ajuste los valores de GFI y CFI deben ser $\geq .95$), raíz cuadrada media del error de aproximación (RMSEA) y SRMR (los valores de RMSEA y SRMR deben ser $\leq .08$ para indicar un buen ajuste).

Los resultados de los tres modelos mostraron falta de ajuste global para explicar los datos, por lo que se analizaron las faltas de ajuste local mediante el uso de índices de modificación y se encontraron asociados con los reactivos 2, 3 y 8. A partir de estos resultados, se decidió modificar el modelo eliminando estos reactivos, ya que alcanzar ajuste mediante la inclusión de los mismos puede reflejar más la idiosincrasia de la muestra que la relación entre variables (MacCallum, Roznowski, & Necowitz, 1992). El modelo modi-

Tabla 4

Índices de ajuste para el análisis factorial confirmatorio de los modelos.

| Modelo | χ^2 | <i>p</i> | χ^2/gl | CFI | GFI | RMSEA | IC 90% | SRMR |
|-------------|-------------------------|----------------|-------------|-----|-----|-------|------------|------|
| 9 reactivos | $\chi^2_{(27)} = 85.39$ | $p < .01^{**}$ | 3.16 | .94 | .92 | .12 | [.09, .15] | .04 |
| 8 reactivos | $\chi^2_{(20)} = 64.01$ | $p < .01^{**}$ | 3.20 | .95 | .93 | .11 | [.08, .14] | .04 |
| 7 reactivos | $\chi^2_{(14)} = 34.39$ | $p < .01^{**}$ | 2.45 | .97 | .96 | .09 | [.05, .13] | .03 |
| 6 reactivos | $\chi^2_{(9)} = 12.56$ | $p = .18$ | 1.39 | .99 | .99 | .05 | [0, .11] | .02 |

**Figura 1**

Análisis factorial confirmatorio de la FSS.

Nota. Solución ajustada

ficado con una dimensión y seis reactivos mostró un adecuado ajuste en todos los índices, como se muestra en la Tabla 4, que incluye los índices de ajuste para todos los modelos.

Las cargas factoriales estandarizadas de la escala, que resultan significativas, positivas y con valores adecuados para la solución ajustada, se muestran en la Figura 1.

Para determinar la existencia de un funcionamiento diferencial de los reactivos de acuerdo con el parámetro sexo del individuo, se realizó una prueba de invarianza factorial de medición. Los resultados muestran una invarianza factorial estricta, ya que no se encontraron diferencias sig-

nificativas en el ajuste de los modelos anidados, tal como lo indica la Tabla 5.

La Tabla 5 incluye los datos de la prueba de diferencias de la χ^2 y cambio en CFI para los modelos anidados.

A continuación, se compararon ambos grupos en la media de la variable latente *fatiga* y no se encontraron diferencias estadísticamente significativas ($Z = -0.35, p = .72$).

Como evidencia de validez concurrente, se calculó la correlación con el puntaje total de la Escala de Impacto de Fatiga y se encontró una correlación media estadísticamente significativa ($r = .59, p < .001$).

Tabla 5

Prueba de diferencia de la χ^2 en los modelos de la dimensión *fatiga*.

| Modelo | χ^2 | gl | $\Delta\chi^2$ | Δgl | p | CFI | ΔCFI |
|---------------|----------|----|----------------|-------------|-----|------|--------------|
| Configuración | 25.36 | 18 | - | - | - | 0.99 | - |
| Métrica | 27.20 | 23 | 2.03 | 5 | .84 | 1 | .01 |
| Fuerte | 32.21 | 28 | 4.89 | 5 | .42 | 1 | 0 |
| Estricta | 39.37 | 34 | 5.10 | 6 | .53 | 1 | 0 |

Discusión

Uno de los objetivos del presente trabajo era validar psicométricamente de manera culturalmente relevante la Escala de Gravedad de la Fatiga en población general de la Ciudad de México. Para ello se buscó partir de una traducción adecuada al lenguaje que no sacrificara la evaluación de los constructos que la escala original propone. La versión final después de su aplicación y análisis cuenta con seis reactivos, un nivel de confiabilidad de $\alpha = .91$ y todos los reactivos se encuentran relacionados significativamente, lo que corrobora que se trata de la medición de un mismo constructo.

