

ANÁLISIS ESTRUCTURAL E INVARIANZA DE MEDICIÓN DEL MBI-GS EN TRABAJADORES PERUANOS

STRUCTURAL ANALYSIS AND MEASUREMENT INVARIANCE OF MBI-GS IN PERUVIAN WORKERS

Manuel Fernández-Arata*, Arturo Juárez García** y César Merino Soto***

Universidad de San Martín de Porres, Perú
Universidad Autónoma del Estado de Morelos, México

Recibido: 16 de febrero de 2015

Aceptado: 29 de abril de 2015

RESUMEN

La medición del burnout ha evolucionado con la creación de varios instrumentos y modelos. El Maslach Burnout Inventory - General Survey (MBI-GS) es uno de estos instrumentos para medir tres constructos definicionales del burnout: (1) agotamiento emocional, (2) eficacia profesional y (3) indiferencia. Fue creado para un amplio rango de ocupaciones, pero pocas veces se ha verificado su estructura latente e invarianza de medición en Latinoamérica. El presente estudio analiza esta estructura latente y la invarianza de medición del MBI-GS en una muestra de 940 trabajadores peruanos de varias ocupaciones. Se aplicó la metodología de ecuaciones estructurales mediante el análisis factorial confirmatorio, así como la invarianza de medición entre varones y mujeres, imponiendo restricciones sucesivamente más estrictas. Los resultados verificaron satisfactoriamente la estructura de tres dimensiones latentes del MBI-GS, y la invarianza de sus parámetros entre hombres y mujeres. Se discute las implicaciones de los resultados.

Palabras clave: Burnout, análisis factorial confirmatorio, invarianza de medición, trabajadores, validez.

ABSTRACT

The measurement of burnout has evolved into the creation of various tools and models. The Maslach Burnout Inventory - General Survey (MBI-GS) is one of these instruments used to measure three definitional constructs of burnout: (1) emotional exhaustion, (2) professional efficiency, and (3) indifference. It was created for a wide range of occupations, but its latent structure and invariance of measurement in Latin America has rarely been verified. The present study analyzes the latent structure and the invariance of measurement of MBI-GS in a sample of 940 Peruvian workers in various occupations. The methodology of structural equations was applied through the confirmatory factor analysis, as well as the invariance of measurement between men and women, imposing restrictions successively more strict. The results satisfactorily verified the structure of three-dimensional latent MBI-GS, and the invariance of its parameters between men and women. The implications of the results are discussed.

Keywords: Burnout, confirmatory factor analysis, invariance of measurement, workers, validity

* mfernandezal@usmp.pe

** arturojuarezg@hotmail.com

*** sikayax@yahoo.com.ar

Una de las escalas para evaluar el síndrome de burnout, en un amplio rango de ocupaciones, es el Maslach Burnout Inventory (MBI) en su versión MBI-General Survey, MBI-GS (Maslach, Jackson y Leiter, 1996), cuya diferenciación respecto a las anteriores versiones ocurre en el ajuste de varios aspectos de los constructos medidos por el MBI-GS que son: agotamiento emocional (AE), indiferencia (IN) y eficacia profesional (EP), principalmente al considerar que la aparición del síndrome se debe a una crisis del trabajador con el propio trabajo y no necesariamente en su relación con las personas (Maslach et al., 1996; Oramas, Gonzáles y Vergara, 2007). En ese sentido, los ítems de AE son más genéricos que los de las versiones anteriores e incluyen fatiga física y emocional, sin hacer referencia a la interacción con personas. También, los ítems de IN abordan las actitudes de indiferencia hacia el trabajo y los roles implicados. Por último, los ítems de EP recogen aspectos sociales y no sociales del trabajo, enfocándose más en las expectativas de éxito (Gil-Monte, 2002).

En las relaciones entre los constructos medidos por el MBI-GS, AE e IN son los constructos que se valoran como esenciales en la definición del burnout y ambos generalmente muestran correlaciones comparativamente más elevadas que con EP. Esto ha sido replicado en numerosos estudios (e.g. Bakker, Demerouti y Schaufeli, 2002; Chirkowska-Smolak y Kleka, 2011; Gil-Monte, 2002; Oramas et al., 2007; Schutte, Toppinnen, Kalimo y Schaufeli, 2000), lo que respalda la interdependencia de ambos en la expresión conductual del burnout, aspectos que también son esenciales para determinar la gravedad de los síntomas (Green, Walkey y Taylor, 1991). Pero sus altas correlaciones también pueden indicar que la validez discriminativa es inferior a la varianza común de ambos, por lo que no se podría garantizar una adecuada diferenciación comportamental que sea útil para la investigación y la práctica profesional (Farrell, 2009). La validez discriminativa es un aspecto de la validez de constructo que permite diferenciar de manera conceptual y práctica los constructos (Fornell y Larcker, 1981; Hair, Anderson, Tatham y Black, 2006), y que generalmente se incluye en los estudios de validez interna de un instrumento. Algunos estudios han encontrado incluso dos fuertes factores, uno integrado por AE e IN y el otro por EP (Grajales, 2001), pero esta estructura es menos

frecuente que la de tres factores, por lo que parece haber un suficiente respaldo para tomar la presunción que la estructura del MBI-GS en otros contextos culturales también puede interpretarse del mismo modo (tres factores). Debido a que el MBI-GS es relativamente nuevo en Latinoamérica comparado con su uso en países angloparlantes, se requiere explorar psicométricamente sus propiedades.

