

# Analysis of the Internal Structure of the Maslach Burnout Inventory (Human Service Survey) in Peruvian Physicians

Análisis de la estructura interna del *Maslach Burnout Inventory (Human Service Survey)* en médicos peruanos

Análises da estrutura interna do *Maslach Burnout Inventory (Human Service Survey)* em médicos peruanos

Gustavo Alexis Calderón-De la Cruz, Mg<sup>1\*</sup>

César Merino-Soto, Mg<sup>2</sup>

**Received:** October 8, 2019 · **Accepted:** February 25, 2020

**Doi:** <https://doi.org/10.12804/revistas.urosario.edu.co/revsalud/a.9275>

**To cite this article:** Calderón-De la Cruz GA, Merino-Soto C. Analysis of the internal structure of the Maslach Burnout Inventory (Human Service Survey) in peruvian Physicians. Rev Cienc Salud. 2020;18(2):1-17. <https://doi.org/10.12804/revistas.urosario.edu.co/revsalud/a.9275>

## Abstract

**Introduction:** The Maslach Burnout Inventory-Human Service Survey (MBI-HSS) is a measuring instrument applied to Peruvian physicians using the National Survey on User Satisfaction of Health Services (Ensalud) in 2014. However, evidence on the validity of its internal structure has not been established to date and it seems necessary to examine due to the various existing proposals related to the configuration of its factors. This study examines the validity of the internal structure of MBI-HSS in Peruvian physicians with secondary data from Ensalud. **Materials and Methods:** The sample consisted of 2222 doctors from different regions of Peru selected from a two-stage and stratified probability sampling. The validity of the internal structure of the MBI-HSS was assessed through exploratory and confirmatory factor analysis, and the reliability was calculated according to the internal consistency (alpha and omega coefficients). In addition, the differential item functioning (DIF) was evaluated according to gender. **Results:** An internal structure of three factors is demonstrated in the MBI-HSS with reduction of three items. The reliability was adequate (between .845 and .918), although it decreased considerably in the presence of correlated errors (between .335 and .517). Regarding the DIF, it was found that item 10 presented variation according to

1 Universidad de San Martín de Porres (Perú).

\* Corresponding author: [gcalderond@usmp.pe](mailto:gcalderond@usmp.pe)

2 Universidad Autónoma del Estado de Morelos (México).

gender. *Conclusion:* The original version (22 items) of the MBI-HSS is not appropriate to evaluate burnout in Peruvian doctors, the proposal of 19 items is viable, although as this is an initial validation study, the results must be replicated.

*Keywords:* Validation; burnout syndrome; physicians; Peru.

## Resumen

*Introducción:* el Maslach Burnout Inventory-Human Service Survey (MBI-HSS) es un instrumento de medición que fue aplicado en médicos peruanos utilizando la Encuesta Nacional de Satisfacción de Usuarios en Salud (Ensusalud) en el año 2014. Sin embargo, la evidencia de validez de su estructura interna no se ha elaborado hasta la actualidad y su examinación parece necesaria debido a las diferentes propuestas que existen sobre su configuración factorial. El presente estudio tiene por objetivo examinar la validez de la estructura interna del MBI-HSS en médicos peruanos con datos secundarios de la Ensusalud. *Materiales y métodos:* la muestra estuvo conformada por 2222 médicos de diferentes regiones de Perú seleccionados a partir de un muestreo probabilístico bietápico y estratificado. La validez de la estructura interna del MBI-HSS fue evaluada a través del análisis factorial exploratorio y confirmatorio; la confiabilidad fue estimada según la consistencia interna (coeficientes alfa y omega); y se evaluó el funcionamiento diferencial de los ítems (DIF) según el sexo. *Resultados:* se demuestra una estructura interna de tres factores en el MBI-HSS con reducción de tres ítems; la confiabilidad fue adecuada (entre 0.845 y 0.918), aunque disminuyó considerablemente en presencia de errores correlacionados (entre 0.335 y 0.517); en cuanto al DIF, se encontró que el ítem 10 presentó variación según el sexo. *Conclusión:* la versión original (22 ítems) del MBI-HSS no es pertinente para la evaluación del burnout en médicos peruanos, siendo viable la propuesta de 19 ítems, aunque al ser un estudio inicial de validación se requerirá una réplica de resultados.

*Palabras clave:* validación; síndrome de burnout; médicos; Perú.

## Resumo

*Introdução:* o Maslach Burnout Inventory-Human Service Survey (MBI-HSS) é um instrumento de medição que foi aplicado em médicos peruanos utilizando o Inquérito Nacional de Satisfação de Usuários em Saúde (Ensusalud) no ano 2014. No entanto, a evidência de validade de sua estrutura interna não se tem elaborado até hoje e seu exame parece necessário devido às diferentes propostas que existem sobre sua configuração fatorial. O presente estudo tem por objetivo, examinar a validade da estrutura interna do MBI-HSS em médicos peruanos com dados secundários do inquérito Ensusalud. *Materiais e métodos:* a amostra esteve conformada por 2222 médicos de diferentes regiões do Peru selecionados a partir de amostragem probabilística bi-etápica e estratificado. A validade da estrutura interna do MBI-HSS foi avaliada através da análise fatorial exploratória e confirmatória; a confiabilidade foi estimada segundo a consistência interna (coeficientes alfa e ômega); para além, avaliou-se o funcionamento diferencial dos itens (DIF) segundo o sexo. *Resultados:* demonstra-se uma estrutura interna de três fatores no MBI-HSS com redução de três itens; a confiabilidade foi adequada (entre 0.845 e 0.918), ainda que diminuíra consideravelmente em presença de erros correlacionados (entre 0.335 y 0.517); quanto ao DIF se encontrou que o item 10 apresentara variação segundo o sexo. *Conclusão:* a versão original (22 itens) do MBI-HSS não é pertinente para a avaliação do Burnout em médicos peruanos sendo viável a proposta de 19 itens, ainda que sendo um estudo inicial de validação se requererá uma réplica de resultados.

*Palavras-chave:* validação; síndrome de burnout; médicos; Peru.