Mediante la validación psicométrica a través del análisis factorial exploratorio de ejes principales, se encontró que la estructura que emerge libremente de la población tiene una dimensión con nueve reactivos como en la estructura de la escala original (Krupp et al., 1989) con una adecuada consistencia interna. Sin embargo, en el análisis factorial confirmatorio se encontró que los reactivos 2 (*El ejercicio aumenta mi fatiga*), 3 (*Me fatigo fácilmente*) y 8 (*La fatiga es uno de los tres síntomas que menos me permiten hacer cosas*) no eran adecuados.

La falta de ajuste de los reactivos 2 y 3 que indican problemas en el funcionamiento físico puede explicarse por dos posibles razones, la primera es que son reactivos que surgieron en poblaciones afectadas por enfermedades crónicas, en particular en pacientes con esclerosis múltiple que entre sus principales síntomas tienen debilidad muscular en las extremidades y dificultad con la coordinación y el equilibrio (Instituto Nacional de Desórdenes Neurológicos y Accidentes Vasculares [NINDS], 2002), lo que no corresponde con el estado de salud de la población general, la segunda es que contestar afirmativamente a estos reactivos puede implicar una disminución de

la capacidad de trabajo, lo que en México no es socialmente aceptado por ser una cultura que valora mucho el trabajo (Díaz-Guerrero, 2003), esta razón también puede explicar que el reactivo 8 no funcione según se esperaba dado que se refiere a la capacidad de hacer cosas.

La extensión de la escala la convierte en una versión corta que incluye sólo los reactivos con las mejores propiedades psicométricas para la población meta. Haber obtenido una versión corta es una ventaja, ya que es más práctica en su aplicación. Es importante destacar que en el presente trabajo se obtuvo una nueva versión de la FSS, por lo que los resultados obtenidos no son comparables con los resultados obtenidos con otras versiones de la misma escala, ya que al cambiar el número de reactivos, la definición del constructo cambia y se circunscribe a los resultados obtenidos en población general de la Ciudad de México.

Una evidencia de validez concurrente es la correlación media con la Escala de Impacto de Fatiga, la cual es importante porque se trata de una escala que mide fatiga en la misma población meta con criterios rigurosos de confiabilidad y validez, además de que ha sido utilizada en otras poblaciones para probar este tipo de validez (Oncu, Atamaz, Durmaz, & On, 2013).

En cuanto al objetivo de comparar por sexo, se observó que existe invarianza factorial estricta entre hombres y mujeres, lo que significa que no existe sesgo que favorezca a ningún grupo. También se encontró que no hay diferencias significativas en el intercepto de la variable latente, lo que sugiere que la fatiga se experimenta independientemente de la variable sexo.

Limitaciones

Es importante resaltar que en este estudio no se utilizaron pruebas físicas para medir la fatiga y

se consideraron sólo medidas de autoinforme, por lo que sería conveniente hacer estudios que las incluyeran, ya que a pesar de que el tratamiento estadístico riguroso de los datos ayuda a controlar el error en las mediciones, se ha observado que en el caso de autoinformes de salud, las personas pueden exagerar o subestimar los síntomas por tener percepciones alteradas debido a su propio estado (Cohen, 1979).

También es necesario mencionar que la información sobre validez presentada en este estudio es pertinente exclusivamente a la medición de la fatiga en población general de la Ciudad de México. Nuevos estudios de validez serían necesarios para conocer las propiedades psicométricas de la escala obtenida en este estudio en otros contextos, con la finalidad de poder utilizarla en los mismos para fines de investigación o diagnóstico. Aún ante la posibilidad de trabajar con población clínica dentro de la Ciudad de México, serían necesarios nuevos estudios de validez; considerando que la estructura de los reactivos podría ser diferente, y que en la presente versión se han eliminado algunos reactivos.

Se debe destacar que la prueba de invarianza demuestra que la escala es una medida insesgada por sexo, y por consiguiente es adecuado usarla para comparar la fatiga entre hombres y mujeres; sin embargo, en caso de utilizarse para otras variables atributivas, debería establecerse primero que funcione de la misma manera para todos los grupos.

Conclusión

Los resultados encontrados muestran que la FSS tiene propiedades psicométricas que la hacen una medida confiable y válida para el propósito de evaluar fatiga en población general de la ciudad de México y puede ser utilizada en investiga-

ciones o evaluaciones que contribuyan a conocer la prevalencia y las consecuencias de la fatiga en la población general.