En Latinoamérica, y hasta la fecha del presente manuscrito, solo dos estudios (Fernández, Merino y Guimet, en prensa; Flores, Merino, Camacho, Juárez y Placencia, en prensa) se han efectuado para verificar la estructura latente del MBI-GS mediante la metodología del modelado de ecuaciones estructurales (Structural Equations Modeling - SEM), específicamente usando el análisis factorial confirmatorio (Confirmatory Factor Analysis - CFA). Este procedimiento permite evaluar las relaciones estructurales internas de un instrumento desde una base hipotética de las mismas, obteniendo parámetros interpretables para la definición del constructo de interés; por lo tanto, es un procedimiento altamente recomendado cuando se quiere probar si un modelo de medición hipotetizado es una razonable figura de cómo se comportan varios parámetros de los ítems y sus variables latentes (Bentler y Dugeon, 1996; Hair et al., 2006). En ambos estudios (Fernández et al., en prensa; Flores et al., en prensa) se replicó satisfactoriamente el modelo de tres dimensiones, pero requirieron de errores correlacionados en algunos ítems y la remoción del ítem 13; no se examinaron sus coeficientes estructurales, y la confiabilidad por consistencia interna para los puntajes fue aceptable (mayor a 0.70). Debe señalarse que ambos estudios se realizaron con un tamaño muestral (menos de 200) que pone un serio límite a la generalización confiable de sus resultados, y la viabilidad de verificar la invarianza de medición; adicionalmente, fueron grupos ocupacionalmente homogéneos que limitan la variabilidad de los resultados. En el resto de la literatura sobre la validez estructural del MBI-GS en Latinoamérica, no hay más evidencias psicométricas sobre la dimensionalidad del MBI-GS en muestras no hispanas usando el CFA. Por otro lado, con muestras hispanas no latinoamericanas que han aplicado la metodología SEM (e.g. Bakker et al., 2002; Chirkowska-Smolak y Kleka, 2011; Leiter y Schaufeli, 1996), se ha reportado tres aspectos que se resumen del siguiente

* mfernandezal@usmp.pe

** arturojuarezg@hotmail.com

*** sikayax@yahoo.com.ar

modo: (a) la estructura latente del MBI-GS generalmente es modelada de manera satisfactoria por tres constructos correspondientes a AE, EP e IN; (b) AE, IN y EP están correlacionados en una dirección y magnitud teóricamente consistentes; (c) pueden aparecer inconsistentes covariaciones entre los términos de error de los ítems dentro del mismo constructo; y (d) los ítems representan bien sus relaciones con sus respectivos constructos. En este último aspecto, sin embargo, el ítem 13 generalmente muestra pobre peso factorial respecto a su constructo (IN), y es un problema que las investigaciones han solucionado removiéndolo de los análisis principales (Bakker et al., 2002; Chirkowska-Smolak y Kleka, 2011; Fernández et al., en prensa; Gil-Monte, 2002; Merino y Fernández, 2014; Oramas et al., 2007; Salanova y Schaufeli, 2000; Schutte et al., 2000).

En el Perú, las investigaciones conducidas hasta la fecha con el MBI han obtenido resultados únicamente con las versiones anteriores del MBI-GS, tales como la MBI-HS (Human Services Survey) (Aguilar y Gutiérrez, 2007; Gomero, Palomino, Ruiz y Llap, 2005; Mariños, Otero, Tomateo y Málaga, 2011) y MBI-ES (Educators Survey) (Fernández, 2002, 2008; Delgado, 2003), y han sido hechas con un enfoque exploratorio aplicando modelos factoriales no apropiados para la medición de constructos, por ejemplo, análisis de componentes principales. La ausencia de evaluaciones psicométricas del MBI-GS con metodología SEM en el Perú y en Latinoamérica, crea una cuestión abierta de estas versiones respecto a su validez interna, la confiabilidad de los puntajes y la invarianza de medición entre los grupos ocupacionales (Merino y Fernández, 2014). Considerando las ventajas del MBI-GS –aplicable a múltiples ocupaciones, dimensionalidad replicable, consistencia interna adecuada, validada en varios idiomas, y que no hay estudios psicométricos robustos desde una perspectiva focalizada en análisis SEM– el presente estudio tiene por objetivos estudiar la estructura interna del MBI-GS y su invarianza de medición (IM) con relación al sexo. Este procedimiento verifica que las propiedades estructurales se mantienen invariantes entre los grupos, lo que representa un argumento para hacer comparaciones de grupo en las estimaciones latentes de medias y covarianzas (Brown, 2006; Byrne, 2008); por ejemplo, la invarianza configuracional (estructura), invarianza débil (métrica) son prerrequisitos para muchos

análisis de comparación de grupos (Byrne, 2008; Steinmetz, 2013). El atributo del género servirá como agrupación para probar la IM, pues la experiencia psicológica en varones y mujeres puede interaccionar de diferente modo de acuerdo al contexto (Hyde, 2005), y no se ha reportado en el presente contexto de investigación alguna característica de IM con alguna de las versiones del MBI-GS.

Método

Participantes

Integraron la muestra 940 participantes entre docentes, serenos y profesionales. Los docentes sumaron un total de 741 (78.8 %), de Educación Primaria 64 % y de Educación Secundaria 36 %, provenientes de escuelas estatales y privadas de la ciudad de Lima Metropolitana. Serenos, 142 (15.1 %) que laboraban como vigilantes en dos municipalidades de la ciudad de Lima, la mayoría con grado de instrucción Secundaria. Profesionales (administrador, psicólogo, abogado, economista), 57 (6.1 %) del área de recursos humanos de empresas de Lima. En cuanto al género, 37.6 % varones y 62.4 % mujeres. Laboraban en entidades estatales el 51.2 % y el 48.8 % en entidades privadas. Los grupos de edad tuvieron la siguiente distribución: el 19.8 % entre veinticinco y veintinueve años; el 19.8 % entre treinta y treinta y cuatro años; el 18.3 % entre treinta y cinco y cuarenta años; el 34.8 % más de cuarenta años y el 7 % no consignaron este dato.

Instrumento

- *Maslach Burnout Inventory General Survey* - MBI-GS (Maslach et al., 1996).

