



## ARTÍCULO ORIGINAL

### Análisis estructural y desarrollo de una versión breve de la versión en español del Inventario de Ansiedad ante Exámenes (TAI-E) en universitarios de Lima

*Structural analysis and development of a short form of the Test Anxiety Inventory – Spanish version (TAI-S) in college students from Lima*

Sergio Alexis Dominguez-Lara <sup>1\*</sup> y Freddy De la Cruz-Contreras <sup>2</sup>

<sup>1</sup> Universidad de San Martín de Porres, Perú

<sup>2</sup> Universidad Nacional Mayor de San Marcos, Perú.

\* Dirección de correspondencia: Sergio Dominguez-Lara. Instituto de Investigación de Psicología, Universidad de San Martín de Porres, Av. Tomás Marsano 242 (5to piso), Lima 34 – Perú. Email: [sdominguezmpcs@gmail.com](mailto:sdominguezmpcs@gmail.com), [sdominguezl@usmp.pe](mailto:sdominguezl@usmp.pe)

Recibido: 06 de febrero de 2017; Revisado: 12 de marzo de 2017; Aceptado: 13 de marzo de 2017; Publicado Online: 14 de marzo de 2017.

#### CITARLO COMO:

Dominguez-Lara, S., & De la Cruz-Contreras, F. (2017). Análisis estructural y desarrollo de una versión breve de la versión en español del Inventario de Ansiedad ante Exámenes (TAI-E) en universitarios de Lima. *Interacciones*, 3(1), 7-17. doi: 10.24016/2017.v3n1.50

#### PALABRAS CLAVE

Ansiedad ante exámenes;  
Confiability;  
Validez;  
Universitarios;  
Análisis factorial;  
Versión breve.

#### RESUMEN

El presente estudio tuvo como objetivos examinar la estructura interna de la versión en español del Inventario de Ansiedad ante Exámenes (TAI-E), y obtener una versión breve. Fue evaluada una muestra de 376 universitarios peruanos (73.1% mujeres) entre 16 y 55 años (M = 20.71). La estructura interna del TAI-E fue analizada mediante el modelamiento de ecuaciones estructurales (análisis factorial confirmatorio), y la versión breve fue obtenida teniendo como base un modelo Bifactor y criterios conceptuales. Fueron evaluados cuatro modelos: ortogonal, oblicuo, bifactor y unidimensional. En cuanto a los resultados, el TAI-E está constituido por una sola dimensión. Del mismo modo, la versión breve muestra indicadores robustos y evidencias de equivalencia empírica con la versión extensa. Finalmente, las dos versiones mostraron coeficientes de confiabilidad con magnitudes elevadas. En conclusiones, las dos versiones estudiadas evidencias propiedades psicométricas favorables que respaldan su uso en el contexto peruano.



## KEY WORDS

Test anxiety;  
Reliability;  
Validity;  
University students;  
Factorial analysis;  
Short form.

## ABSTRACT

The aim of the present study is to examine the internal structure of the Test Anxiety Inventory-Spanish version (TAI-S) and have a brief version. It was assessed of a sample of 376 Peruvian university students (73.1% women) between 16 and 55 years old ( $M = 20.71$ ). The internal structure of TAI-S was analyzed by using structural equation modeling (factorial confirmatory analysis) and the short version was obtained on the basis of a Bifactoral model and conceptual approaches. Four models were evaluated: orthogonal, oblique, bifactor and one-dimensional. As the result, the TAI-S is constituted by a single dimension. Likewise, the short version shows robust indicators and based on the empirical evidence with the long version. Finally, the two versions showed reliability coefficient with high magnitudes. In conclusion, the two version studied show favorable psychometric properties that support its use in the Peruvian context.