También puede contribuir en el diagnóstico, lo que permitiría a los profesionales de la salud atender este problema desde sus etapas tempranas y contribuir al desarrollo de programas de prevención de la fatiga.

Referencias

- Ali, M. R., Khaleque, A., Khanam, M., Al-Shatti, A., & Ahmed, R. U. (1994). Critical flicker frequency of mentally retarded and normal persons. *Perceptual and Motor Skills*, 79(3), 1235-1238. doi: 10.2466/pms.1994.79.3.1235
- Armutlu, K., Keser, İ., Korkmaz, N., Akbıyık, D. İ., Sümbüloğlu, V., Güney, Z., & Karabudak, R. (2007). Psychometric study of Turkish version of Fatigue Impact Scale in multiple sclerosis patients. *Journal of the Neurological Sciences*, 255(1-2), 64-68. doi: 10.1016/j.jns.2007.01.073
- Bernal-Vargas, L., Riveros-Munévar, F., Vinaccia-Alpi, S., & Quiceno-Sierra, J. M. (2017). Estructura factorial y consistencia interna de la Escala de Severidad de Fatiga en población colombiana con enfermedades crónicas. *Enfermería Global*, 16(2), 37. doi: 10.6018/eglobal.16.2.255821
- Beurskens, A. J., Bültmann, U., Kant, I., Vercoulen, J. H., Bleijenberg, G., & Swaen, G. M. (2000). Fatigue among working people: Validity of a questionnaire measure. *Occupational and Environmental Medicine*, 57(5), 353-357. Recuperado de <https://www.ncbi.nlm.nih.gov/pubmed/10769302>
- Chalder, T., Berelowitz, G., Pawlikowska, T., Watts, L., Wessely, S., Wright, D., & Wallace, E. P. (1993). Development of a fatigue scale. *Journal of Psychosomatic Research*, 37(2), 147-53. Recuperado de <https://www.ncbi.nlm.nih.gov/pubmed/8463991>
- Cohen, F. (1979). Personality, stress, and the develop-

- ment of physical illness. En G. C. Stone, F. Cohen, N. E. Adler & Associates. *Health Psychology - A Handbook. Theories, applications, and challenges of a psychological approach to the healthcare system.* (pp. 1-32). San Francisco: JosseyBass.
- De Croon, E. M. (2006). Psychometric properties of the Need for Recovery After Work Scale: Test-retest reliability and sensitivity to detect change. *Occupational and Environmental Medicine*, 63(3), 202-206. doi: 10.1136/oem.2004.018275
- DeShon, R. P. (2004). Measures are not invariant across groups without error variance homogeneity. *Psychology Science*, 46(1), 137-149. Recuperado de http://www.pabst-publishers.de/psychology-science/1-2004/abstract_08.html
- Díaz-Guerrero, R. (2003). *Psicología del Mexicano* (6ª reimp.). México: Trillas.
- Duarte-Ayala, R. E. (2017). *Fatiga, estilos de enfrentamiento y calidad de vida en médicos residentes* (Tesis doctoral). Universidad Nacional Autónoma de México, Facultad de Psicología, México. Recuperado de <http://132.248.9.195/ptd2017/junio/513011441/Index.html>
- Duarte-Ayala, R. E., Velasco-Rojano, Á. E., Sánchez-Sosa, J. J., & Reyes-Lagunes, L. I. (2017). Adaptación y validación de la Escala de Impacto de Fatiga. *Acta de Investigación Psicológica*, 7(1), 2585-2592. doi: 10.1016/j.aiiprr.2017.02.002
- Duarte, R. E., Velasco, A. E., & Reyes-Lagunes, L. I. (2016). Validación psicométrica de la Escala de Gravedad de Fatiga. En C. Gómez, J. E. Hernández-Sánchez, R. Díaz-Loving, L. I. Reyes-Lagunes & S. Rivera (Eds.), *Aportaciones Actuales de la Psicología Social. Volumen III.* (pp. 103-111) Tabasco, México: Asociación Mexicana de Psicología Social AMEPSO. Universidad Autónoma Juárez de Tabasco.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272-299. doi: 10.1037/1082-989x.4.3.272
- Fuentes-Márquez, S., Senín-Calderón, C., Rodríguez-Testal, J. F., & Carrasco, M. A. (2015). Perceived experience of fatigue in clinical and general population: Descriptors and associated reactivities. *Spanish Journal of Psychology*, 18(9), 1-8. doi: 10.1017/sjp.2015.11
- Gök, K., Cengiz, G., Erol, K., & Özgöçmen, S. (2016). The Turkish Version of Multidimensional Assessment of Fatigue and Fatigue Severity Scale is reproducible and correlated with other outcome measures in patients with systemic sclerosis. *Archives of Rheumatology*, 31(4), 329-332. doi: 10.5606/ArchRheumatol.2016.5909
- Gottschalk, M., Kümpfel, T., Flachenecker, P., Uhr, M., Trenkwalder, C., Holsboer, F., & Weber, F. (2005). Fatigue and regulation of the hypothalamo-pituitary-adrenal axis in multiple sclerosis. *Archives of Neurology*, 62(2), 277-280. doi: 10.1001/archneur.62.2.277
- Instituto Nacional de Desórdenes Neurológicos y Accidentes Vasculares [NINDS] (2002). *Esclerosis Múltiple: Esperanza en la Investigación.* Bethesda, Maryland: Departamento de Salud y Servicios Humanos de los Estados Unidos. Recuperado de https://espanol.ninds.nih.gov/trastornos/span_esclerosis.pdf
- Jackson, D. L. (2003). Revisiting sample size and number of parameter estimates: Some support for the N:q hypothesis. *Structural Equation Modeling*, 10(1), 128-141. doi: 10.1207/S15328007SEM1001_6
- Johansson, S., Kottorp, A., Lee, K. A., Gay, C. L., & Lerdal, A. (2014). Can the Fatigue Severity Scale 7-item version be used across different patient populations as a generic fatigue measure –A comparative study using a Rasch model approach. *Health Quality of Life Outcomes*, 12(1), 24. doi: 10.1186/1477-7525-12-24
- Kenny, D. A., & Milan, S. (2012). Identification: A nontechnical discussion of a technical issue. En R. H. Hoyle (Ed.), *Handbook of Structural Equation Modeling* (pp. 145-163). New York: Guilford Press.
- Kos, D., Kerckhofs, E., Carrea, I., Verza, R., Ramos, M., & Jansa, J. (2005). Evaluation of the Modified