Esta versión contiene 16 ítems clasificados de la siguiente manera: Agotamiento emocional (cinco ítems), Indiferencia (cinco ítems) y Eficacia profesional (seis ítems). Los ítems se valoran en una escala tipo Likert que refleja la frecuencia con la que experimentan la situación que se describe en el ítem y tiene siete grados que oscilan entre 0 (Nunca) y 6 (Todos los días). Se usó la versión en español, traducida para América Latina (Juárez et al., 2011). Como en otros estudios estructurales respecto al ítem 13 (del puntaje Indiferencia), este fue retirado en los análisis porque en la literatura peruana (Fernández et al., en prensa; Merino y Fernández, 2014) e internacional (Bakker et al.,

* mfernandez1@usmp.pe

** arturojuarezg@hotmail.com

*** sikayax@yahoo.com.ar

2002; Chirkowska-Smolak y Kleka, 2011; Gil-Monte, 2002; Oramas et al., 2007; Salanova y Schaufeli, 2000; Schutte et al., 2000), no demostró buenas propiedades correlacionales apropiadas en su constructo. Esta decisión también fue respaldada por las correlaciones ítem-test preliminares de este ítem obtenidas en la presente muestra, las que mostraron estar cerca de cero en la muestra total y en las submuestras relevantes (por ejemplo, sexo).

Procedimiento

Se aplicó la metodología de ecuaciones estructurales mediante el análisis factorial confirmatorio (Jöreskog, 1969) para verificar varios modelos sobre los parámetros de medición del MBI-GS. El método utilizado fue el de máxima verosimilitud con el escalamiento de Satorra y Bentler, 1994 (SB- χ^2), un procedimiento que demuestra ser efectivo bajo diferentes intensidades de distribución no normal de los ítems (Boomsma, 2000; Lei y Wu, 2012; Tong y Bentler, 2013), y que reduce la discrepancia entre el efecto de las distribuciones de las variables y la distribución de la función de ajuste aplicada (Bentler y Dugeon, 1996). Por otro lado, el número de opciones de respuesta de los ítems del MBI-GS fue siete, que se consideró como una característica suficiente para su aproximación a variables de tipo continuas (Remthulla, Brosseau-Liard y Savalei, 2012), con lo que se disminuiría la distinción estadística entre variables categóricas y continuas y no se restringiría seriamente la varianza de los ítems. Se usó la matriz de covarianzas S como unidad de análisis (expuesta en la Tabla 1), estimada mediante el programa EQS 6.2 (Bentler y Wu, 2012); este programa también fue utilizado para todos los análisis SEM.

El procedimiento CFA se inició imponiendo restricciones estructurales que son habituales (MacCallum y Austin, 2000), como la covariación cero entre los términos de error de los ítems y de los factores, y que cada ítem es influenciada por una única variable latente. El primer indicador de cada factor fue fijado con 1.0; como es probable que algunas de estas restricciones requieran ser relajadas durante el análisis (Boomsma, 2000), se hicieron reespecificaciones que mejoren el ajuste dentro de un marco de modelamiento a posteriori. Estas se hicieron primero con base estadística, verificando los índices de Lagrange (Sörbom, 1989) y segundo,

examinando la racionalidad de los mismos (Boomsma, 2000; Lei y Wu, 2012). Los índices de ajuste que se usaron para evaluar descriptivamente los modelos probados (ver Resultados) fueron el *Comparative Fit Index*, CFI³ 0.95 (Bentler, 1990), *Standardized Root Mean Square Residual*, SRMR ≤ 0.08 (Jöreskog y Sörbom, 1986) y *Root Mean Square Error of Approximation*, RMSEA ≤ 0.05 (Steiger y Lind, 1980), con sus intervalos de confianza al 90 % (McDonald, 1989). Esta combinación de índices de ajuste parece estar mejor respaldada por la literatura metodológica empírica (Jackson, Gillaspay y Purc-Stephenson, 2009) para tomar decisiones sobre los modelos evaluados. Por otro lado, la estimación de la confiabilidad se hizo mediante el coeficiente α (Cronbach, 1951) y sus intervalos de confianza (Romano, Kromrey, Owens y Scott, 2011), y el coeficiente ω (McDonald, 1999), usando los resultados factoriales del modelamiento CFA.

La validez convergente se refiere al monto promedio de varianza extraída (*Average Variance Extracted* - AVE), el mismo que se satisface si es > 0.50 (Fornell y Larcker, 1981); sin embargo, este criterio puede considerarse muy conservador y se puede aceptar valores levemente menores (Hair et al., 2006), especialmente en un contexto en que los constructos muestran acumulada evidencia empírica y respaldo teórico respecto a la moderada o elevada varianza común, como usualmente es el caso entre AE e IN. La validez discriminativa se evaluó mediante la comparación de la raíz cuadrada de AVE con la correlación interfactorial de los otros constructos (ϕ), mediante la cual se satisface este tipo de evidencia si $\sqrt{AVE} > \phi$ (correlación interfactorial).

Finalmente, para verificar la invarianza de medición en la comparación de varones y mujeres, se procedió secuencialmente, desde el modelo con menos restricciones hasta el modelo con más restricciones, un enfoque que es habitualmente conducido en la literatura metodológica (Brown, 2006). Las restricciones aplicadas a las pruebas de invarianza de los ítems serán respecto al número de variables latentes (invarianza configuracional), a las cargas factoriales λ_i (invarianza métrica), los interceptos τ_i (invarianza escalar) y los términos de error ϵ_i (invarianza residual). Esencialmente se eligió al sexo como criterio de diferenciación de los grupos porque es la variable que mostraba una cantidad balanceada de sujetos, pero

* mfernandezal@usmp.pe

** arturojuarezg@hotmail.com

*** sikayax@yahoo.com.ar

además por su efecto en varios constructos relacionados con el estrés (Hyde, 2005; Kajantie y Phillips, 2006).

Resultados

Análisis preliminar

En la Tabla 1 se muestra las características estadísticas de los ítems. Se observa que la intensidad de respuesta es más elevada en los ítems de EP e IN, pues están más cerca de las opciones extremas de respuesta. Esta característica impactó también en los elevados valores de asimetría y curtosis de estos mismos ítems. Los ítems de AE, por el contrario, muestran relativamente menos asimetría y excesos de curtosis. Se observa que el ítem 13 tiene menor covariación que el resto de los ítems y la más alta dispersión de respuestas; adicionalmente, su asimetría y curtosis tienen características diferentes comparadas con los ítems de su misma escala.