## Introduction

Freudenberger and Maslach formally began the study of burnout syndrome during the 1970s. Considered a chronic response to work-related stress, burnout is explained as a progressive symptomatology of emotional exhaustion, depersonalization, and low personal accomplishment (1-7). Highly demanding jobs together with decreasing resources enable the onset of burnout, and its consequences impact workers' health and organizations' growth (8, 9). In addition, the effects resulting from burnout have proven to cause issues in workers' physiological (e.g., heart disease) and psychological (e.g., insomnia) conditions, in addition to affecting their performance at work (e.g., increased absenteeism) (10).

Burnout is particularly sensitive in jobs where health care activities are inherent to the position. As an analogy to the typical teacher–student relationship, medical work also requires constant interaction with patients (7, 11). This iterative process gives rise to overloading work demands, and work–family conflicts, job insecurity, and interpersonal problems emerge, which are considered sources of burnout (12-14). Demands at work may vary based on the socio-cultural and working scenario for physicians; however, in the particular case of burnout, there is agreement on its high prevalence across different regions of the world (15). Therefore, detecting this psychosocial risk is especially important in medical contexts, particularly where research reports on this topic are still considered to be on the rise (e.g., Peru). This emerging situation is associated with insufficient methodological accuracy of publications addressing this subject as well as low number of studies on the metric evidence concerning the validity of the instruments used (16). This study specifically assessed the Maslach Burnout Inventory (MBI).

The MBI is a psychological measurement instrument used to assess burnout. In its third edition, three versions were created targeting a) health professionals and social service workers (MBI-Human Service Survey, *MBI-HSS*), b) teachers (MBI-Educators Survey, *MBI-ED*), and c) general workers (MBI-General Survey, *MBI-GS*) (11). *MBI-HSS* was the first version of the instrument developed from clinical experiences and comprised 22 items referring to the constant worker–client (patient) interaction explaining the burnout symptomatology through three dimensions: Emotional Exhaustion (*EE*, the feeling of being exhausted and overwhelmed physically and psychologically; it represents the burnout central dimension), Depersonalization (*DE*, a cynical and detached attitude towards work and colleagues), and Low Personal Accomplishment (*PA*, tendency to negatively assess personal achievements at work; this dimension is positively assessed) (17).

*MBI-HSS* is one of the most widely used instruments globally (18, 19). In the case of Peru, is one of the countries that continuously reports its application in investigating burnout. *MBI-HSS*, however, poses questions that affect its usefulness, associated with its translations and the divergence between theory and practice (20). This has occurred mainly in Spanish-speaking

contexts, thus resulting in a deviation from the instrument's conception and understanding the construct (21).

The original version of MBI-HSS recommended the exclusion of items 12 (PA) and 16 (EE) for internal structure verification purposes, as these presented significant factor loadings in EE and DE dimensions, respectively (11). These recommendations obtained greater empirical evidence after the systematic and meta-analytical review conducted among female nurses (22). Several studies also reported inconsistency in the MBI-HSS internal structure across different work activities (e.g., nurses and other jobs), concluding the need to eliminate some items or modify their dimensions (23-29). These results and issues were also present in medical samples (30-32).

For several years now, the Peruvian government has recommended the study of burnout and its causes, in response to scarce research (33). In this endeavor, in 2014, the Instituto Nacional de Estadística e Informática (National Institute of Statistics and Information Technology, INEI) and the National Health Superintendency (Susalud) participated in the burnout assessment (through the use of MBI-HSS) as part of the User Satisfaction Survey (Ensusalud), a physician survey with a free database (<http://portal.susalud.gob.pe/blog/base-de-datos-2014/>). Including MBI-HSS can serve as a proxy for information related to the psychosocial discomfort of Peruvian physicians, mainly due to the assessment's national scope. However, the MBI-HSS metrics featured in this study have not been verified, and given that this instrument's measurement issues are globally known, it is not only important but also necessary to assess whether the representation of burnout by the MBI-HSS items remains consistent with the theory.

In this regard, a search for validation studies of MBI-HSS was conducted in databases such as Scopus, PubMed, and Scientific Electronic Library Online (SciELO) in Peru, using keywords in Spanish and English according to the following categories: a) construct: burnout syndrome; b) study type: *validación* validation; c) sample: *médicos, doctor* (physicians; doctor), and d) nationality: *peruano* (Peruvian), which resulted in the total absence of validation studies among the Peruvian population. A complementary free search was conducted using Google, identifying one study that obtained evidence of content validity from the assessment performed by three judges who reached a certain degree of agreement (34). However, no further details were provided regarding the judges' expertise or the kind of analysis used to assess the agreement and its results, casting doubts on this validation process and probably on the conclusions derived from the study's main objective. Such efforts can be considered emergent and preliminary in a context of scientific accuracy.

Despite the lack of instrumental publications and inadequate validity procedures for burnout in Peruvian physicians, empirical reports present results on the prevalence of burnout and its association with social and demographic variables by using MBI-HSS (35-37). Barring the report by Solis et al. that examined the content validity of MBI-HSS, the remaining studies justified the

use of this instrument with the validation process performed in contexts other than Peru, a premise adopted assuming cultural similarity (38). This practice, known as validity and reliability induction, is not recommended, as it involves transferring the validity or reliability evidence from one construct developed in one context to a new one, without enough rationality or the verified application sample and the suitability of its assessment (39). Taking the recommendations from the American Education Research Association, the American Psychological Association, and the National Council of Educational Measurement into account, the priority of assessing the internal structure of measurement instruments in contexts not previously examined is an essential requirement and a condition for other validation sources subsequent to the internal structure definition (40).

As the characterization of burnout is critical to clinical intervention, and with the above considerations in mind, this manuscript examines the factorial structure of MBI-HSS in Peruvian physicians using secondary data analysis. Moreover, the invariance of the measurement will be tested based on the item response theory through the analysis of differential item functioning (DIF) based on gender.

## Materials and Methods

### Research Design

This is an instrumental research design, which consists of applying one or more methodological strategies to examine the metric properties of a measurement (41, 42).

### Sample

Peruvian physicians were selected from a two-stage stratified probability sampling (<http://portal.susalud.gob.pe/wp-content/uploads/archivo/encuesta-sat-nac/2014/FICHA-TECNICA.pdf>). A total of 2228 physicians participated in this study, of which 6 were withdrawn as their MBI-HSS scores were not recorded in the survey database, leaving a total of 2222 participants for this study's analysis.