A pesar del tiempo, el estudio de la ansiedad ante los exámenes (AE; Mandler & Sarason, 1952) sigue llamando la atención de los investigadores en diferentes lugares del mundo debido a su importancia dentro del proceso enseñanza-aprendizaje en los distintos niveles educativos, y con especial relevancia en el contexto universitario (Onyeizugbo, 2010). La AE puede definirse como un rasgo de personalidad específico para exhibir estados de ansiedad de forma más intensa y frecuente de lo habitual, con preocupaciones que interfieren con la atención, concentración y realización de exámenes (Spielberger, 1980a; Spielberger & Vagg, 1995).

Desde el enfoque dual (Liebert & Morris, 1967; Spielberger, 1980a), la AE se manifiesta de forma fisiológica (aquí: *Emocionalidad*) y cognitiva (aquí: *Preocupación*). La *Emocionalidad* hace referencia a la percepción de activación afectiva o fisiológica al momento de la evaluación; y la *Preocupación* consiste en pensamientos recurrentes sobre las consecuencias, personales o sociales, de un posible mal desempeño en el examen. Ambos componentes interactúan, mermando progresivamente el desempeño del estudiante en la situación de evaluación (Spielberger, 1980a; Valero, 1999; Furlán, Sánchez, & Heredia, 2009; Sub & Phabha, 2003).

Por otro lado, la AE se asocia consistentemente con un pobre desempeño del estudiante en las evaluaciones (Chapell et al., 2005; Farooqui, Ghani, & Spielberger, 2012; Heredia, Piemontesi, Furlán, & Hodapp, 2008; Sarason & Stoops, 1978; Seipp, 1991; Sub & Phabha, 2003), aunque otras investigaciones solo han encontrado esa asociación con algunos de los componentes de la AE, como la *Interferencia* (Heredia et al., 2008; Piemontesi & Heredia, 2009; Piemontesi, Heredia, & Furlán, 2009). Asimismo, se ha observado una relación directa de la AE con el agotamiento emocional académico (Dominguez-Lara, 2014a), estrategias de aprendizaje poco profundas y memorísticas (Furlán et al., 2009), baja autoeficacia (Dominguez-Lara, Villegas, Cabezas,

Aravena, & De la Cruz, 2013; Onyeizugbo, 2010), pobre auto-concepto académico (Iroegbu, 2013), entre otras.

Existen diversos instrumentos para evaluar de la AE (ver Ali & Mohsin, 2013; Furlán, 2006). La mayoría coincide en la presencia de un factor general (AE) y los dos componentes del modelo dual, mientras que otros instrumentos agregan dimensiones derivadas teóricamente a partir de falencias observadas en algunos instrumentos. Por ejemplo, el componente *Interferencia* es conceptualmente distinto a *Preocupación*, pero ambos conforman un componente cognitivo de mayor jerarquía (Hodapp, 1991; Heredia et al., 2008). Entre los instrumentos disponibles, el *Test Anxiety Inventory* (TAI; Spielberger, 1980b; Spielberger, González, Taylor, Algaize, & Anton, 1978) es uno de los más usados a nivel mundial para la evaluación de la AE (Chapell et al., 2005). El TAI posee una base teórica sólida (Spielberger, 1980a, 1980b; Spielberger & Vagg, 1995), y cuenta con múltiples estudios empíricos e instrumentales en población universitaria de diferentes contextos culturales. Sin embargo, los estudios que utilizaron la versión en español del TAI (Bauermeister, Collazos, & Spielberger, 1983) en el contexto peruano, se centraron en población escolar (p.e., Aliaga, Ponce, Bernaola, & Pecho, 2001) y solo existe un estudio instrumental (Villegas, Dominguez-Lara, Sotelo, & Sotelo, 2015), pero realizado bajo el marco del análisis factorial exploratorio.

