

Medición con Ítem Único del Agotamiento Emocional Académico en Estudiantes Universitarios Peruanos: Evidencias de Validez y Confiabilidad

Single-Item Measure of Academic Emotional Exhaustion in Peruvian College Students: Evidences of Validity and Reliability

Sergio Alexis Dominguez-Lara¹ y César Merino-Soto²

Resumen

El objetivo del presente estudio fue obtener evidencias de validez y confiabilidad del Ítem Único de Agotamiento Emocional Académico (IUAEA). Fue evaluada una muestra total de 227 estudiantes universitarios (155 mujeres) entre 17 y 56 años ($M=23.52$, $DE=5.972$), y una submuestra de 116 sujetos (79 mujeres) para el estudio de estabilidad temporal. Además del IUAEA, fue utilizada la Escala de Cansancio Emocional (ECE) como criterio externo. Las evidencias de validez fueron analizadas con ecuaciones estructurales a través del modelamiento conjunto del IUAEA con los ítems del ECE; correlaciones entre el IUAEA y el ECE; y un análisis comparativo de su relación con aspectos sociodemográficos (sexo, actividad laboral). El análisis de confiabilidad realizado muestra resultados satisfactorios en cuanto a la consistencia interna y estabilidad temporal del IUAEA. Los resultados son consistentes con otros estudios acerca de la validez de la medida de ítem único. Se discuten los resultados y sus implicancias.

Palabras clave: ítem único, agotamiento emocional, validez, confiabilidad

Abstract

The aim of this study was to obtain evidence of validity and reliability of the Single Item Academic Emotional Exhaustion (IUAEA). Was assessed a total sample of 227 college students (155 women) between 17 and 56 years ($M=23.52$, $SD=5.972$), and a subsample of 116 subjects (79 women) for the temporal stability study. In addition to the IUAEA, it was used the Escala de Cansancio Emocional (ECE) as external criterion. Evidence of validity were analyzed with structural equation modeling through the whole modeling of IUAEA with items of ECE; correlations between IUAEA and ECE; and a comparative analysis of their relationship with sociodemographic aspects (sex, work activity). Reliability analysis performed shows satisfactory results in terms of internal consistency and temporal stability of IUAEA. The results are consistent with other studies on the validity of the measure single item. The results and their implications are discussed.

Keywords: single-item, emotional exhaustion, validity, reliability

¹ Magíster en Psicología Clínica y de la Salud. Docente, Universidad de San Martín de Porres. Instituto de Investigación de Psicología, Universidad de San Martín de Porres, Av. Tomás Marsano 242 (5to piso), Lima 34, Perú. Tel.: 0051988053909. Correo: sdominguezmpcs@gmail.com

² Magíster en Psicología Educativa. Docente, Universidad de San Martín de Porres. Instituto de Investigación de Psicología, Universidad de San Martín de Porres, Av. Tomás Marsano 242 (5to piso), Lima 34, Perú. Tel.: 005961626454. Correo: sikayax@yahoo.com.ar

Introducción

En la etapa universitaria existen eventos que son inherentes a dicho periodo formativo y que a la vez son percibidos como estresantes por el estudiante (Dominguez-Lara & Merino-Soto, 2015a; Hu & Schaufeli, 2009), como el excesivo trabajo para la casa, los exámenes finales y estudiar para los mismos (Martín-Monzón, 2007), los horarios (Escribano-Barreno & Díaz-Morales, 2013) e incluso evaluaciones periódicas (Peñacoba & Moreno, 1999). Estos eventos exigen al estudiante que despliegue estrategias de afrontamiento e involucramiento académico (Medrano, Moretti, & Ortiz, 2015), y si estas no son efectivas, traería como consecuencia el *burnout*.

El *burnout* puede definirse como la experiencia de *agotamiento emocional* (AE) crónico que lleva a la pérdida de motivación y que suele progresar hasta sentimientos de inadecuación y fracaso (Maslach, Schaufeli, & Leiter, 2001). Según algunos planteamientos se considera que dos de los componentes del modelo de Burnout (Maslach et al., 2001), *despersonalización* y *falta de realización personal*, no se observan significativamente en el *burnout* académico (BA; Caballero, Hederich, & Palacio, 2009; Martínez & Marques-Pinto, 2005), mientras que en otros estudios se indica que estos se manifiestan de forma distinta en dicho ámbito (Schaufeli, Martínez, Marques-Pinto, Salanova, & Bakker, 2002). En tal sentido, el *cinismo* *indiferencia* indicaría un desinterés y falta de compromiso hacia las clases, los estudios, etc.; y la *ineficacia* haría referencia al sentimiento de incompetencia y falta de confianza en las habilidades de estudio por parte del estudiante (Boada-Grau, Merino-Tejedor, Sánchez-García, Prizmic-Kuzmica, & Vigil-Colet, 2015; Moyano, & Riaño-Hernández, 2013; Schaufeli et al., 2002). Empíricamente se ha corroborado que estos componentes también pueden hallarse en adolescentes escolarizados, y es una conducta desadaptativa diferenciada de otras expresiones emocionales como la ansiedad y depresión (Moyano & Riaño-Hernández, 2013; Salmela-Aro, Kiuru, Leskinen, & Nurmi, 2009; Salmela-Aro, Savolainen, & Holopainen, 2009).

La evaluación y estudio del AE, considerado el núcleo del *burnout* (Rohland, Kruse, & Rohrer, 2004; Rosales, & Rosales, 2013; Van Dierendonck, Schaufeli, & Sixma, 1994), es importante porque en población universitaria se relaciona directamente con la ansiedad y depresión (Aluja, 1997; Caballero, Gutiérrez, & Palacio, 2015; Dominguez-Lara, 2013a; Neveu, 2007), estrés académico (Yussef, 2016), estrategias disfuncionales de regulación emocional (Dominguez-Lara, 2018a), e inversamente con la autoeficacia académica (Dominguez-Lara, 2013a) y la autoestima (González & Landero, 2007).