According to the sociodemographic characteristics, a greater proportion of male physicians (76.2%) between the ages of 24 and 83 years ( $M = 45.41$ ;  $SD = 11.061$ ) were found to work at the Ministry of Health-Regional Government institutions (MINSA-GR; 45.1%), ESSALUD (47%), private clinics (6.3%), the Peruvian Armed Forces (FFAA), and the Peruvian National Police (PNP; 1.6%). Their work experience ranged from 1 to 53 years ( $M = 16.50$ ;  $SD = 10.00$ ), having an indefinite, appointed, and permanent contract type (45.7%), undetermined according to Legislative Decree No. 728 (24.6%), special regime of administrative service contracting (CAS)

(13.1%), fixed-term contract (6.8%), professional fees (4.7%), and other (5%). It is also known that 43.7% of physicians hold other healthcare-related positions and 66% are specialized in their field of work.

The city of residence was distributed as follows: Lima (26%), Arequipa (8.6%), La Libertad (6.8%), Lambayeque (6.5%), Ica (5%), Ancash (4.7%), Cusco, (4.1%), Tacna (3.6%), Junin (3.4%), Piura (3.2%), Huánuco (2.9%), San Martín (2.7%), Cajamarca (2.7%), Apurimac (2.7%), Puno (2.6%), Ayacucho (2.3%), Loreto (2.2%), Tumbes (1.9%), Ucayali (1.6%), Moquegua (1.6%), Amazonas (1.5%), Callao (1.3%), Madre de Dios (1.1%), Huancavelica (0.7%), and Cerro de Pasco (0.4%). The physicians' marital status was classified as married (63.7%), single (23%), cohabitating (6.7%), divorced or separated (5.9%), and widowed (0.7%). The rest of the detailed information can be found in the aforementioned database.

## Instrument

### *Maslach Burnout Inventory Human Service Survey (MBI-HSS, 11)*

This instrument is composed of 22 items integrated into 3 dimensions: Emotional Exhaustion (EE, 9 items; “me siento emocionalmente agotado por mi trabajo”), Depersonalization (DE, five items; “Creo que trato a algunos de mis pacientes como si fueran objetos impersonales”) and Personal Accomplishment (PA, 8 items; “Trato muy eficazmente los problemas de los pacientes”). Each item can be answered using a 7 points Likert scale ranging from 0 (never) to 6 (every day). It could not identify the origin of the translation or adaptation used of the MBI-HSS for this national survey.

## Analysis

**Descriptive Statistics.** Descriptive statistics of the Mean (M), Standard Deviation (SD), Asymmetry (g1), and Kurtosis (g2) were examined.

**Internal Structure.** To examine the evidence of validity of the internal structure, the study sample was initially randomly divided into two groups.

In the first group (n = 1089), the Exploratory Factor Analysis (EFA) was carried out through the Factor program, version 10.7.01 (43). To fulfill the objective, the covariance matrix was specified according to the ordinal range of response, and the extraction of unweighted least squares was used with the prominent rotation method. Horn's parallel analysis was applied for factor identification (44). The adequacy of the correlation matrix was calculated using the Kaiser-Meyer-Olkin (KMO) and Bartlett's sphericity test. Based on this result, the researchers evaluated the factorial simplicity to verify the amount of variance of the factor loadings in relation to their dimension and compared with the other dimensions, verified at the item (ISF-F) and factor (ISF-F) levels. This analysis was conducted using the *SIMLOAD* program (45).

In the second sample ( $n = 1133$ ), the resulting EFA model was checked through a confirmatory factor analysis (CFA), implemented by means of the EQS program, version 6.2 (46). The poly-correlation matrices were used for factorial extraction. In view of the rejection of multivariate normality through the Mardia test (222.7499), the maximum likelihood estimation method was implemented with the Satorra-Bentler ( $SB-\chi^2$ ) correction. The selected adjustment indexes in the study were RMSEA ( $\leq 0,05$ ; 90% CI), CFI ( $\geq 0,95$ ) and SRMR ( $\leq 0,05$ ) (47). The acceptance criterion for factor loadings was expected to be  $>0.40$  (48).

In addition, the possible incorrect specifications of the analyzed models were assessed by checking the Sörbom Modification Indices (MI) (49). This procedure included evaluating their magnitude by using the SSV-SPSS program, considering the following as criteria: minimum factorial load of 0.40, statistical power at 0.75, and minimum parameter change (EPCMIN) at 20 (50, 51).

**Differential Item Functioning (DIF).** The monotonic association coefficient, *partial gamma* ( $\gamma^p$ ), was implemented (52). The use of this coefficient is suggested to identify the conditional relationship between items and grouping variables (e.g., gender), controlling the score of the examined scale (53-55). The levels for results qualification (54, 55) were weak ( $|0.00|$  to  $|0.150|$ ), moderate ( $|0.16|$  to  $|0.30|$ ), and strong ( $>|0.31|$ ).

**Reliability.** This was evaluated through internal consistency by means of the  $\alpha$  coefficient and the  $\omega$  coefficient (56, 57). To consider an adequate coefficient of internal consistency, values greater than or equal to 0.70 were expected.

## Results

### Internal Structure

The KMO result was satisfactory (0.90) and Bartlett's sphericity test was significant ( $p < 0.001$ ), being suitable to perform the EFA.

**Exploratory Factor Analysis.** The EFA (table 1) showed the replication of a three-dimensional structure of MBI-HSS; three items were removed for further analysis because one of them was a Heywood case (Burnout 1), whereas the other two cases (Burnout 16; Burnout 21) presented factor loadings below the expected criterion. The EFA was again replicated finding that the factor loadings were greater than 0.40. This result was corroborated with that obtained with the ISF-I, which showed that the items were influenced by their latent dimensions. This result was corroborated at the factor level (ISF-F).