La mayoría de los estudios instrumentales llevados a cabo con el TAI reportan la presencia de un factor general (AE) y dos específicos, *Emocionalidad* y *Preocupación*, interpretables de forma independiente (Ali & Mohsin, 2013; Bauermeister et al., 1983; Papantoniou, Moraitou, & Filippidou, 2011; Spielberger, 1980b; Szafranski, Barrera, & Norton, 2012; Taylor & Deane, 2002; Ware, Galassi, & Dew, 1990). Sin embargo, algunos estudios no apoyan la unidimensionalidad del TAI (Oktedalen & Hagtvet, 2011), y un trabajo relativamente reciente con muestra peruana reporta la presencia de una estructura de dos factores: *Ansiedad durante el*

examen y Ansiedad ante la situación de examen ( $r = .587$ ; Villegas et al., 2015). De estos factores, solo el primero se focaliza en la situación de examen y tiene compatibilidad con el factor general de AE que fue hallado en otras investigaciones (Ali & Mohsin, 2013; Bauermeister et al., 1983; Papantoniou et al., 2011; Spielberger, 1980b).

A partir de esos hallazgos, pueden realizarse reflexiones metodológicas relevantes a la medición de un constructo complejo en sí mismo. Algunos estudios instrumentales (Oktedalen & Hagtvet, 2011; Papantoniou et al., 2011; Ware & Galassi, 1990) no interpretan la elevada correlación interfactorial ( $\phi = .800$ ,  $\phi = .906$ ,  $\phi = .883$ , respectivamente) como indicador de multicolinealidad (Arias, 2008) y fuente potencial de información redundante a causa de la presencia de un factor general más robusto e interpretable, en lugar de dos específicos. De igual forma, en otros estudios fue empleada la rotación ortogonal Varimax (Ali & Mohsin, 2013; Bauermeister et al., 1983; Szafranski et al., 2012), cuyos supuestos base (e.g., factores independientes) enmascaran la relación entre factores y se puede concluir erróneamente la presencia de independencia factorial. Además, en algunos estudios empíricos no fue discutida la elevada correlación interfactorial ( $\phi = .74$ ; Chapell et al., 2005), lo que puede llevar al lector a considerar que, efectivamente, el TAI está compuesto por dos factores específicos (*Emocionalidad* y *Preocupación*) y uno general (AE). Finalmente, independientemente de la consideración de dos factores, no se justifica mediante métodos empíricos la obtención de un puntaje total (Suma de todos los ítems del TAI; o solo de los que pertenecen a *Emocionalidad* y *Preocupación*).

En este sentido, son necesarios análisis adicionales y más especializados para la evaluación de la dimensionalidad del TAI, ya que, si se considera la presencia de dos factores cuando realmente uno es suficiente, se estaría cometiendo un error de especificación (Hu & Bentler, 1998) independientemente de las magnitudes de los índices de ajuste u otro indicador de bondad del modelo (Dominguez-Lara, 2016a).

Ante esta situación, algunos investigadores implementaron estrategias (e.g. modificar algunos ítems) para minimizar el aparente solapamiento entre *Preocupación* y *Emocionalidad* (TAI-G; Hodapp, 1991). A pesar de ello, las correlaciones se mantienen entre moderadas y altas (Harpell & Andrews, 2012; Heredia et al., 2008; Mowbray, Jacobs, & Boyle, 2015; Ringeisen, Buchwald, & Hodapp, 2010). Del mismo modo, el puntaje total fue justificado a través de modelos jerárquicos indirectos (e.g., Keith, Hodapp, Schermelleh-Engel, & Moosbrugger, 2003; Mowbray et al., 2015), aunque este no es el procedimiento más adecuado para evaluar la influencia

directa del constructo de orden superior sobre los ítems (Canivez, 2016). En efecto, el modelamiento bifactor o jerárquico directo (Canivez, 2016; Reise, 2012) es más apropiado para evaluar simultáneamente cuánta varianza de cada ítem del TAI es explicada por el factor general y los factores específicos. De este modo, podría valorarse la viabilidad de un factor general o de dos factores específicos del TAI, como sugiere la literatura antecedente.