Si bien existen medidas estandarizadas para la evaluación del *burnout académico* (Rosales & Rosales, 2013), su extensión puede ser una limitación para su inclusión dentro de baterías de tests psicológicos en algunos contextos educativos. Por ello, es necesario contar con medidas breves válidas y confiables, y las medidas de ítem único podrían ser una opción adecuada. Éstas tienden a prescindir de los ítems redundantes, reducir la fatiga, frustración y aburrimiento de responder repetidamente ítems muy parecidos (Robins, Hendin, & Trzesniewski, 2001; van Hooff, Geurts, Kompier, & Taris, 2007). Además, están recomendados para constructos unidimensionales y con adecuada consistencia interna (Robins et al., 2001; Loo, 2001; von Hooff et al., 2007). No obstante, sus limitaciones más importantes es que son amplios en cuestión de su contenido y no es posible tomar decisiones importantes a partir de sus resultados (Loo, 2001; Fuch & Diamantopoulos, 2009). Asimismo, se postula que no es posible calcular su confiabilidad por consistencia interna (Petrescu, 2013), pero desarrollos metodológicos recientes indican lo contrario (Postmes, Haslam, & Jans, 2012; Wanous, & Hudy, 2001; Wanous & Reichers, 1996), aceptando magnitudes $\geq .70$ como adecuadas (Wanous & Hudy, 2001; Wanous, Reichers & Hudy, 1997). En este contexto, y dado el carácter multidimensional del BA, es necesario enfocar la evaluación en la AE por ser la primera etapa de ese síndrome.

Existen antecedentes que avalan la propuesta de la medición del AE con ítem único. Por ejemplo, West y su equipo (West, Dyrbye, Sloan, & Shanafelt, 2009; West, Dyrbye, Satele,

Sloan, & Shanafelt, 2012) utilizaron un ítem correspondiente a la dimensión AE del *Maslach Burnout Inventory* (MBI; Maslach, Jackson, & Leiter, 1996) en profesionales de medicina (“*How often do you feel burned out from your work?*”). En base a una estrategia de validez concurrente, West et al. (2012) correlacionaron el puntaje del ítem con la subescala completa (excluyendo el ítem), y desde el cual se obtuvieron correlaciones positivas y estadísticamente significativas. Por otro lado, Rohland et al. (2004) crearon un ítem único de *burnout* en universitarios y profesionales de medicina, hallando correlaciones positivas y estadísticamente significativas con la subescala AE del MBI. En este punto se destaca que una estrategia analítica habitual es correlacionar el ítem único con una medida extensa del mismo constructo.

Propuesta de medición y análisis del ítem único de agotamiento emocional en universitarios

El *Ítem Único de Agotamiento Emocional Académico* (IUAEA) propuesto en el presente estudio es una medida global del AE académico, es decir, la experiencia de agotamiento asociada con los deberes comúnmente presentes en la vida universitaria de los estudiantes durante el semestre académico. Esta medida se viabiliza por la unidimensionalidad del constructo (Fontana, 2011; González & Landero, 2007; Dominguez-Lara, 2013a, 2014). El contenido propuesto para el ítem es “*Me siento emocionalmente agotado (tensión, angustia, preocupación) por las tareas (trabajos, exposiciones, exámenes, etc.) que demanda mi vida académica*”. Este contiene principalmente aclaraciones del significado del AE (“...*tensión, angustia, preocupación*...”), que van de acuerdo al marco conceptual de base (Maslach et al., 2001). Del mismo modo, se hace referencia a las labores realizadas en el ámbito académico (“...*trabajos, exposiciones, exámenes, etc.*...”) a fin de minimizar la posibilidad de que el estudiante responda el ítem en base a su experiencia con otro tipo de actividades (trabajo actual); además está contextualizado en una amplia perspectiva de vivencia universitaria del estudiante (“... *mi vida académica*”), considerando que es una evaluación global, no situacional, del AE en la universidad. Estas aclaraciones fueron necesarias para focalizar la

experiencia de AE en los estudiantes de manera similar.

Si bien los estudios con medidas de ítem único no son recientes en psicología, el uso del modelamiento de ecuaciones estructurales (SEM, por sus siglas en inglés) para la obtención de evidencias de validez es escaso (Postmes et al., 2012) y su aplicación en esta situación puede revelar patrones correlacionales no evaluados previamente en la literatura, los cuales se basaron en puntajes observados. Esta metodología SEM tiene ventajas importantes para corroborar hipótesis de validez, pues el análisis se orienta por especificaciones *a priori* basadas en la teoría desde un marco multivariado. En la presente investigación se estudió el IUAEA mediante el análisis SEM para la obtención de evidencias de validez, así como su confiabilidad por estabilidad temporal.

Método

Participantes

Se obtuvieron dos grupos de participantes para examinar las evidencias de validez y confiabilidad. La primera muestra formó parte del estudio de validez, y estuvo conformada por 227 estudiantes de una facultad de psicología de una universidad privada (155 mujeres), con edades comprendidas entre 17 y 56 años ($M=23.52$, $DE=5.972$), de los cuales 116 (51.1%) no trabajan, 100 (44.1%) trabajan, y 11 (4.8%) no consignaron ese dato en la encuesta. La edad de los participantes fue similar respecto al género ($t_{(216)}=.742$; $p>.05$; $d=.12$). Todos fueron predominantemente de nivel socioeconómico medio, nacidos en Lima Metropolitana (Perú) y residentes de varias zonas (distritos) dentro de esta misma ciudad. Todos cursaban entre el 1er y 4to año de estudios. Los criterios de exclusión fueron: no ser peruanos de nacimiento, no aceptar participar y un número sustancial de ítems o información demográfica sin responder.

Para obtener evidencias de confiabilidad por estabilidad, en la segunda muestra se contó con 116 sujetos (79 mujeres; una persona no consignó el dato), con edades comprendidas entre 17 y 56 años ($M=23.88$, $DE=6.45$). La edad de los participantes fue similar respecto al género

($t_{(113)}=1.614$; $p>.05$; $d=.33$). Este segundo grupo provino de la muestra de validez que aceptó participar en esta segunda ocasión. Entre ambos no hubo diferencias sustanciales en la distribución de sus características.