Modelamiento

Modelo de medición. Para probar los modelos se planteó primero una estrategia en dos fases: especificación a priori (identificar el modelo de acuerdo a hipótesis) y a posteriori (en que potencialmente se hacen nuevas especificaciones que mejorarían el ajuste de los modelos, basados en los problemas de ajuste hallados). El único modelo a priori (M_a) fue el de tres variables latentes oblicuas (AE, EP e IN), interpretado como la diferenciación conceptual entre AE, EP e IN, el mismo que tiene más respaldo en la literatura como el que logra un ajuste satisfactorio. No se probarían más modelos (por ejemplo, el de factores ortogonales o unidimensional) debido a que estos no tienen sustancial respaldo en la literatura y son menos justificables racionalmente. El análisis anterior se realizó para toda la muestra y para los subgrupos generados por el sexo, pues esto serviría como los modelos básicos para verificar después la invarianza de medición. Los resultados del ajuste (Tabla 2) en la muestra total indican que el modelo de tres dimensiones (M_a) se ajustó solo moderadamente a los datos, pues existió discrepancia entre CFI y los índices RMSEA y SRMR. Examinando los índices de modificación se halló una reespecificación que produciría un sustancial cambio en el $SB-\chi^2$ del modelo evaluado ($D = 171.05$, $p_{Hancock} < 0.001$), consistente en establecer la correlación entre los términos

de error de los ítems 14 y 15, generando el modelo M_b . Conceptualmente estos ítems reflejan una actitud de indiferencia y distanciamiento hacia el propio trabajo así como al valor y significado que se le puede conceder; representan igualmente una actitud defensiva hacia las agotadoras demandas de la actividad laboral (Gil-Monte, 2002; Maslach et al., 1996; Moreno, Rodríguez y Escobar, 2001; Oramas et al., 2007). La modificación, realizada en un marco a posteriori, produjo una correlación moderadamente alta entre estos ítems ($r = 0.429$, $Cov = .739$, $Z = 8.598$, $p < 0.001$) y un ajuste superior comparado con el M_a (Tabla 3).

Invarianza de medición. En el análisis anterior se utilizó la muestra total, pero para iniciar el estudio de la invarianza de medición, se procedió a buscar el modelo de línea base entre los grupos que se compararon (Milfont y Fischer, 2010). Los resultados del modelamiento en varones y mujeres (Tabla 2) también verificaron que, para ambos grupos, el modelo M_a , requería una reespecificación importante correspondiente también a la covariación entre los ítems 14 y 15; por lo tanto, se evaluó también el ajuste de esta reespecificación (M_b), presentada en la Tabla 2, lográndose buenos resultados ya que los índices de ajuste mejoraron. Por otro lado, se observa que el ajuste del modelo M_a generalmente fue mejor en el grupo de varones comparado con el de mujeres, de tal modo que la reespecificación (covariación entre los ítems 14 y 15) produjo efectos más grandes en el grupo de mujeres.

El análisis de la invarianza de medición entre hombres y mujeres consideró también probar la equivalencia de la covarianza de error entre los ítems 14 y 15, dado que esta especificación ocurrió en varones y mujeres. Los resultados se muestran en la Tabla 3, hallándose que la invarianza configuracional es satisfactoria y converge con la literatura respecto a las tres variables latentes. En función a esta línea base configuracional, se probó la invarianza de las cargas factoriales (invarianza métrica), un nivel de invarianza más restringido que la anterior. En este aspecto se halló también un aceptable ajuste, ya que los índices de ajuste no cambiaron significativamente. Seguido, se examinó la invarianza del intercepto de los ítems (invarianza escalar), obteniendo resultados satisfactorios. Fundamentalmente, los resultados reportan primero, que los índices de ajuste SRMR y RMSEA se

* mfernandezal@usmp.pe

** arturojuarezg@hotmail.com

*** sikayax@yahoo.com.ar

Tabla 1
Matriz de varianzas y covarianzas de los ítems del MBI-GS

	MBI B1	MBI B2	MBI B3	MBI B4	MBI B6	MBI B5	MBI B7	MBI B10	MBI B11	MBI B12	MBI B16	MBI B8	MBI B9	MBI 14	MBI 15	MBI 13
MBI 1	2.485															
MBI 2	1.908	2.895														
MBI 3	1.637	1.886	2.716													
MBI 4	1.417	1.515	1.693	2.478												
MBI 6	1.303	1.308	1.388	1.415	2.195											
MBI 5	-0.052	0.175	0.002	-0.110	-0.162	2.238										
MBI 7	-0.192	-0.110	-0.234	-0.289	-0.212	0.607	1.914									
MBI 10	-0.185	-0.024	-0.133	-0.266	-0.269	0.538	0.563	1.270								
MBI 11	-0.203	-0.050	-0.186	-0.237	-0.226	0.438	0.484	0.522	0.901							
MBI 12	-0.250	-0.157	-0.235	-0.294	-0.259	0.483	0.585	0.468	0.584	1.221						
MBI 16	-0.391	-0.300	-0.366	-0.350	-0.407	0.613	0.667	0.582	0.531	0.558	1.520					
MBI 8	0.892	0.785	0.831	0.811	0.797	-0.206	-0.436	-0.333	-0.321	-0.363	-0.478	1.678				
MBI 9	0.933	0.877	0.953	0.921	0.881	-0.208	-0.382	-0.292	-0.342	-0.430	-0.490	1.241	1.602			
MBI 14	0.356	0.294	0.313	0.454	0.615	-0.234	-0.442	-0.426	-0.309	-0.377	-0.454	0.555	0.681	2.093		
MBI 15	0.448	0.317	0.428	0.499	0.526	-0.378	-0.544	-0.368	-0.333	-0.399	-0.424	0.774	0.784	1.142	2.171	
MBI 13	0.169	0.337	0.305	0.425	0.362	0.128	0.004	0.046	0.097	0.072	0.076	0.346	0.314	0.438	0.438	4.312
M	2.23	2.67	1.88	1.76	1.07	4.73	4.92	5.25	5.34	4.95	5.10	0.86	0.90	0.90	0.92	3.82
DE	1.577	1.702	1.648	1.574	1.482	1.496	1.384	1.127	950	1.105	1.233	1.296	1.266	1.447	1.474	2.077
As.	0.577	0.371	0.720	0.772	1.466	-1.278	-1.567	-2.250	-2.009	-1.290	-1.860	1.740	1.634	1.747	1.751	-0.532
Cu.	-0.465	-0.788	-0.379	-0.207	1.408	0.930	2.208	5.981	5.343	1.892	3.828	2.652	2.231	2.394	2.330	-1.103