- Fatigue Impact Scale in four different European countries. *Multiple Sclerosis*, 11(1), 76-80. doi: 10.1191/1352458505ms1117oa
- Kline, R. (2016). *Principles and practice of structural equation modeling* (4^a ed.). New York, NY: The Guilford Press.
- Krupp, L. B., LaRocca, N. G., Muir-Nash, J., & Steinberg, D. (1989). The Fatigue Severity Scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. *Archives of Neurology*, 46(10), 1121-1123. doi: 10.1001/archneur.1989.00520460115022
- Lerdal, A., & Kottorp, A. (2011). Psychometric properties of the Fatigue Severity Scale—Rasch analyses of individual responses in a Norwegian stroke cohort. *International Journal of Nursing Studies*, 48(10), 1258-1265. doi: 10.1016/j.ijnurstu.2011.02.019
- Lerdal, A., Kottorp, A., Gay, C., Aouizerat, B. E., Portillo, C. J., & Lee, K. A. (2011). A 7-item version of the Fatigue Severity Scale has better psychometric properties among HIV-infected adults: An application of a Rasch model. *Quality of Life Research*, 20(9), 1447-1456. doi: 10.1007/s11136-011-9877-8
- MacCallum, R. C., & Hong, S. (1997). Power analysis in covariance structure modeling using GFI and AGFI. *Multivariate Behavioral Research*, 32(2), 193-210. doi: 10.1207/s15327906mbr3202_5
- MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin*, 111(3), 490-504. Recuperado de <http://www.apa.org/pubs/journals/bul>
- Magnusson, D. (1998). *Teoría de los tests*. México: Trillas.
- Malta-Purim, K. S., Bittencourt-Gumarães, A. T., Kapp-Titski, A. C., & Leite, N. (2016). Sleep deprivation and drowsiness of medical residents and medical students. *A Revista do Colégio de Cirurgiões*, 43(6), 438-444. doi: 10.1590/0100-69912016006005
- Millsap, R. E., & Olivera-Aguilar, M. (2012). Investigating measurement invariance using confirmatory factor analysis. En R. H. Hoyle (Ed.), *Handbook of Structural Equation Modeling* (pp. 380-392). New York, NY: Guilford Press.
- Montero, I., & León, O. G. (2007). A guide for naming research studies in Psychology. *International Journal of Clinical and Health Psychology*, 7(3), 847-862.
- Neisa, C., Tangarife, A., & Godoy, L. (2014). Variables moderadoras de la relación fatiga laboral y accidentes de trabajo en personal operativo. En J. Orejuela (Ed.), *Psicología de las Organizaciones y del Trabajo* (pp. 429-446). Cali: Editorial Bonaventuriana.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Teoría psicométrica*. McGraw-Hill: New York.
- Oncu, J., Atamaz, F., Durmaz, B., & On, A. (2013). Psychometric properties of Fatigue Severity and Fatigue Impact Scales in postpolio patients. *International Journal of Rehabilitation Research*, 36(4), 339-345. doi: 10.1097/mrr.0b013e3283646b56
- Orrantia-Daniel, M. C. G., & De la Vega-Bustillos, E. (2003). Trabajo, fatiga, calidad y productividad. *IV Congreso Nacional de Estudios del Trabajo*. Colegio de Sonora. Hermosillo, Sonora. Recuperado de <http://www.izt.uam.mx/sotraem/Documentos/Documentos/Amet2003/templates/res/tema%209/GILBERTO%20ORRANTIA-ENRIQUE%20DE%20LA%20VEGA.pdf>
- Ottoneo, M., Pellicciari, L., Giordano, A., & Foti, C. (2016). Rasch analysis of the Fatigue Severity Scale in Italian subjects with multiple sclerosis. *Journal of Rehabilitation Medicine*, 48(7), 597-603. doi: 10.2340/16501977-2116
- Reyes-Lagunes, L. I. (1993). Las redes semánticas naturales, su conceptualización y su utilización en la construcción de instrumentos. *Revista de Psicología Social y Personalidad*, 9(1), 81-97.
- Reyes-Lagunes, L. I., & García y Barragán, L. F. (2008). Procedimiento de validación psicométrica culturalmente relevante: Un ejemplo. En S. Rivera-Aragón, R. Díaz-Loving, R. Sánchez-Aragón & L. I. Reyes-Lagunes (Eds.), *La Psicología Social en México. Vol. XII* (pp. 625-636). México: Asociación Mexicana de Psicología Social.
- Rosseel, Y. (2012). *Lavaan: An R package for structural*