Instrumentos

Ítem Único de Agotamiento Emocional Académico (IUAEA; Dominguez-Lara, & Merino-Soto, presente estudio). Es una medición global del AE académico. El contenido del ítem es *Me siento emocionalmente agotado (tensión, angustia, preocupación) por las tareas (trabajos, exposiciones, exámenes, etc.) que demanda mi vida académica.* Las respuestas al IUAEA se realizan mediante un escalamiento ordinal de cinco puntos, los cuales van desde *Muy en desacuerdo* hasta *Muy de acuerdo*. Este escalamiento de cinco puntos fue elegido debido a que brinda más información acerca de las características evaluadas y proporciona mayor estabilidad de respuestas al realizar dos aplicaciones (Dominguez-Lara, 2013b; Lozano, García-Cueto, & Muñiz, 2008; Nunnally & Bernstein, 1994; Preston & Colman, 2000).

Escala de Cansancio Emocional (ECE; Fontana, 2011). Consta de 10 ítems que se puntúan de uno a cinco (de *raras veces a siempre*), considerando los 12 últimos meses de vida estudiantil. Para el presente estudio se utilizó la versión adaptada (Dominguez-Lara, 2013a, 2014a). En la presente investigación evidenció adecuada unidimensionalidad: $SB-\chi^2_{(35)}=68.533$ ($p<.01$), CFI=.984, RMSEA (IC 90%)=.065 (.042, .088), SRMR=.052; e indicadores de confiabilidad favorables ($\alpha=.876$, $\omega=.893$).

Procedimiento

El contacto y la solicitud de participación voluntaria fueron realizados de forma presencial, explicando a los participantes que se les contactará por una segunda ocasión para realizar el cuestionario nuevamente. Del mismo modo, se les garantizó la confidencialidad de los resultados, a fin de elevar la tasa de respuestas positivas. Los estudiantes que aceptaron participaron firmaron un consentimiento informado. La aplicación fue estandarizada y en horario regular de clases; todos los estudiantes presentes y que aceptaron

participar voluntariamente fueron incluidos en el estudio. Se mantuvo el mismo orden de las pruebas en cada grupo. El proyecto fue aprobado por el Instituto de Investigación de psicología de la institución en la cual se llevó a cabo el estudio, teniendo en consideración los aspectos éticos y procedimentales de la investigación.

Plan de análisis. De forma preliminar a cada procedimiento, se examinaron las características distribucionales de los puntajes del IUAEA, a través de sus estadísticos descriptivos. Del mismo modo, se estimó el posible *efecto de piso* (Terwee et al., 2007). Esto podría esperarse dada la tendencia que presentan los instrumentos que evalúan aspectos vinculados con psicopatología a generar respuestas negativas, las cuales pueden ser asimétricamente fuertes y positivas. Finalmente, se calculó el índice estandarizado de asimetría (SSI; Malgady, 2007) para evaluar el grado de asimetría del puntaje del IUAEA.

La evaluación de la validez estructural del IUAEA en la primera muestra se hizo para obtener evidencias de la estructura interna de forma conjunta con una medida extensa de AE. Se aplicó la metodología del modelamiento de ecuaciones estructurales (SEM) con el programa EQS 6.2 (Bentler & Wu, 2012): método de Máxima Verosimilitud (MV), correlaciones policóricas (Lee, Poon, & Bentler, 1995) considerando que es una medida ordinal (Dominguez-Lara, 2014b). Se usó la prueba general de bondad de ajuste χ^2 con la corrección de Satorra y Bentler ($SB-\chi^2$; 1994), el RMSEA ($\leq.05$; Steiger & Lind, 1980), y el CFI ($\geq.95$; Bentler, 1990).

En el proceso de modelamiento, se evaluó un modelo unidimensional (M_1) que incluía los diez ítems del ECE y el IUAEA, todos ellos influidos por la misma variable latente (es decir, AE). La racionalidad de este procedimiento radica en que el IUAEA también es una medida de AE, por lo cual debería compartir la variable latente común con los demás ítems. En el segundo modelo evaluado (M_2) se fijó la igualdad de los coeficientes de configuración (tau-equivalencia), indicando que todos los ítems (incluyendo el IUAEA) evalúan con la misma intensidad el AE. Para este modelo, en el caso se solicitó los índices de modificación (IM; Sörbom, 1989) a fin de

examinar las potenciales reespecificaciones, y se realizaron verificando qué ocurriría si alguna restricción impuesta (igualdad de cargas factoriales) debe ser relajada para mejorar el ajuste. La decisión fue tomada con base estadística de acuerdo con la significación estadística y el cambio significativo en el CFI ($\Delta_{CFI} < -.01$; Cheung & Rensvold, 2002) y NNFI ($\Delta_{NNFI} < -.01$; Gignac, 2007), así como con la racionalidad teórica de las mismas (Boomsma, 2000; Lei & Wu, 2012).

A fin de obtener evidencias de validez mediante la relación con criterios externos a nivel de puntajes observados, se realizaron correlaciones bivariadas del IUAEA con los ítems del ECE, cuantificando la relación conjunta a través del índice *Dm* (Dominguez-Lara, 2013c; Taras & Kline, 2010), esperando indicadores elevados que refrenden esta relación teórica; así como con el puntaje total a través de la *r*.

El índice *Dm* hace referencia a una medida de la desviación absoluta de las correlaciones observadas con respecto a los valores hipotetizados, siendo su rango de cero a uno, lo que indica que mientras más cercano a uno sea, existe una mayor precisión de las hipótesis realizadas: la unidad es obtenida cuando cada correlación observada es similar a las correlaciones hipotetizadas.