**Table 1.** Descriptive Statistics and Exploratory Factor Analysis of MBI-HSS (n = 1089)

	Descriptive Analysis				First Exploratory Factor Analysis			Second Exploratory Factor Analysis			
	M	SD	As	Ku	F1 (EE)	F2 (DE)	F3 (PA)	F1 (EE)	F2 (DE)	F3 (PA)	ISF-I
MBI 1	2.21	1.731	.485	-.903	<b>1.02</b>	-.245	.092				
MBI 2	1.06	1.446	1.507	1.515	.765	-.076	-.085	.719	-.024	-.080	.980
MBI 3	1.34	1.500	1.242	.804	<b>.889</b>	-.143	-.054	.896	-.148	-.055	.955
MBI 6	1.36	1.502	1.263	.906	<b>.560</b>	.226	.035	<b>.611</b>	.133	-.001	.933
MBI 8	1.75	1.806	.942	-.285	<b>.880</b>	.020	.081	<b>.891</b>	.001	.075	.989
MBI 13	1.04	1.492	1.632	1.961	<b>.519</b>	.169	-.079	<b>.527</b>	.163	-.077	.845
MBI 14	2.10	1.982	.623	-.970	<b>.654</b>	.004	.114	<b>.665</b>	-.024	.105	.962
MBI 16	.99	1.365	1.810	3.024	<b>.390</b>	.357	-.064				
MBI 20	1.29	1.609	1.382	1.027	<b>.444</b>	.112	-.077	.480	.071	-.086	.923
MBI 5	.68	1.170	2.157	4.606	.070	<b>.643</b>	-.071	.114	<b>.577</b>	-.109	.896
MBI 10	1.40	1.599	1.227	.587	.119	<b>.839</b>	.152	.004	<b>.976</b>	.152	.965
MBI 11	1.35	1.658	1.331	.837	.310	<b>.564</b>	.101	.289	<b>.585</b>	.090	.700
MBI 15	.43	1.018	3.452	13.537	-.021	<b>.573</b>	-.169	.076	<b>.434</b>	-.219	.684
MBI 22	1.04	1.375	1.736	2.734	.025	<b>.448</b>	-.047	.040	<b>.423</b>	-.071	.946
MBI 4	5.08	1.545	-2.048	3.296	.128	.042	<b>.710</b>	.106	.060	<b>.711</b>	.957
MBI 7	5.13	1.314	-2.273	5.093	.085	.018	<b>.775</b>	.073	.036	<b>.784</b>	.984
MBI 9	5.07	1.407	-2.137	4.299	.145	-.039	<b>.783</b>	.086	.024	<b>.797</b>	.981
MBI 12	4.97	1.401	-1.911	3.375	-.365	.107	<b>.562</b>	-.387	.130	<b>.565</b>	.540
MBI 17	5.08	1.374	-2.125	4.252	.086	-.113	<b>.747</b>	.076	-.101	<b>.754</b>	.959
MBI 18	5.06	1.328	-2.049	4.081	.067	-.073	<b>.787</b>	.079	-.094	<b>.777</b>	.963
MBI 19	4.88	1.456	-1.658	2.234	-.158	.045	<b>.670</b>	-.128	-.003	<b>.641</b>	.943
MBI 21	3.22	2.166	-.110	-1.478	-.321	.121	<b>.300</b>				
Correlations											
F1 (EE)	--	--	--	--	--	--	--	1.000			
F2 (DE)	--	--	--	--	--	--	--	.785	1.000		
F3 (PA)	--	--	--	--	--	--	--	-.465	-.541	1.000	
isf	--	--	--	--	--	--	--	.918	.772	.964	

M = Mean; SD = Standard Deviation; As = Asymmetry; Ku: Kurtosis; EE = Emotional Exhaustion; DE = Depersonalization; PA = Personal Accomplishment; ISF-I = Item Level Factorial Simplicity Index; ISF-F = Factor Level Factorial Simplicity Index.

**Confirmatory Factor Analysis.** Concerning CFA (table 2), adjustment rates were partially adequate with the 19-item structure obtained from the exploratory analysis,  $SB-\chi^2_{(149)} = 720.016$  ( $p < 0.001$ ), CFI = 0.977; RMSEA = 0.058 (CI = 0.054, 0.062), and SRMR = 0.063. Factor loadings were moderate and high (between 0.53 and 0.82); the correlation between dimensions was theoretically consistent; however, EE and DE showed high covariation. Regarding the fit

improvement specifications (modification indexes, MI), the MIS reported 79 pairs of correlated errors between items, which affected the change in  $SB-\chi^2$ . When evaluating the practical significance of these MI, 19 (11.17%) of these pairs were relevant, but only 16 of them were selected because they were present within the same factor. The result was an improvement in the fit indexes of the  $SB-\chi^2_{(116)} = 396.734$  ( $p < 0.001$ ),  $CFI = 0.987$ ,  $RMSEA = 0.046$  ( $CI = 0.041, 0.051$ ), and  $SRMR = 0.055$ . The application of these new parameters indicated the existence of non-factor systematic variance.

**Differential Item Functioning (DIF).** Table 2 shows the results of the DIF evaluation. Barring item 10, the  $\gamma^p$  coefficients in the rest were not statistically significant ( $p > 0.40$ ), and were predominantly small in magnitude, indicated by point estimates and their confidence interval estimates. The homogeneity of the  $\gamma^p$  estimated at each score level (see *Homogeneity- $\chi^2$*  section) indicated that direction and size tend to be similar.

**Reliability.** Using a calculation of  $\alpha$  and  $\omega$ , the reliability was acceptable ( $>0.70$ ), but the presence of correlated errors required adjusting these estimates using the Raykov method (58). When modeling the pairs of correlated errors (16 pairs), the reliability estimates ( $\omega_{corrected}$ ) decreased considerably in all three dimensions (table 2). Discrepancies were also observed between  $\alpha$  and  $\omega$ , in which  $\alpha$  was always lower than uncorrected  $\omega$ .