* mfernandezal@usmp.pe

** arturojuarezg@hotmail.com

*** sikayax@yahoo.com.ar

Tabla 2
Evaluación de los modelos para MBI-GS, en la muestra total y submuestras

	SB- χ^2 (gl)	CFI	SRMR	RMSEA (IC 90 %)
Total				
M _a	443.4370 (87)	0.917	0.061	0.066 (0.060, 0.072)
M _b	301.331 (86)	0.950	0.054	0.052 (0.045, 0.058)
Varones				
M _a	205.724 (87)	0.962	0.064	0.062 (0.051, 0.073)
M _b	176.523 (86)	0.971	0.061	0.055 (0.043, 0.066)
Mujeres				
M _a	368.893 (87)	0.907	0.063	0.074 (0.067, 0.082)
M _b	221.799 (86)	0.955	0.055	0.052 (0.043, 0.060)

M_a: modelo de tres factores oblicuos (AE e IN, y EP). M_b: modelo M_areespecificado con ítem 14 ↔ ítem 15.

Tabla 3
Invarianza de medición del MBI-GS (varones vs. mujeres)

Tipos de invarianza	SB- χ^2 (gl)	CFI (Δ_{CFI})	SRMR	RMSEA (IC 90 %)
Configuracional	387.448 (172)	0.949	0.058	0.052 (0.045, 0.058)
Métrica	410.446 (182)	0.947 (0.002)	0.062	0.051 (0.045, 0.058)
Escalar	428.513 (194)	0.951 (0.004)	0.063	0.051 (0.044, 0.057)

mantuvieron en magnitudes aceptables (< 0.064 y < 0.053 , respectivamente), además de producirse solo ligeros cambios entre el CFI de los modelos (< 0.005), diferencias que en el CFI son similares o más pequeñas que los criterios de cambio del CFI propuestos por Chen, 2007 (< 0.01) y Cheung y Rensvold, 2002 (< 0.005); y segundo, estos resultados indican que las propiedades de medición se mantienen satisfactoriamente invariantes en los criterios de invarianza probados, es decir, configuracional, métrica y escalar entre varones y mujeres.

Parámetros de los ítems. Se reportaron también las cargas estructurales (Tabla 4), información que permite evaluar mejor la relación de cada ítem con todos los constructos modelados (Graham, Graham y Thompson, 2003; Thompson, 1997). Los coeficientes estructurales de cada ítem en su propio factor (iguales a las cargas factoriales o coeficientes β) en general, fueron mayores que las cargas estructurales en los otros factores, aspecto que fue más notorio en EP. Sin embargo, los ítems de AE e IN mostraron coeficientes estructurales en los otros

* mfernandezal@usmp.pe

** arturojuarezg@hotmail.com

*** sikayax@yahoo.com.ar

LIBERABIT: Lima (Perú) 21(1): 9-20, 2015

factores, cerca de 0.50, indicando que la interpretación de tales ítems puede ser factorialmente compleja.

Todas las cargas factoriales fueron estadísticamente significativas ($Z > 9.0$), excepto los ítems 14 y 15 en IN, e ítem 5 en EP, el resto de ítems (75 %) logró lo que Steenkamp y Van Trijp (1991) llamaron el *débil criterio de convergencia* (prueba estadísticas $Z > 1.96$) y *fuerte*

criterio de convergencia del constructo (cargas factoriales ≥ 0.50).

Confiabilidad

Los resultados de la confiabilidad (Tabla 4), estimados desde el enfoque de la teoría clásica de los test (coeficiente α [Cronbach, 1951]) y el modelamiento factorial (coeficiente

Tabla 4
Parámetros de medición de los ítems, correlaciones interfactoriales y confiabilidad

	Coeficientes β (cargas)				Coeficientes estructura		
	AE	EP	IN	h^2	AE	EP	IN
B1	0.804			0.647	0.804	-0.182	0.498
B2	0.800			0.641	0.800	-0.181	0.495
B3	0.817			0.668	0.817	-0.185	0.506
B4	0.761			0.579	0.761	-0.172	0.471
B6	0.712			0.506	0.761	-0.172	0.471
B5		0.456		0.208	-0.103	0.456	-0.21
B7		0.559		0.313	-0.127	0.559	-0.257
B10		0.627		0.393	-0.142	0.627	-0.288
B11		0.740		0.548	-0.168	0.74	-0.34
B12		0.681		0.464	-0.154	0.681	-0.313
B16		0.649		0.421	-0.147	0.649	-0.298
B8			0.828	0.686	0.513	-0.381	0.828
B9			0.912	0.831	0.564	-0.419	0.912
B14			0.400	0.160	0.248	-0.184	0.4
B15			0.473	0.224	0.293	-0.218	0.473
ϕ							
AE	1				-	-	-
EP	-0.227	1			-	-	-
IN	0.619	-0.460	1		-	-	-
α	0.885	0.774	0.771		-	-	-
ω	0.886	0.790	0.761 (0.698)*		-	-	-
r_{ii}	0.606	0.384	0.465		-	-	-
AVE	0.608	0.391	0.475		-	-	-
\sqrt{AVE}	0.779	0.625	0.689		-	-	-

AE: agotamiento emocional. EP: eficacia profesional. IN: indiferencia. r_{ii} : correlación inter-ítem promedio. AVE: varianza extraída promedio. α y ω : coeficientes de confiabilidad. ϕ : correlación interfactorial.