Por otro lado, para considerar una variable adicional, se calcularon las diferencias entre varones y mujeres en cuanto a los puntajes de IUAEA y ECE. Si bien la diferencia se expresada mediante el estadístico *t*, este se transformó a *r* (Abrami, Cohen, & d'Apollonia, 1988) con el objetivo de comparar esos coeficientes de correlación a través de un método específico para el caso de correlaciones con un elemento en común (en este caso, la variable *Sexo* y *Condición laboral*) por medio de sus intervalos de confianza (IC) (Zou, 2007). De este modo, si el IC de las diferencias contiene el cero, se concluye que estas no son estadísticamente significativas ($H_0: r_{IUAEA} - r_{ECE} = 0$). Un procedimiento similar se llevará a cabo comparando estudiantes que trabajan y no trabajan en ambas medidas (IUAEA y ECE).

Con la muestra total se estimó la consistencia interna del IUAEA en base a un método de aproximación (Dominguez-Lara & Merino-Soto,

2017; Wanous, & Hudy, 2001; Wanous & Reichers, 1996). Con la segunda muestra (parte de la muestra total que fue evaluada luego de cuatro semanas) fueron llevados a cabo tres procedimientos: 1) La estimación de la confiabilidad de los puntajes con el método de estabilidad temporal, procedimiento recomendado para estos casos (Denissen, Geenen, Selfhout, & Van Aken, 2007), mediante el coeficiente r_s ; 2) Se contrastaron los coeficientes α del IUAEA de los dos momentos mediante la comparación de sus IC (Dominguez-Lara & Merino-Soto, 2015b); 3) Se comparó el desempeño del grupo a través de dos momentos usando la *t* de Student para muestras relacionadas, esperando diferencias estadísticamente no significativas.

Resultados

Evidencias de Validez

Análisis descriptivos preliminares. En cuanto al aspecto descriptivo univariado, el IUAEA presentó una tendencia hacia la respuesta central ($M=3.053$; $DE=.925$), un grado de asimetría y curtosis aceptable ($g_1=-.273$; $g_2=-.541$; $SSI \leq .25$; Malgady, 2007), evidenciando ausencia tanto de *efecto de techo* (4.846%) como de *efecto de piso* (2.643%).

Evidencias de validez interna convergente mediante análisis SEM. En primera instancia se evaluó el ajuste del modelo unidimensional (M_1) que incluía los diez ítems del ECE y el UCE, todos ellos influidos por la misma variable latente (AE). De forma previa se analizó la normalidad multivariada, y al examinar el coeficiente de Mardia (1970) se observó un indicador de 16.44, el cual puede considerarse adecuado ($< .70$; Rodríguez & Ruiz, 2008). Los índices de ajuste estuvieron alrededor de los puntos de corte recomendados por la literatura (Tabla 1), lo cual indica que los once ítems (diez del ECE y el IUCE), se ven afectados por la misma variable latente. Cabe mencionar que el coeficiente de configuración para el IUAEA dentro de dicho modelo fue de .543.

Posteriormente se evaluó el ajuste del segundo modelo (M_2) que supone la igualdad de los coeficientes de configuración (modelo equivalente-tau), el cual indica que todos los ítems

Tabla 1. Análisis de los modelos de medición del IUAEA

	SB- χ^2 (gl)	CFI	Δ_{CFI}	NNFI	Δ_{NNFI}	SRMR	RMSEA (IC 90%)	AIC	CAIC
Dimensionalidad									
M ₁	77.41 (44)	.985	-	.982	-	.051	.058 (.036, .079)	-10.590	-205.288
Equivalencia métrica									
M ₂	117.9517 (54)	.972	-.013	.972	-.010	.093	.072 (.054, .090)	9.952	-228.996
M ₃	98.6422 (53)	.980	.008	.979	.007	.087	.062 (.042, .080)	-7.358	-241.880
M ₄	81.3806 (52)	.987	.007	.986	.007	.074	.050 (.027, .070)	-22.619	-252.717

Nota. M₁: modelo unidimensional (ECE+IUAE); M₂: Modelo unidimensional con igualdad de cargas factoriales; M₃: M₂ + residuales correlacionados de los ítems 5 y 6. M₄: M₃ + ítem 7 liberado de la restricción de igualdad.

evaluados (incluyendo el IUAEA), evalúan del mismo modo el AE. Si bien los índices de ajuste pueden considerarse adecuados, los IM sugieren correlacionar los residuales de los ítems cinco (*Tengo dolor de cabeza y otras molestias que afectan a mi rendimiento académico*) y seis (*Hay días que noto más la fatiga y me falta energía para concentrarme*), lo cual resulta teóricamente válido considerando los aspectos comunes que poseen vinculados a la ansiedad fisiológica; además dichos índices sugieren liberar la restricción de igualdad de la carga factorial del ítem siete (*Me siento emocionalmente agotado por mis estudios*), por presentar una carga factorial superior a las de los demás ítems. Esto no va en contra de la hipótesis inicial sobre el IUAEA, ya que el cambio sugerido no lo afecta directamente, es decir, la igualdad estadística de su coeficiente de configuración se mantiene con nueve de los 10 ítems del ECE. Se probaron los modelos considerando un evento a la vez.

Al evaluar el modelo tau-equivalente con los residuales relacionados de los ítems cinco y seis (M₃), fueron obtenidos índices de ajuste que superaron al modelo anterior, inclusive la covarianza entre residuales no resulta despreciable ($\delta_{56}=.246$; $r=.437$; $p<.05$). Finalmente, al considerar además la liberación del ítem siete (M₄), el ajuste mejora llegando a superar a M₁, aunque no de manera determinante, por lo cual la no-invarianza en el ítem siete podría explicarse mejor por el efecto del error de muestreo, pero no por aspectos relacionados al constructo.

Evidencias de validez del IUAEA con criterios externos. Al correlacionar el IUAEA con cada uno de los ítems de ECE fue obtenido un índice *Dm* de

magnitud elevada ($Dm=.834$; Dominguez-Lara, 2013c). Además, la correlación entre IUAEA y el puntaje total del ECE fue de $r=.496$ (IC95% .391, .588), denotando una magnitud del efecto mediano (Ellis, 2010).

Por otro lado, luego de calcular las diferencias de acuerdo al sexo y situación laboral de los puntajes de IUAEA y ECE, los estadísticos fueron transformados a *r*, y comparados (r_{IUAEA} vs r_{ECE}) a través de sus IC (Zou, 2007). De este modo, el IC de cada comparación contiene el cero, por lo tanto las diferencias no fueron estadísticamente significativas. Adicionalmente, se halló que la magnitud del efecto para tales diferencias fue insignificante (Cohen, 1992).