**Table 2.** Confirmatory Factor Analysis of MBI-HSS (n = 1133) and DIF

	EE ( $\lambda$ )	DE ( $\lambda$ )	PA ( $\lambda$ )	$h^2$	Differential item functioning		
					$\gamma^p$	CI 95%	Homogeneity- $\chi^2$ (gl = 9)
MBI 2	.759			.576	-.116	-.234, .002	15.67
MBI 3	.786			.618	.141	.027, .255	6.23
MBI 6	.724			.525	-.022	-.138, .093	4.58
MBI 8	.801			.642	.136	.027, .245	13.32
MBI 13	.666			.443	.067	-.052, .186	1.56
MBI 14	.612			.374	-.000	-.103, .103	13.52
MBI 20	.621			.386	-.062	-.176, .052	11.15
MBI 5		.751		.564	-.093	-.225, .040	13.80
MBI 10		.795		.632	.315**	.202, .427	9.76
MBI 11		.815		.664	.053	-.070, .176	6.92
MBI 15		.701		.491	-.200	-.370, -.029	1.22
MBI 22		.531		.282	-.106	-.226, .014	12.53
MBI 4			.618	.382	.006	-.123, .135	13.18
MBI 7			.708	.501	-.104	-.236, .028	3.40
MBI 9			.687	.472	-.032	-.162, .067	1.05

Continue

	EE ( $\lambda$ )	DE ( $\lambda$ )	PA ( $\lambda$ )	h <sup>2</sup>	Differential item functioning		
					$\gamma^p$	CI 95%	Homogeneity- $\chi^2$ (gl = 9)
MBI 12			.687	.473	-.141	-.259, -.022	10.09
MBI 17			.778	.605	.204	.076, .329	12.18
MBI 18			.819	.671	.064	-.063, .191	12.37
MBI 19			.770	.593	.016	-.104, .137	7.63
Correlation							
EE	1						
DE	.811	1					
PA	-.491	-.581	1				
Reliability							
$\alpha$	.834	.738	.794				
$\omega$	.918	.845	.909				
$\omega_{corrected}$	.447	.335	.517				

$\lambda$  = Factor Loadings; h<sup>2</sup> = Communality;  $\alpha$  = Cronbach's alpha;  $\omega$  = Omega.  $\gamma^p$ : partial gamma. \*\*p < 0.01.

## Discussion

This study examined the validity of the internal structure of MBI-HSS in Peruvian physicians using secondary data collected by the 2014 Ensuslud survey. The results indicated the presence of a three-dimensional structure equal to the original proposal, although three items were eliminated due to the introduction of a Heywood case (item 1, EE) and low factor loadings (item 16, EE; item 21, PA) (11, 17). When considering the resulting factorial model of 19 items, the correlation between EE and DE was stronger, while the correlation between EE and PA was lower but around 0.50. In a more general context for our result interpretation, the decision to maintain or eliminate items affected the size of the correlations between the factors and the factor loadings. The item removals carried out in this study as well as in the previous studies suggest that the correlations between the constructs of MBI-HSS obtained in our study (and from other studies) require careful interpretation and may be of limited in terms of generalization, caused by content variations in the instrument's structure (23, 25, 59, 60). An additional implication of this inconsistent reduction in MBI-HSS is that the interpretation of the scores is between studies that removed different items.

Along with the problem of MBI-HSS content be altered by eliminating items, the inclusion of correlated errors indicates the inadequacy of the factor in retaining the variance that explains the responses to the items. The existence of correlated errors in this study may be a consequence of the application context, the reactions of those examined, the item content,

or the interaction between these components. It was impossible to detect which of these had the greatest effect on the present sample, but it is clear that the measurement quality of  $_{\text{MBI-HSS}}$  is compromised.

Another methodological consequence of adding correlated errors was calculating reliability. Indeed, reliability for internal consistency (uncorrected for correlated errors) reported acceptable levels, but after modeling the pairs of correlated errors, the coefficient  $\omega$  of each factor considerably decreased. This effect is not unknown in the methodological literature, and it is highly advisable to include the correlated errors to obtain greater accuracy when calculating reliability (58). Therefore, these adjusted estimates are theoretically more appropriate and indicate that variance attributed to items to their latent factor is less than 0.70 (the criterion chosen), even less than 0.60. One implication of this result is that error variance around the score leads to an inaccurate interpretation of scores as well as low replicability, thus making it difficult to guarantee the reproducibility of an examinee's score. Furthermore, the greater magnitude of the uncorrected  $\omega$  coefficient (which represents a model wherein items vary in their degree of validity with respect to their factor) was also observed compared with coefficient  $\alpha$  for each factor and suggests that a coefficient other than  $\alpha$  may be needed to estimate the measurement error in the version validated herein.

Given the representativeness of the sample and the metric results obtained,  $_{\text{MBI-HSS}}$  in its current state (original version or modified by this study) does not appear to be an option for calculating burnout in Peruvian physicians. This result is similar to other validation studies in nurses regarding item reduction (30, 31). Considering the recommendations of Maslach, Jackson and Leiter (11) on removing items 12 and 16, it appears that this removal should be extended to more than two items. In this sense, it is necessarily relevant to verify the factor structure of  $_{\text{MBI-HSS}}$  in each medical group, and by implication, in each Peruvian work group in which the instrument is used. This type of evaluation reduces the risk of inducing validity when it is not justified (39).

It was not apparent that differential functioning of items ( $_{\text{DIF}}$ ) existed in the adapted version, which validates the comparisons that can be made between men and women in  $_{\text{MBI-HSS}}$ . Although only one item was detected (item 10 of Depersonalization), the reason for this differentiated response is unclear and requires a mixed methods study involving preliminary quantitative analysis and post hoc qualitative analysis. The homogeneity assessment of the  $\gamma^p$  coefficients suggests that the variability due to  $_{\text{DIF}}$  is uniform and could suggest the absence of non-uniform  $_{\text{DIF}}$ . However, the detection of  $_{\text{DIF}}$  in this item entails only one step to determine if the bias in its interpretation is present. Overall, and given the magnitude of the gamma coefficients, it can be concluded that the version obtained in this study does not produce apparent significant differential performance.

In addition, the findings of this study open an overview of the adequate use required with  $_{\text{MBI-HSS}}$ , especially when this measurement instrument presents the most reports for the

burnout assessment. For Peruvian healthcare professionals in particular, this first report on internal structure indicates that the use of the original 22-item version can be disregarded and as a result, previous studies with MBI-HSS could potentially be disregarded as well. This situation may lead to greater compromise if these studies have entailed making decisions that affect the participants involved, thus leading to ethical implications about the inappropriate use of instruments involved in employment or clinical decision-making.