* coeficiente corregido por errores correlacionados.

* mfernandezal@usmp.pe

** arturojuarezg@hotmail.com

*** sikayax@yahoo.com.ar

ω [McDonald, 1999]), arrojaron que estos estadísticos muestran valores muy similares entre sí. La magnitud de los mismos fue mayor a 0.75, por lo tanto los puntajes pueden considerarse relevantes para la investigación de grupo. La precisión poblacional de los coeficientes α se estimó por medio de intervalos de confianza (95 %), con el método exacto de Koning y Frances (2003); para AE, EP e IN fueron respectivamente: 0.873, 0.896; 0.751, 0.796; y 0.746, 0.794. Respecto al coeficiente ω para IN, se hizo una corrección de su estimación para tomar en cuenta los errores correlacionados (Lucke, 2005; Raykov, 2001) entre los ítems 14 y 15. Se obtuvo un coeficiente de menor magnitud aunque aún cercano a 0.70, sugiriendo que el impacto del error correlacionado sobre la confiabilidad puede ser sustancial.

Discusión

Se halló que los coeficientes estructurales de los ítems de EP fueron más diferenciados respecto a los otros factores en los que sus coeficientes fueron pequeños. Pero comparado con los coeficientes estructurales de los ítems de AE e IN, estos fueron menos diferenciados entre ellos, sugiriendo una complejidad factorial, es decir, que la interpretación de los ítems no es simple, pues muestran cargas factoriales elevadas en otros factores además del propio factor (Fleming y Merino, 2005; Merino y Grimaldo, 2010). Aunque la mayor magnitud de los ítems en sus propios factores indica que representan el constructo pretendido, existe una cantidad de varianza compartida que no puede dejar de tomarse en cuenta. Este resultado es coherente con la literatura, en que se tiende a encontrar una mayor covariación entre los factores AE e IN y esta complejidad puede ser razonable para comprender el funcionamiento de ambos constructos (Merino y Grimaldo, 2010), pero que no ha sido explorado en el nivel de los ítems pues los estudios revisados no reportan los coeficientes estructurales. Es probable que los resultados de los coeficientes estructurales sirvan para replantear nuevos ítems de AE o IN de acuerdo a la magnitud de estos coeficientes (ver por ejemplo, Merino, 2013), o expandir la definición de ambos constructos.

Por otro lado, la similaridad de los coeficientes de confiabilidad α y ω permite deducir que las cargas factoriales pueden considerarse similares, y que los ítems

correspondientes mantienen similar importancia en la definición cuantitativa del constructo, especialmente para AE. En este factor puede afirmarse que satisface el modelo tau-equivalente para definir su modelo de confiabilidad, lo que significa que el coeficiente α es un estimador aceptable (Zinbarg, Revelle, Yovel y Li, 2005). En la situación de EP, las cargas fueron moderadas, y su rango de variación fue comparativamente mayor que AE, aspecto que impactó en su bajo monto de varianza promedio (AVE). Dada la similaridad entre sus coeficientes α y ω , puede aceptarse también un modelo tau-equivalente para estimar la confiabilidad (McDonald, 1989; Zinbarg et al., 2005). Respecto a IN, el rango de variabilidad de sus cargas factoriales fue mayor entre dos pares de ítems (8 y 9, 14 y 15), aunque este efecto se canceló debido que no se produjo una diferencia sustancial en sus coeficientes α y ω . Pero esta diferencia debe tomarse en cuenta en la interpretación de este constructo, pues es un resultado que parece no encontrarse en otros estudios y puede representar algo propio de la muestra de participantes.

Respecto a la covariación entre los errores de los ítems 14 y 15, esto no solo parece ser consecuencia de relaciones conceptuales adicionales a la varianza común de su constructo, pues también pueden provenir del modo en que están fraseados, dado que ambos se encuentran consecutivamente presentados en el cuestionario, o por una interacción entre estos posibles motivos. Estos dos aspectos han sido reconocidos como fuentes de varianza única (Brown, 2006), y que pueden ser modelados como ocurrió en el presente estudio. Si el ordenamiento de estos ítems es una razonable explicación, una forma de evitar que esta varianza única aparezca sería alterar su orden de presentación dentro del cuestionario. Y con respecto al ítem 13, parece que su futura inclusión en el MBI-GS debe considerar reformularlo o sencillamente removerlo, pues nuestros resultados continúan cuestionando su contribución a la validez del instrumento.

La metodología aplicada permitió obtener las bases estadísticas para evaluar los modelos que mejor representan el comportamiento estructural de los ítems. Sin embargo, esto debe entenderse como aproximaciones más que representaciones exactas (Bentler y Dugeon, 1996) de la forma como interaccionan las conductas muestreadas por los ítems. Por otro lado, considerando que las

* mfernandezal@usmp.pe

** arturojuarezg@hotmail.com

*** sikayax@yahoo.com.ar

variaciones de muestreo son menores cuando se analizan datos en muestras relativamente grandes, es plausible que la estructura del MBI-GS es una fuerte hipótesis generalizable, y que se pueden obtener elevados parámetros de los ítems, una condición que sugiere una apropiada recuperación de los factores (Costello y Osborne, 2005; MacCallum, Widaman, Zhang y Hong, 1999; Preacher y MacCallum, 2002).