Confiabilidad

Consistencia interna. Para este análisis fue utilizada la propuesta de Wanous y su equipo (Wanous & Hudy, 2001; Wanous & Reichers, 1996), quienes estimaron la confiabilidad de las medidas de ítem único a partir de una modificación de la fórmula de desatenuación (Nunnally & Bernstein, 1994):

$$\hat{r}_{xy} = \frac{r_{xy}}{\sqrt{r_{xx} \cdot r_{yy}}}$$

Donde \hat{r}_{xy} es la correlación estimada entre IUAEA y el puntaje total del ECE; r_{xy} es la correlación entre IUAEA y ECE; r_{xx} es la confiabilidad por consistencia interna del ECE; y r_{yy} es la confiabilidad por consistencia interna del IUAEA. La correlación estimada (\hat{r}_{xy}) podría considerarse 1.00, es decir, una relación perfecta entre ambos, o .90 como algo más conservador (Dolbier et al., 2005), pero se utilizó una aproximación

Tabla 2. Estadísticos descriptivos y comparación de magnitud del efecto

Sexo	n	M	DE	t	r'	IC 95% r	IC 95% $r_{IUAEA} - r_{ECE}$	q
Muestra 1								
IUAEA								
Varón	72	20.99	6.234	-3.559**	.231	.104, .351		
Mujer	155	24.53	7.299				-.059, .196	.066
ECE								
Varón	72	2.83	1.021	-2.464**	.162	-.032, .286		
Mujer	155	3.15	.861					
Muestra 2								
IUAEA								
Trabaja	100	3.09	.842	.518	.034	-.010, .167		
No trabaja	116	3.01	.982				-.157, .111	.023
ECE								
Trabaja	100	23.26	6.576	.852	.057	-.077, .189		
No trabaja	116	23.44	7.438					

Nota. M = Media aritmética; DE=Desviación estándar; t=t de Student; r'=Coeficiente de correlación de Pearson estimado; q: q de Cohen.

metaanalítica que indica una magnitud promedio para la relación entre ítems únicos y sus medidas extensas en ciencias sociales: $r_{xy}=.64$ (Postmes et al., 2012). En este sentido, considerando que el α del ECE fue .876; la correlación entre IUAEA y ECE, .496, y $r_{xy}=.64$, la confiabilidad por consistencia interna del puntaje del IUAEA (α_{IUAEA}) asciende a .686, cercana al límite propuesto ($\geq .70$; Wanous et al., 1997) para la muestra total.

Estabilidad temporal. Un análisis descriptivo de los estadísticos en el primer y segundo momento indica que el patrón de respuesta se mantuvo constante, así como la distribución de las alternativas. Fue obtenido un coeficiente r_s moderado y estadísticamente significativo ($r_s=.539$; IC95% .396, .657; $p<.001$). Del mismo modo, se examinó si el monto estimado de error de medición se mantuvo constante en las dos aplicaciones. Así, fueron comparados los coeficientes α del IUAEA del primer (α_1) y segundo momento (α_2). El α_1 y α_2 para el ECE fue de .873 y .900 ($\chi^2_{(1)}=2.559$, $p=.110$). Entonces, al realizar el cálculo del apartado anterior el α_1 y α_2 del IUAEA fueron .871 (IC95% .819, .909) y .977 (IC95% .967, .984), y respectivamente. Ambos coeficientes resultaron de magnitud adecuada ($\geq .70$; Wanous et al., 1997), pero el monto de error estimado no fue el mismo a lo largo de las aplicaciones. Finalmente, desde el punto de vista de comparación de grupos relacionados, no se encontraron diferencias estadísticamente significativas entre las dos aplicaciones ($t_{(115)}=.638$, $p=.525$).

Tabla 3. Resultados descriptivos y de confiabilidad del IUAEA

	M	DE	g_1	SSI	g_2	%Min	%Máx
Primer momento	3.05	.873	-.261	.01	-.517	3.448	1.724
Segundo momento	3.00	.942	-.318	.02	-.634	6.034	1.724

Nota. N=116; Min: Valor mínimo. Máx: Valor máximo. M: media aritmética. DE: desviación estándar. g_1 : asimetría de Fisher. g_2 : curtosis de Fisher. SSI: Índice estandarizado de Asimetría.

Discusión

El objetivo del presente estudio fue la validación de una medida de ítem único del AE académico en el contexto universitario, a fin de contar con una medida ultrabreve que permita una evaluación rápida y efectiva. Fueron implementados procedimientos que no son utilizados habitualmente para obtener evidencias de validez con este tipo de medidas, como el análisis SEM (aspectos estructurales y modelos de medición), y el cálculo de su confiabilidad.

Respecto a la validez, a nivel de variables latentes, el análisis realizado da cuenta de la validez convergente entre el IUAEA y los ítems de la ECE. Esta situación coincide conceptualmente con la literatura, ya que no fueron hallados estudios preliminares que evalúen la relación entre las medidas de ítem único de burnout y medidas más extensas a nivel estructural. Posteriormente, en el modelo base fue impuesta la igualdad de coeficiente de configuración. Los resultados apoyaron la hipótesis preliminar, salvo por uno de

los ítems del ECE que fue liberado de esa restricción. No obstante, si bien se halló una magnitud moderada del coeficiente de configuración del IUAEA dentro de la estructura evaluado, esto puede explicarse por la generalidad de su contenido. En síntesis, desde la metodología SEM, el IUAEA evalúa con la misma intensidad el AE académico que los ítems de una medida extensa (ECE).

Con los puntajes observados, el patrón relacional entre IUAEA y ECE fue similar al descrito previamente, y además coincidente con los hallazgos de estudios anteriores (Rohland et al., 2004.; West, Dyrbye, Sloan et al., 2009; West, Dyrbye, Satele et al., 2012), aunque en estos no establecieron procedimientos que consideren a modo de hipótesis la correlación esperada entre el ítem único y la medida más extensa, como en el presente estudio. En tal sentido, el uso de hipótesis previas supone un avance respecto al estudio de medidas de ítem único.