- equation modeling. *Journal of Statistical Software*, 48(2), 1-36. doi: 10.18637/jss.v048.i02
- Schmitt, T. A. (2011). Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment*, 29(4), 304-321. doi: 10.1177/0734282911406653
- Torres-Harding, S. R., Jason, L. A., & Taylor, R. R. (2002). Fatigue severity, attributions, medical utilization, and symptoms in persons with chronic fatigue. *Journal of Behavioral Medicine*, 25(2), 99-113. Recuperado de <https://link.springer.com/article/10.1023/A:1014850819995>
- Torres-Harding, S. R., Mason-Shutter, J., & Jason, L. A. (2008). Fatigue Among Spanish- and English-Speaking Latinos. *Social Work in Public Health*, 23(5), 55-72. doi: 10.1080/19371910802053232
- Useche, S., Cendales, B., & Gómez, V. (2017). Measuring fatigue and its associations with job stress, health and traffic accidents in professional drivers: The case of BRT operators. *EC Neurology*, 4(4), 103-118.
- Van't Leven, M., Zielhuis, G.A., Van der Meer, J. W., Verbeek, A. L., & Bleijenberg, G. (2009). Fatigue and chronic fatigue syndrome-like complaints in the general population. *European Journal of Public Health*, 20(3), 251-257. doi: 10.1093/eurpub/ckp113
- Viner, R., & Christie, D. (2005). Fatigue and somatic symptoms. *British Medical Journal*, 330, 1012-1015. doi: 10.1136/bmj.330.7498.1012
- West, S. G., Taylor, A. B., & Wu, W. (2012). Model fit and model selection in structural equation modeling. En R. H. Hoyle (Ed.), *Handbook of Structural Equation Modeling* (pp. 209-231). New York: Guilford Press.
- Whitehead, L. (2009). The measurement of fatigue in chronic illness: A systematic review of unidimensional and multidimensional fatigue measures. *Journal of Pain and Symptom Management*, 37(1), 107-128. doi: 10.1016/j.jpainsymman.2007.08.019
- Yoshitake, H. (1971). Three characteristic patterns of subjective fatigue symptoms. *Ergonomics*, 21(3), 231-233. doi: 10.1080/00140137108931236