Questioning whether burnout continues to be one of the most frequent problems in the medical profession is less likely. The impact of its prevalence in Peru cannot yet be detailed, but it can be deduced from international studies that this psychosocial risk deserves to be addressed because of its levels, what causes it, and the consequences that it may entail for physicians and their organization.

The limitations of this study include lack of disaggregation according to the specific occupation of the medical staff; presentation of other factorial models for understanding burnout (e.g., bi-factor model); and lack of verification of other evidence of validity such as exploring its relationship with other constructs by verifying its relationship with other tests with a similar purpose or by evaluating the degree of prediction it provides in constructs such as depression, anxiety, somatic symptoms and sleep problems (61, 62). Furthermore, the type of DIF analyzed was uniform, and a formal examination of the non-uniform DIF is required. Finally, no analysis was applied directly calculating the number of factor cross loadings (e.g., ESEM: *exploratory structural equations modeling*), which may give a complementary or different perspective on the analysis performed herein (63).

Finally, we would like to add two final recommendations: first, the potential need to obtain an abbreviated version with generalizable and robust metric qualities and that guarantees the appropriate interpretation of scores. Indeed, with a balanced number of items representative of the core content of constructs measured by MBI-HSS, this goal is of scientific and practical value. Second, inferring the validity of MBI-HSS compromises the ethics of the researcher and the ineffectiveness of the use of an instrument intended to identify a problem requiring clinical and psychosocial care.

## Author Contributions

**G**ustavo Alexis Calderón-De la Cruz: a) Study concept and design, data acquisition and analysis, b) Article planning, c) Final review.

César Merino-Soto: a) Data analysis and interpretation, b) Review of significant intellectual content, c) Final article approval.

## Conflict of Interests

None declared.

## References

1. Freudenberger HJ. Staff burn-out. *J Soc Issues*. 1974;30(1):159-65. Doi: [10.1111/j.1540-4560.1974.tb00706.x](https://doi.org/10.1111/j.1540-4560.1974.tb00706.x)
2. Freudenberger HJ. The staff burn-out syndrome in alternative institutions. *Psychotherapy: Theory Research & Practice*. 1975;12(1):73-82. Doi: [10.1037/h0086411](https://doi.org/10.1037/h0086411)
3. Maslach C. Burned-out. *Human Behavior*. 1976;5(9):16-22.
4. Maslach C. Job burn-out. How people cope. *Public Welfare*. 1978;36(2):56-8.
5. Maslach C. The client role in staff burn-out. *J Soc Issues*. 1978;34(4):111-24. Doi: [10.1111/j.1540-4560.1978.tb00778.x](https://doi.org/10.1111/j.1540-4560.1978.tb00778.x)
6. Maslach C, Schaufeli WB. Historical and conceptual development of burnout. En Schaufeli WB, Maslach C, Marek T, editors. *Professional Burnout: Recent Developments in Theory and Research*. Washington, DC: Taylor & Francis; 1993. p. 1-16.
7. Maslach C, Schaufeli WB, Leiter MP. Job burnout. *Annu Rev Psychol*. 2001;52:397-422. Doi: [10.1146/annurev.psych.52.1.397](https://doi.org/10.1146/annurev.psych.52.1.397)
8. Alarcon GM. A meta-analysis of burnout with job demands, resources, and attitudes. *J VocatBehav*. 2011;79(2):549-62. Doi: [10.1016/j.jvb.2011.03.007](https://doi.org/10.1016/j.jvb.2011.03.007)
9. Lee RT, Ashforth BE. A meta-analytic examination of the correlates of the three dimensions of job burnout. *J Appl Psychol*. 1996;81(2):123-33. Doi: [10.1037/0021-9010.81.2.123](https://doi.org/10.1037/0021-9010.81.2.123)
10. Salvagioni DAJ, Melanda FN, Mesas AE, González AD, Gabani FL, Andrade SMD. Physical, psychological and occupational consequences of job burnout: A systematic review of prospective studies. *PLoS One*. 2017;12(10):e0185781. Doi: [10.1371/journal.pone.0185781](https://doi.org/10.1371/journal.pone.0185781)
11. Maslach C, Jackson SE, Leiter MP. *Maslach Burnout Inventory (3rd ed.)*. Palo Alto, CA: Consulting Psychology Press; 1996.
12. Leiter MP, Frank E, Matheson TJ. Demands, values, and burnout: relevance for physicians. *Can Fam Physician*. 2009;55(12):1224-5.e12256.
13. Adám S, Györffy Z, Susánszky É. Physician burnout in Hungary: A Potential Role for Work-Family Conflict. *J Health Psychol*. 2008;13(7):847-56. Doi: [10.1177/1359105308095055](https://doi.org/10.1177/1359105308095055)
14. Azam K, Khan A, Alam MT. Causes and adverse impact of physician burnout: A systematic review. *J Coll Physicians Surg Paki: JCPSP*. 2017;27(8):495-501.
15. Kumar S. Burnout and doctors: Prevalence, prevention and intervention. *Healthcare (Basel)*. 2016;4(3). Pii: E37. Doi: [10.3390/healthcare4030037](https://doi.org/10.3390/healthcare4030037)
16. Calderón-De la Cruz GA, Merino-Soto C, Juárez-García A, Dominguez-Lara S, Fernández-Arata M. ¿Es replicable la estructura factorial del Maslach Burnout Inventory Human