Por otro lado, con respecto a la relación con criterios externos de corte sociodemográfico (sexo, situación laboral), al comparar los grupos según los puntajes del IUAEA y ECE, no se halló diferencia estadísticamente significativa en cuanto a los resultados observados, lo cual es evidencia adicional de la equivalencia empírica entre las dos medidas.

Finalmente, los puntajes del IUAEA evidenciaron indicadores de confiabilidad adecuados, tanto a nivel de varianza verdadera acumulada, como de estabilidad temporal. Con relación al primer punto, fueron integrados dos procedimientos: el uso de la fórmula de desatenuación, y la incorporación de una magnitud correlacional estimada a partir de datos reales. Aunque los coeficientes variaron de una aplicación a otra, esta estimación está vinculada con las fluctuaciones en los puntajes de la medida extensa (ECE), por lo que sería conveniente otros estudios que aborden el IUAEA de forma independiente en cuanto a la estimación de la confiabilidad (p.e., SEM con ítem único; Petrescu, 2013).

Sobre la estabilidad temporal, la correlación fue moderada entre las dos aplicaciones, observándose similitudes tanto a nivel distribucional como de patrón de respuesta. No

obstante, es poco realista esperar correlaciones muy elevadas debido al carácter dinámico e insidioso del burnout, la variabilidad inherente a la muestra de estudio, y el amplio del contenido del IUAEA, debido a que el afronte de las exigencias académicas puede variar según la actividad (exámenes orales, exámenes escritos, exposiciones, etc.) y la etapa del ciclo académico.

En cuanto a las implicancias prácticas, en vista de la evidencia favorable sobre la relación del IUAEA y el ECE, el ítem único podría ser utilizado como prueba de *screening* en situaciones que involucren evaluaciones extensas en las cuales la brevedad de los instrumentos y el poco tiempo disponible sean aspectos importantes a tomar en cuenta para no disminuir la tasa de respuestas de los evaluados.

En cuanto a las limitaciones, además de las inherentes a las medidas de un solo ítem, el uso de una muestra con características particulares (estudiantes de psicología; mayor porcentaje de mujeres) puede llevar a resultados sesgados. De este modo, se recomienda ampliarla en cuanto a características y carreras profesionales a fin de generalizar los hallazgos, dado que es plausible argumentar que la experiencia de burnout no sea similar entre distintas carreras o tipo de institución educativa (pública o privada). Finalmente, el análisis SEM llevado a cabo involucró de forma conjunta al IUAEA y a la ECE; no obstante, para que los resultados obtenidos no sean dependientes del método analítico utilizado, se requerirá incluir estudio complementario de la dimensionalidad del IUAEA (Anderson & Gerbing, 1988), ya que existe evidencia favorable acerca del uso de SEM en otras medidas de ítem único (Dominguez-Lara, 2018b). Asimismo, dado que la experiencia de AE no es ajena al ámbito superior universitario (Salmela-Aro et al., 2008), también podrían explorarse las bondades psicométricas del IUAEA en otros grupos de edad, así como poner a prueba la invarianza de medición. Por otro lado, la mortandad muestral de los participantes en la aplicación del retest, puede generar que los participantes restantes sean autoseleccionados y poseer con características idiosincrásicas y no representativas de la muestra inicial, además de restringir la varianza. El impacto sobre la correlación test-retest es desconocido, pero una replicación permitirá responder a esta pregunta.

En conclusión, el IUAEA presenta evidencias de validez respecto a su asociación con medidas extensas de AE a nivel estructural y de puntuaciones observadas, y con relación a variables sociodemográficas, así como indicadores de confiabilidad en magnitudes aceptables.

Referencias

- Abrami, P. C., Cohen, P. A., & d'Apollonia, S. (1988). Implementation problems in meta-analysis. *Review of Educational Research*, 58, 151-179. doi:10.3102/00346543058002151
- Aluja, A. (1997). Burnout profesional en maestros y su relación con indicadores de salud mental. *Boletín de Psicología*, 55, 47-61.
- Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A Review and recommended two-step approach. *Psychological Bulletin*, 103(3), 411-423. doi:10.1037/0033-2909.103.3.411
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238-246. doi:10.1037/0033-2909.107.2.238
- Bentler, P. M., & Wu, E. J. C. (2012). *EQS 6.2 for windows* [Statistical Program]. Encino, CA: Multivariate Software, Inc.
- Boada-Grau, J., Merino-Tejedor, E., Sánchez-García, J., Prizmic-Kuzmica, A., & Vigil-Colet, A. (2015). Adaptation and psychometric properties of the SBI-U scale for Academic Burnout in university students. *Anales de Psicología*, 31(1), 290-297. doi:10.6018/analesps.31.1.168581
- Boomsma, A. (2000). Reporting analyses of covariance structures. *Structural Equation Modeling*, 7(3), 461-483. doi:10.1207/S15328007SEM0703_6
- Caballero, C., Hederich, C., & Palacio, J. (2009). El burnout académico: Delimitación del síndrome y factores asociados a su aparición. *Revista Latinoamericana de Psicología*, 42(1), 131-146.
- Caballero, C., Gutiérrez, O., & Palacio, J. (2015). Relación del burnout y el engagement con depresión, ansiedad y rendimiento académico en estudiantes universitarios. *Salud Uninorte*, 31(1), 59-69. doi:10.14482/sun.31.1.5085
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233-255. doi:10.1207/S15328007SEM0902_5
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159. doi:10.1037/0033-2909.112.1.155
- Denissen, J., Geenen, R., Selfhout, M., & Van Aken, M. (2007). Single-item big five ratings in a social network design. *European Journal of Personality*, 22(1), 37-54. doi: 10.1002/per.662
- Dolbier, C. L., Webster, J. A., McCalister, K. T., Mallon, M. W., & Steinhardt M. A. (2005). Reliability and validity of a single-item measure of job satisfaction. *American Journal of Health Promotion*, 19(3), 194-198. doi:10.4278/0890-1171-19.3.194
- Dominguez-Lara, S. (2018a). Agotamiento emocional académico en estudiantes universitarios: ¿Cuánto influyen las estrategias cognitivas de regulación emocional? *Educación Médica*. 19(2), 96-103. doi:10.1016/j.edumed.2016.11.010
- Dominguez-Lara, S. (2018b). Ítem único de ansiedad ante exámenes: Análisis con modelos de ecuaciones estructurales. *Enfermería Clínica*, 28(2), 143-144. doi:10.1016/j.enfcli.2017.03.013
- Dominguez-Lara, S. (2013a). Análisis psicométrico de la Escala de Cansancio Emocional en estudiantes de una universidad privada. *Revista Digital de Investigación en Docencia Universitaria*, 7(1), 45-55. doi:10.19083/ridu.7.186
- Dominguez-Lara, S. (2013b). ¿Ítems Politémicos o Dicotómicos? Un estudio empírico con una escala unidimensional. *Revista Argentina de Ciencias del Comportamiento*, 5(3), 30-37.
- Dominguez-Lara, S. (2013c). Propuesta para el cálculo del Índice Dm para la validez de los ítems. *Interdisciplinaria*, 30(2), 297-303.
- Dominguez-Lara, S. (2014a). Escala de Cansancio Emocional: Estructura factorial y validez de los ítems en estudiantes de una universidad privada. *Avances en Psicología*, 21(1), 89-97.
- Dominguez-Lara, S. (2014b). ¿Matrices Policóricas/Tetracóricas o Matrices Pearson? Un estudio metodológico. *Revista Argentina de Ciencias del Comportamiento*, 6(1), 39-48.