Los hallazgos respecto a la invariancia de medición del MBI-GS fueron buenos, indicando que sus propiedades estructurales se mantienen con pocos cambios entre las muestras de acuerdo al sexo. Pero hay otras consecuencias prácticas de la IM lograda, pues esto también permite que las comparaciones usando los puntajes y las medias observadas entre los grupos sean apropiadas de realizar (Steinmetz, 2013). La IM obtenida en la presente muestra indica que el instrumento puede generar respuestas levemente variantes entre los grupos, al menos entre varones y mujeres, pero sin aparente impacto sobre las propiedades de medición. Las propiedades de medición puestas a prueba permiten establecer que, en ambos grupos, puede identificarse la misma configuración factorial, que los ítems representan a su constructo de manera similar, y el escalamiento de las respuestas también es equivalente entre varones mujeres.

Los resultados del presente estudio dejan abiertas varias cuestiones, las mismas que forman parte de sus limitaciones; por ejemplo, el desbalance del tamaño muestral en los trabajadores de la muestra, y que estos provienen de Lima. Asimismo, no se probó directamente el modelo de medición apropiado para la confiabilidad. Incluso con estas limitaciones, estos hallazgos tienen implicancias relevantes en las mediciones que se realicen en la población como un criterio importante pero no único para realizar despistajes. Además, se requiere enfatizar el cambio conceptual que ocurre en la evaluación del MBI-GS con respecto a las otras versiones, esto es, una crisis en las relaciones de la persona con su trabajo sin que necesariamente esto signifique también una crisis en la relación de la persona con los otros miembros del centro laboral.

Referencias

- Aguilar, A. & Gutiérrez, R. (2007). El fenómeno de burnout en médicos de dos hospitales de la ciudad de Lima-Perú. *Revista de Psicología Herediana*, 2(1), 1-13.
- Bakker, A. B., Demerouti, E., & Schaufeli, W. B. (2002). Validation of the Maslach Burnout Inventory - General Survey: An Internet study. *Anxiety, Stress, and Coping*, 15, 245-260.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238-246.
- Bentler, P. M. & Dugeon, P. (1996). Covariance structure analysis: statistical practice, theory, and directions. *Annual Review of Psychology*, 47, 563-592.
- Bentler, P. M. & Wu, E. J. C. (2012). *EQS 6.2 for windows (Version 6.1) [Statistical Program]*. Encino, CA: Multivariate Software, Inc.
- Boomsma, A. (2000). Reporting analyses of covariance structures. *Structural Equation Modeling*, 7(3), 461-483.
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford Press.
- Byrne, B. M. (2008). Testing for multigroup equivalence of a measuring instrument: A walk through the process. *Psicothema*, 20(4), 872-882.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14(3), 464-504.
- Cheung, G. W. & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233-255.
- Chirkowska-Smolak, T. & Kleka, P. (2011). The Maslach Burnout Inventory-General Survey: Validation across different Occupational groups in Poland. *Polish Psychological Bulletin*, 42(2), 86-94.
- Costello, A. B. & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment Research & Evaluation*, 10(7). Recuperado de <http://pareonline.net/pdf/v10n7a.pdf>
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.
- Delgado, A. (2003). El síndrome de burnout en profesores de Educación Secundaria de Lima Metropolitana. *Revista de Investigación en Psicología*, 6(2), 26-47.
- Farrell, A. M. (2009). Insufficient discriminant validity: A comment on Bove, Pervan, Beatty and Shiu (2009). *Journal of Business Research*, 63(3), 324-327.
- Fernández, M. (2002). *Realidad psicossocial del maestro de Primaria*. Lima: Fondo de Desarrollo Editorial-Universidad de Lima.