- Services Survey (MBI-HSS) en la profesión de enfermera del Perú?: un estudio nacional. *Enferm Clin*. 2020. Doi: [10.1016/j.enfcli.2019.12.013](https://doi.org/10.1016/j.enfcli.2019.12.013)
17. Maslach C, Jackson SE. *MBI: Maslach Burnout Inventory. Manual*. Palo Alto, CA: Consulting Psychologists Press; 1981.
  18. Schaufeli WB, Enzmann D. *The burnout companion to study and research: A critical analysis*. London: Taylor & Francis; 1998.
  19. Worley JA, Vassar M, Wheeler DL, Barnes LLB. Factor structure of scores from the Maslach Burnout Inventory: A review and meta-analysis of 45 exploratory and confirmatory factor-analytic studies. *Educ Psychol Meas*. 2008;68(5):797-823. Doi: [10.1177/0013164408315268](https://doi.org/10.1177/0013164408315268)
  20. Squires A, Finlayson C, Gerchow L, Cimiotti JP, Matthews A, Schwendimann R, et al. Methodological considerations when translation “burnout”. *Burn Res*. 2004;1(2): 59-68. Doi: [10.1016/j.burn.2014.07.001](https://doi.org/10.1016/j.burn.2014.07.001)
  21. Kulakova O, Moreno B, Garrosa E, Sanchez MO, Aragón A. Universalidad del constructo del Maslach Burnout Inventory en el contexto latinoamericano. *Acta de Investigación Psicológica*. 2014;7:2679-90. Doi: [10.1016/j.aiprr.2017.05.001](https://doi.org/10.1016/j.aiprr.2017.05.001)
  22. Loera B, Converso D, Viotti S. Evaluating the psychometric properties of the Maslach Burnout Inventory-Human Services Survey (MBI-HSS) among Italian Nurses: how many factors must a researcher consider? *PLoS ONE*. 2014;9(12):e114987. Doi: [10.1371/journal.pone.0114987](https://doi.org/10.1371/journal.pone.0114987)
  23. Lee, H, Chien T, Yen M. Examining factor structure of Maslach Burnout Inventory among nurses in Taiwan. *J Nurs Management*. 2012;21(4):648-56. Doi: [10.1111/j.1365-2834.2012.01427.x](https://doi.org/10.1111/j.1365-2834.2012.01427.x)
  24. Kanste O, Miettunen J, Kyngäs H. Factor structure of the Maslach Burnout Inventory among Finnish nursing staff. *Nurs Health Sci*. 2006;8:201-7. Doi: [10.1111/j.1442-2018.2006.00283.x](https://doi.org/10.1111/j.1442-2018.2006.00283.x)
  25. Olivares-Faúndez VE, Mena-Miranda L, Jélvez-Wilke C, Macía-Sepúlveda F. Validez factorial del Maslach Burnout Inventory Human Services (MBI-HSS) en profesionales chilenos. *Univ Psychol*. 2013;13(1):145-59. Doi: [10.11144/javeriana.UPSY13-1.vfmb](https://doi.org/10.11144/javeriana.UPSY13-1.vfmb)
  26. Poghosyan L, Aiken LH, Sloane DM. Factor structure of the Maslach Burnout Inventory: an analysis of data from large scale cross-sectional surveys of nurses from eight countries. *Int J Nurs Stud*. 2009;46(7):894-902. Doi: [10.1016/j.ijnurstu.2009.03.004](https://doi.org/10.1016/j.ijnurstu.2009.03.004)
  27. Vicente CS, Oliveira RA, Maroco J. Análise fatorial do Inventário de Burnout de Maslach (MBI-HSS) em profissionais portugueses. *Psicologia, Saúde & Doenças*. 2013;14(1):152-67.
  28. Gil-Monte PR. Factorial validity of the Maslach Burnout Inventory (MBI-HSS) among Spanish professionals. *Rev Saúde Pública*. 2005;39(1):1-8.
  29. Valente SDS, Wang YP, Menezes PR. Structural validity of the Maslach Burnout Inventory and influence of depressive symptoms in banking workplace: Unfastening the occupational conundrum. *Psychiatry Res*. 2018;267: 68-174. Doi: [10.1016/j.psychres.2018.05.069](https://doi.org/10.1016/j.psychres.2018.05.069)
  30. Córdoba L, Tamayo JA, Gonzáles MA, Martínez MI, Rosales A, Barbato SH. Adaptation and validation of the Maslach Burnout Inventory-Human Services Survey in Cali, Colombia. *Colomb Med*. 2011;42(3):286-93.

31. Naudé JIP, Rothmann S. The validation of the Maslach Burnout Inventory-Human Services Survey for emergency medical technicians in Gauteng. SAJIP. 2004;30(3):21-8.
32. Matejić B, Milenović M, KisićTepavčević D, Simić D, Pekmezović T, Worley JA. Psychometric properties of the Serbian version of the Maslach Burnout Inventory-Human Services Survey: A validation study among anesthesiologists from Belgrade teaching hospitals. The Scientific World Journal. 2015;1-8. Doi: [10.1155/2015/903597](https://doi.org/10.1155/2015/903597)
33. Ministerio de Trabajo y Promoción del Empleo. Informe técnico de los factores de riesgo psicosocial en trabajadores de Lima Metropolitana. Lima; 2014.
34. Solís-Cóndor R, Tantalean-del Águila M, Burgos-Aliaga R, Chambi-Torres J. Agotamiento profesional: prevalencia y factores asociados en médicos y enfermeras en siete regiones del Perú. An Fac Med. 2017;78(3):270-6. Doi: [10.15381/ana-les.v78i3.13757](https://doi.org/10.15381/ana-les.v78i3.13757)
35. Arteaga-Romaní A, Junes-Gonzales W, Navarrete-Saravia A. Prevalencia del síndrome de burnout en personal de salud. Rev Méd Panacea. 2014;4(2):40-4.
36. Maticorena-Quevedo J, Beas R, Anduaga-Beramendi A, Mayta-Trista P. Prevalencia del síndrome de burnout en médicos y enfermeras del Perú, Ensusalud 2014. Rev Peru Med Exp Salud Pública. 2016;33(2):241-7. Doi: [10.17843/rpmesp.2016.332.2170](https://doi.org/10.17843/rpmesp.2016.332.2170)
37. Beas R, Anduaga-Beramendi A, Maticorena-Quevedo J, Mayta-Tristán P. Factores asociados con el síndrome de Burnout en médicos y enfermeras, Perú. 2014. Rev Fac Cien Med Univ Nac Córdoba. 2017;74(4):331-7. Doi: [10.31053/1853.0605.v74.n4.16344](https://doi.org/10.31053/1853.0605.v74.n4.16344)
38. Díaz F, Gómez CI. La investigación sobre el síndrome de burnout en Latinoamérica entre 2000 y el 2010. Psicología desde el Caribe. 2016;33(1):113-31.
39. Merino-Soto C, Calderón-De la Cruz GA. Validez de estudios peruanos sobre estrés y burnout. Rev Peru Med Exp Salud Publica. 2018;35(2):353-4. Doi: [10.17843/rpme-sp.2018.352.3521](https://doi.org/10.17843/rpme-sp.2018.352.3521)
40. American Educational Research Association, American Psychological Association, National Council on Measurement in Education. Standards for Educational and Psychological Testing. Washington, DC: American Educational Research Association; 2014.
41. Ato M, López JJ, Benavente A. Un sistema de clasificación de los diseños de investigación en psicología. Anal Psicol. 2013;29(3):1038-59. Doi: [10.6018/analeps.29.3.178511](https://doi.org/10.6018/analeps.29.3.178511)
42. Montero I., Leon O. A guide for naming research studies in Psychology. Int J Clin Health Psychol. 2007;7(3):847-62.
43. Lorenzo-Seva U, Ferrando PJ. Factor: A computer program to fit the exploratory factor analysis model. Behav Res Methods. 2006;38:88-91. Doi: [10.3758/BF03192753](https://doi.org/10.3758/BF03192753)
44. Horn JL. A rationale and test for the number of factors in factor analysis. Psychometrika. 1965;32:179-85. Doi: [10.1007/BF02289447](https://doi.org/10.1007/BF02289447)
45. Fleming JS. Computing measures of simplicity of fit for loadings in factor-analytically derived scales. Behav Res Methods Instrum Comput. 2003;34:520-4.
46. Bentler PM, Wu EJC. EQS 6.2 for windows [Statistical Program]. Encino: Multivariate Software, Inc; 2012.