- Dominguez-Lara, S., & Merino-Soto, C. (2015a). Una versión breve del *Cognitive Emotional Regulation Questionnaire*: Análisis estructural del CERQ-18 en estudiantes universitarios limeños. *Revista Peruana de Psicología y Trabajo Social*, 4(1), 25-36.
- Dominguez-Lara, S., & Merino-Soto, C. (2015b). ¿Por qué es importante reportar los intervalos de confianza del coeficiente alfa de Cronbach? *Revista Latinoamericana de Ciencias Sociales, Niñez y Juventud*, 13(2), 1326-1328.
- Dominguez-Lara, S., & Merino-Soto, C. (2017). Fiabilidad por consistencia interna de medidas de un solo ítem. *Actas Urológicas Españolas*, 41(3), 213. doi:10.1016/j.acuro.2016.04.003
- Ellis, P. (2010). *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results*. Cambridge, UK: Cambridge University Press.
- Escribano-Barreno, C., & Díaz-Morales, J. F. (2013). Rendimiento académico en adolescentes matutinos y vespertinos. *Revista Iberoamericana de Diagnóstico y Evaluación – e Avaliação Psicológica*, 36(2), 163-182.
- Fontana, S. (2011). Estudio preliminar de las propiedades psicométricas de la escala de desgaste emocional para estudiantes universitarios. *Revista Argentina de Ciencias del Comportamiento*, 3(2), 44-48.
- Fuch, C., & Diamantopoulus, A. (2009). Using single-item measures for construct measurement in management research. *Die Betriebswirtschaft*, 69(2), 195 - 2010.
- Gignac, G.E. (2007). Multi-factor modeling in individual differences research: Some suggestions and recommendations. *Personality and Individual Differences*, 42, 37-48. doi:10.1016/j.paid.2006.06.019
- González, M., & Landero, R. (2007). Escala de Cansancio Emocional (ECE) para estudiantes universitarios: Propiedades psicométricas en una muestra de México. *Anales de Psicología*, 23(2), 253-257.
- Hu, Q., & Schaufeli, W. B. (2009). The factorial validity of the Maslach Burnout Inventory – Student Survey in China. *Psychological Reports*, 105, 304-408. doi:10.2466/PRO.105.2.394-408
- Lee, S.-Y., Poon, W.-Y., & Bentler, P. M. (1995). A two-stage estimation of structural equation models with continuous and polytomous variables. *British Journal of Mathematical and Statistical Psychology*, 48, 339-358. doi:10.1007/BF02294660
- Lei, P.-W., & Wu, Q. (2012). Estimation in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 164-179). New York: Guildford Press.
- Lozano, L. M., García-Cueto, E., & Muñiz, J. (2008). Effect of the number of response categories on the reliability and validity of rating scales. *European Journal of Research Methods for the Behavioral and Social Sciences*, 4, 73-79. doi:10.1027/1614-2241.4.2.73
- Loo, R. (2001). A caveat on using single-item versus multiple-item scales. *Journal of Managerial Psychology*, 17(1), 68-75.
- Malgady, R. (2007). How skew are psychological data? A standardized index of effect size. *The Journal of General Psychology*, 134(3), 355-359. doi:10.3200/GENP.134.3.355-360
- Martínez, I., & Marques-Pinto, A. (2005). Burnout en estudiantes universitarios de España y Portugal y su relación con variables académicas. *Aletheia*, 21, 21-30.
- Maslach, C., Jackson, S. E., & Leiter, M. P. (1996). *Maslach Burnout Inventory Manual* (3rd ed.). Palo Alto, CA: Consulting Psychologists Press.
- Maslach, C., Schaufeli, W., & Leiter, M. (2001). Job Burnout. *Annual Review Psychology*, 52, 397-422. doi:10.1146/annurev.psych.52.1.397
- Mardia, K. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57, 519-530.
- Martín-Monzón, I. (2007). Estrés académico en estudiantes universitarios. *Apuntes de Psicología Colegio Oficial de Psicología*, 25(1), 87-99.
- Medrano, L., Moretti, L., & Ortiz, A. (2015). Medición del Engagement Académico en Estudiantes Universitarios. *Revista Iberoamericana de Diagnóstico y Evaluación – e Avaliação Psicológica*, 40(2), 114-123.
- Moyano, N., & Riaño-Hernández, D. (2013). Burnout escolar en adolescentes españoles: Adaptación y validación del School Burnout Inventory. *Ansiedad y Estrés*, 19(1), 95-103.