* mfernandezal@usmp.pe

** arturojuarezg@hotmail.com

*** sikayax@yahoo.com.ar

- Fernández, M. (2008). Burnout, autoeficacia y estrés en maestros peruanos: tres estudios fácticos. *Ciencia & Trabajo*, 10(30), 120-125.
- Fernández, M., Merino, C., & Guimet, M. (en prensa). Propiedades psicométricas del Maslach Burnout Inventory-General Survey en una muestra de docentes de Lima, Perú. En R. León (Ed.), *Homenaje a Reynaldo Alarcón*.
- Flemming, J. & Merino, C. (2005). Medidas de simplicidad y ajuste factorial: Un enfoque para la construcción y revisión de escalas derivadas factorialmente. *Revista de Psicología*, 23(2), 252-266.
- Flores, C. A., Merino, C., Camacho, A., Juárez, A., & Placencia, O. (en prensa). Síndrome de burnout en instructores comunitarios: Propiedades psicométricas del Maslach Burnout Inventory General Survey (MBI-GS). En A. Juárez, (Ed.), *Evaluación psicosocial en el trabajo. Modelos, escalas y propiedades psicométricas en muestras mexicanas*.
- Fornell, C. & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18, 39-50.
- Gil-Monte, P. R. (2002). Validez factorial de la adaptación al español del Maslach Burnout Inventory-General Survey. *Salud pública de México*, 44(1), 36.
- Gomero, R., Palomino, J., Ruiz, F., & Llap, C. (2005). El síndrome de burnout en personal sanitario de los hospitales de la empresa minera de Southern Perú Copper Corporation: estudio piloto. *Revista Médica Herediana*, 16(4), 233-238.
- Graham, J. M., Graham, A. C., & Thompson, B. (2003). Consequences of not interpreting structure coefficients in published CFA research: A reminder. *Structural Equation Modeling*, 10, 142-153.
- Grajales, T. (2001). *Estudio de validez factorial del Maslach Burnout Inventory versión española en una población de profesionales mexicanos*. Recuperado de <http://www.tgrajales.net/mbivalidez.pdf>
- Green, D. E., Walkey, F. H., & Taylor, A. J. W. (1991). The three-factor structure of the Maslach Burnout Inventory. *Journal of Social Behavior and Personality*, 6, 453-472.
- Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (2006). *Multivariate Data Analysis* (6th ed.). New York: Prentice Hall.
- Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, 60, 581-592.
- Jackson, D. L., Gillaspay, J. A. Jr., & Purc-Stephenson, R. (2009). Reporting Practices in Confirmatory Factor Analysis: An Overview and Some Recommendations. *Psychological Methods*, 14, 6-23.
- Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, 34, 183-202.
- Jöreskog, K. G. & Sörbom, D. (1986). *LISREL VI: Analysis of linear structural relationships by maximum likelihood and least squares methods*. Mooresville, IN: Scientific Software.
- Juárez, A., García, J., Camacho, A., Gómez, V., Vera, A., Fernández, M., & García, I. (2011). *Traducción y adaptación del MBI-GS en población latina*. Documento no publicado.
- Kajantie, E. & Phillips, D. I. (2006). The effects of sex and hormonal status on the physiological response to acute psychosocial stress. *Psychoneuroendocrinology*, 31(2), 151-178.
- Koning, A. J. & Frances, P. H. (2003). *Confidence intervals for Cronbach's alpha coefficient values*. ERIM Report Series Reference No. ERS-2003-041-MKT. Recuperado de <http://ssrn.com/abstract=423658>
- Lei, P.-W. & Wu, Q. (2012). Estimation in structural equation modeling. En R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 164-179). New York: Guildford Press.
- Leiter, M. P. & Schaufeli, W. B. (1996). Consistency of the burnout construct across occupations. *Anxiety, Stress, and Coping*, 9, 229-243.
- Lucke, J. F. (2005). «Rassling the hog»: the influence of correlated item error on internal consistency, classical reliability, and congeneric reliability. *Applied Psychological Measurement*, 29(2), 106-125.
- MacCallum, R. C. & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology* 51, 201-226.
- MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, 4, 84-99.
- Mariños, A., Otero, M., Tomateo, J., & Málaga, G. (2011). Coexistencia de síndrome de burnout y síntomas depresivos en médicos residentes. Estudio descriptivo transversal en un hospital nacional de Lima. *Revista Médica Herediana*, 22(4), 162-168.
- Maslach, C., Jackson, S. E., & Leiter, M. P. (1996). *Maslach Burnout Inventory Manual* (3rd ed.). Palo Alto, CA: Consulting Psychologists Press.
- McDonald, R. P. (1989). An index of goodness of fit based on noncentrality. *Journal of Classification*, 6, 97-103.
- McDonald, R. P. (1999). *Test theory: A unified approach*. Mahwah, NJ: Erlbaum.
- Merino, C. (2013). Análisis factorial confirmatorio de grupo múltiple del EPQR-BV: Una versión muy abreviada del EPQR. *Internacional Journal of Psychological Research*, 6(2), 79-88.
- Merino, C. & Fernández, M. (noviembre, 2014). *Modelos latentes e invarianza de medición del MBI-GS en el Perú*. Ponencia en el V Congreso Iberoamericano y XI Nacional de profesionales y estudiantes de Psicología, Lima.

* mfernandezal@usmp.pe

** arturojuarezg@hotmail.com

*** sikayax@yahoo.com.ar

- Merino, C. & Grimaldo, M. (2010). Complejidad factorial y conductas moralmente controversiales. *Revista Argentina de Ciencias del Comportamiento*, 3(3), 38-43.
- Milfont, T. L. & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research*, 3, 112-131.
- Moreno, B., Rodríguez, R., & Escobar, E. (2001). La evaluación del burnout profesional. Factorialización del MBI-GS. Un análisis preliminar. *Ansiedad y estrés*, 7(1), 69-78.
- Oramas, A., González, A., & Vergara, A. (2007). El desgaste profesional. Evaluación y factorialización del MBI-GS. *Revista Cubana de Salud y Trabajo*, 8(1), 37-45.
- Preacher, K. J. & MacCallum, R. C. (2002). Exploratory factor analysis in behavior genetics research: factor recovery with small sample sizes. *Behavior Genetics*, 32, 153-161.
- Raykov, T. (2001). Bias in coefficient alpha for fixed congeneric measures with correlated errors. *Applied Psychological Measurement*, 25(1), 69-76.
- Remthulla, M., Brosseau-Liard, P. E., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods*, 17(3), 354-373.
- Romano, J. L., Kromrey, J. D., Owens, C. M., & Scott, H. M. (2011). Confidence interval methods for coefficient alpha on the basis of discrete, ordinal response items: Which one, if any, is the best? *The Journal of Experimental Education*, 79(4), 382-403.
- Salanova, M. & Schaufeli, W. B. (2000). Exposure to information technologies and its relation to burnout. *Behavior & Information Technology*, 19, 385-392.
- Satorra, A. & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. En A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (pp. 399-419). Thousand Oaks, CA: Sage.
- Schutte, N., Toppinnen, S., Kalimo, R., & Schaufeli, W. B. (2000). The factorial validity of the Maslach Burnout Inventory-General Survey across occupational groups and nations. *Journal of Occupational and Organizational Psychology*, 73, 53-66.
- Sörbom, D. (1989). Model modification. *Psychometrika*, 54, 371-384.
- Steenkamp, J. B. & Van Trijp, H. C. M. (1991). The use of LISREL in validating marketing constructs. *International Journal of Research in Marketing*, 8, 283-299.
- Steiger, J. H. & Lind, J. M. (June, 1980). *Statistically based tests for the number of common factors*. Paper presented at the annual meeting of the Psychometric Society, Iowa City, IA.
- Steinmetz, H. (2013). Analyzing observed composite differences across groups: Is partial measurement invariance enough? *Methodology*, 9(1), 1-12.
- Thompson, B. (1997). The importance of structure coefficients in structural equation modeling confirmatory factor analysis. *Educational and Psychological Measurement*, 57, 5-19.
- Tong, X. & Bentler, P. M. (2013). Evaluation of a new mean scaled and moment adjusted test statistic for SEM. *Structural Equations Modeling*, 20(1), 148-156.
- Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's α , Revelle's β , and McDonald's ω_h : Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*, 70, 123-133.

* Universidad de San Martín de Porres, Perú.

** Universidad Autónoma del Estado de Morelos, México.

* mfernandezal@usmp.pe

** arturojuarezg@hotmail.com

*** sikayax@yahoo.com.ar

LIBERABIT: Lima (Perú) 21(1): 9-20, 2015