47. Satorra A, Bentler PM. Corrections to test statistics and standard errors in covariance structure analysis. En von Eye A, Clogg CC, editors, *Latent variables analysis: Applications for developmental research*. Thousand Oaks, CA: Sage; 1994. p. 399-419.
48. Brown TA. *Confirmatory factor analysis for applied research*. 2<sup>nd</sup> ed. New York: The Guilford Press; 2015.
49. Sörbom, D. Model modification. *Psychometrika*. 1989;54(3):371-84. Doi: [10.1007/BF02294623](https://doi.org/10.1007/BF02294623)
50. Domínguez-Lara S, Merino-Soto C. Evaluación de las malas especificaciones en modelos de ecuaciones estructurales. *Rev Argent Cienc Comport*. 2018;10(2):19-24. Doi: [10.32348/1852.4206.v10.n2.19595](https://doi.org/10.32348/1852.4206.v10.n2.19595)
51. Saris WE, Satorra A, van der Veld WM. Testing structural equation modeling or detection of misspecifications? *Struct Equ Modeling*. 2009;16(4):561-82. Doi: [10.1080/10705510903203433](https://doi.org/10.1080/10705510903203433)
52. Davis JA. A Partial Coefficient for Goodman and Kruskal's Gamma. *J Am Stat Assoc*. 1967;62(317):189-93. Doi: [10.1080/01621459.1967.10482900](https://doi.org/10.1080/01621459.1967.10482900)
53. Kreiner S. Analysis of multidimensional contingency tables by exact conditional tests. *Scandinavian Journal of Statistic*. 1987;14:97-112.
54. Schnohr CW, Makransky G, Kreiner S, Torsheim T, Hofmann F, De Clercq B, Elgar FJ, Currie C. Item response drift in the Family Affluence Scale: A study on three consecutive surveys of the Health Behaviour in School-aged Children (HBSC) survey Measurement. 2013;46, 3119-3126. Doi: [10.1016/j.measurement.2013.06.016](https://doi.org/10.1016/j.measurement.2013.06.016)
55. Schnohr CW, Kreiner S, Due EP, Currie C, Boyce W, Diderichsen F. Differential Item Functioning of a Family Affluence Scale: Validation Study on Data from HBSC 2001/02. *Soc Indic Res*. 2007;89(1):79-95. Doi: [10.1007/s11205-007-9221-4](https://doi.org/10.1007/s11205-007-9221-4)
56. Cronbach L. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;16:297-334. Doi: [10.1007/BF02310555](https://doi.org/10.1007/BF02310555)
57. McDonald RP. *Test theory: A unified treatment*. Mahwah, N.J.: L. Erlbaum Associates; 1999.
58. Raykov T. Point and interval estimation of reliability for multiple-component measuring instruments via linear constraint covariance structure modeling. *Structural Equation Modeling*. 2004;11(3):452-83. Doi: [10.1207/s15328007sem1103](https://doi.org/10.1207/s15328007sem1103)
59. Chao SF, McCallion P, Nickle T. Factorial validity and consistency of the Maslach Burnout Inventory among staff working with persons with intellectual disability and dementia. *J Intellect Disabil Res*. 2011;55(5):529-36. Doi: [10.1111/j.1365-2788.2011.01413.x](https://doi.org/10.1111/j.1365-2788.2011.01413.x)
60. Samaranyake DBDL, Seneviratne S.R. Validity of the Maslach Burnout Inventory-Human Services Survey among Sri Lankan nursing officer. *Psychol Stud*; 2011;57(1):101-11. Doi: [10.1007/s12646-011-0135-5](https://doi.org/10.1007/s12646-011-0135-5)
61. Yu K, Lee, SM, Nesbit EA. Development of a culturally valid Counselor Burnout Inventory for Korean counselor. *Measurement and Evaluation in Counseling and Development*. 2008;41(3):152-61. Doi: [10.1080/07481756.2008.11909827](https://doi.org/10.1080/07481756.2008.11909827)

62. Pisanti R, Lombardo C, Lucidi F, Violani c, Lazzari D. Psychometric properties of the Maslach Burnout Inventory for Human Services among Italian nurses: a test of alternative models. *J Adv Nurs*. 2012;69(3):697-707. Doi: [10.1111/j.1365-2648.2012.06114.x](https://doi.org/10.1111/j.1365-2648.2012.06114.x)
63. Asparouhov T, Muthén B. Exploratory structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*. 2009;16(3):397-438. Doi: [10.1080/10705510903008204](https://doi.org/10.1080/10705510903008204)