- Neveu, J. (2007). Jailed resources: Conservations of resources theory as applied to *burnout* among prison guards. *Journal of Organizational Behavior*, 28(1), 21-42. doi:10.1002/job.393
- Nunnally, J., & Bernstein, I. (1994). *Psychometric Theory* (3rd Ed.). New York: McGraw-Hill.
- Peñacoba, C., & Moreno, B. (1999). La escala de estresores universitarios (EEU). Una propuesta para evaluar el estrés en grupos de poblaciones específicas. *Ansiedad y Estrés*, 5(1), 61-78.
- Petrescu, M. (2013). Marketing research using single-item indicators in structural equation models. *Journal of Marketing Analytics*, 1(2), 99-117. doi:10.1057/jma.2013.7
- Postmes, T., Haslam, A., & Jans, L. (2012). A single-item measure of social identification: Reliability, validity, and utility. *British Journal of Social Psychology*, 53, 597-617. doi:10.1111/bjso.12006
- Preston, C., & Colman, A. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, 104, 1-15. doi:10.1016/S0001-6918(99)00050-5
- Robins, R., Hendin, H., & Trzeaniewski, K. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin*, 27, 151-161. doi:10.1177/0146167201272002
- Rodriguez, M., & Ruiz, M. (2008). Atenuación de la asimetría y de la curtosis de las puntuaciones observadas mediante transformaciones de variables: Incidencia sobre la estructura factorial. *Psicológica*, 29, 205-227.
- Rohland, B., Kruse, G., & Rohrer, J. (2004). Validation of a single-item measure of burnout against the Maslach Burnout Inventory among physicians. *Stress and Health*, 20, 75-79. doi:10.1002/smi.1002
- Rosales, Y., & Rosales, F. (2013). Hacia un modelo bidimensional del Síndrome de Burnout en estudiantes universitarios. *Ciência & Saúde Coletiva*, 19(12), 4767-4775. doi:10.1590/1413-812320141912.18562013
- Salmela-Aro, K., Kiuru, N., Leskinen, E., & Nurmi, J. E. (2009). School-Burnout Inventory (SBI) – Reliability and validity. *European Journal of Psychological Assessment*, 25(1), 48-57. doi:10.1027/1015-5759.25.1.48
- Salmela-Aro, K., Kiuru, N., Pietikäinen, E., & Jokela, J. (2008). Does school matter? The role of school context in adolescents' school-related burnout. *European Psychologist*, 13(1), 12-23. doi:10.1027/1016-9040.13.1.12
- Salmela-Aro, K., Savolainen, H., & Holopainen, L. (2009). Depressive symptoms and school burnout during adolescence: Evidence from two cross-lagged longitudinal studies. *Journal of Youth Adolescence*, 38, 1316-1327. doi:10.1007/s10964-008-9334-3
- Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (pp. 399-419). Thousand Oaks, CA: Sage.
- Schaufeli, W. B., Martínez, I. M., Marques, A., Salanova, M., & Bakker, A.B. (2002). Burnout and engagement in university students. A cross – national study. *Journal of Cross – Cultural Psychology*, 33(5), 464-481. doi:10.1177/0022022102033005003
- Sörbom, D. (1989) Model modification. *Psychometrika*, 54, 371-384. doi:10.1007/BF02294623
- Steiger, J. H., & Lind, J. C. (1980, May). *Statistically based tests for the number of common factors*. Paper presented at the annual meeting of the Psychometric Society, Iowa City, IA
- Taras, V., & Kline, T. (2010). Scale validation via quantifying item validity using the Dm Index. *Psychological Reports*, 107(2), 535-546. doi:10.2466/03.PR0.107.5.535-546
- Terwee, C. B., Bot, S. D. M., de Boer, M. R., van der Windt, D.A.W.M., Knol, D. L., Dekker, J., Bouter, L. M., & de Vet, H. C. W. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, 60(1), 34-42. doi:10.1016/j.jclinepi.2006.03.012
- Van Dierendonck, D., Schaufeli, W., & Sixma, H. (1994). Burnout among general practitioners:

- A perspective from equity theory. *Journal of Social and Clinical Psychology, 13*, 86-100.
- Van Hooff, M., Geurts, S., Kompier, M., & Taris, T. (2007). "How fatigued do you currently feel?" Convergent and discriminant validity of a single-item fatigue measure. *Journal of Occupational Health, 49*, 224-234.
- Wanous, J. P., & Hudy, M. J. (2001). Single-item reliability: A replication and extension. *Organizational Research Methods, 4*(4), 361-375. doi:10.1177/109442810144003
- Wanous, J. P., & Reichers, A. E. (1996). Estimating the reliability of a single-item measure. *Psychological Reports, 78*, 631-634. doi:10.2466/pr0.1996.78.2.631
- Wanous, J. P., Reichers, A. E., & Hudy, M. J. (1997). Overall job satisfaction: How good are single-item measures? *Journal of Applied Psychology, 82*(2), 247-252. doi:10.1037//0021-9010.82.2.247
- West, C., Dyrbye, L., Sloan, J., & Shanafelt, T. (2009). Single item measures of emotional exhaustion and depersonalization are useful for assessing burnout in medical professionals. *Journal of General Internal Medicine, 24*(12), 1318-1321. doi:10.1007/s11606-009-1129-z
- West, C., Dyrbye, L., Satele, D., Sloan, J., & Shanafelt, T. (2012). Concurrent validity of single-item measures of emotional exhaustion and depersonalization in burnout assessment. *Journal of General Internal Medicine, 27*(11), 1445-1452. doi:10.1007/s11606-012-2015-7
- Yussef, F. (2016). Medical student stress, burnout and depression in Trinidad y Tobago. *Academic Psychiatry, 40*(1), 69-75. doi:10.1007/s40596-015-0468-9
- Zou, G. Y. (2007). Toward using confidence intervals to compare correlations. *Psychological Methods, 12*(4), 399-413. doi:10.1037/1082-989X.12.4.399