Entonces, el primer objetivo del presente reporte fue analizar la estructura factorial de la versión en español del TAI (TAI-E) en universitarios limeños a través de la evaluación de los diferentes modelos de medición disponibles. Esto es importante porque la AE es descrita como el obstáculo más fuerte en contextos educacionales (Enright, Baldo, & Wykes, 2000) y frente al cual los estudiantes necesitan apoyo para manejarlo apropiadamente (Chapell et al., 2005). De este modo, el tema tiene relevancia social, pues en la actualidad es una dificultad significativa que enfrenta los estudiantes universitarios, por lo cual el desarrollo de estrategias de intervención es una prioridad para el mantenimiento de la salud mental y general de dicha población.

Por otro lado, a pesar de la brevedad del TAI-E (20 ítems), su inclusión en baterías de evaluación se puede dificultar por la extensión conjunta de los instrumentos que podrían emplearse, por lo cual es conveniente contar con una versión abreviada (TAI-EB) estadísticamente equivalente la versión extensa. Con ello, puede facilitarse la recopilación de información de diversas variables en el marco de estudios de investigación de mayor envergadura, procesos de tamizaje, o de selección de participantes para programas de intervención al interior de las instituciones universitarias.

Existen disponibles dos versiones breves del TAI para ser utilizadas en estudiantes universitarios que fueron derivadas de procedimientos analíticos diferentes. Para la obtención de una de ellas fue empleado un enfoque factorial (Spielberger et al., 1978), y en la otra uno descriptivo por medio de la correlación ítem-test corregida y correlación bivariada con otra medida de ansiedad ( $\alpha = .87$ ,  $n = 333$ ; Taylor & Deane, 2002). Sobre el primer estudio, el uso de rotación ortogonal utilizada puede impedir una correcta interpretación de los constructos analizados; y con relación al segundo, independientemente de los criterios externos, deben considerarse estrategias internas de validación que permitan concluir acerca de la equivalencia entre ambas versiones, y al no reportarse esa información, no existen evidencias que puedan avalar su uso. Entonces, el segundo objetivo es derivar una versión breve a partir del mejor modelo de medición de la primera parte del presente estudio.

## MÉTODO

### Diseño

Es un estudio instrumental, cuyo objetivo es el estudio de las propiedades psicométricas del TAI-E, así como de la versión breve derivada de este (Ato, López, & Benavente, 2013).

### Participantes

La muestra estuvo formada por 376 estudiantes universitarios (73.1% mujeres) peruanos, quienes cursan del primer al séptimo ciclo de la carrera de psicología de una universidad privada ubicada en Lima Metropolitana. Sus edades estuvieron comprendidas entre los 16 y 55 años ( $M = 20.71$ ,  $DE = 4.89$ ). Todos ellos se encontraban solteros al momento de realizar la encuesta. No se encontraron diferencias estadísticamente significativas entre varones y mujeres en cuanto a la edad ( $t_{(374)} = .031$ ,  $d = .000$ ).

### Instrumento

**Inventario de Ansiedad ante Exámenes (TAI-E; Bauermeister, Collazos, & Spielberger, 1983).** Fue utilizada la versión adaptada a universitarios limeños (Villegas et al., 2015) que consta de 19 ítems con cuatro alternativas de respuesta (*Casi nunca, Algunas veces, Frecuentemente y Casi siempre*). Los ítems se presentan en el mismo orden que el original (Spielberger, 1980b), excepto por el primero que fue eliminado en un estudio previo (Villegas et al., 2015). El TAI-E evalúa las dos dimensiones descritas en el modelo dual de la AE: *Emocionalidad* (ítems originales 2, 8, 9, 10, 11, 15, 16, y 18) y *Preocupación* (ítems originales 3, 4, 5, 6, 7, 14, 17, y 20).

### Procedimiento

El TAI-E fue aplicado a los estudiantes universitarios en el horario habitual de clase. De forma preliminar a la evaluación, se explicó que la evaluación formaba parte de una investigación y que la participación era voluntaria a los estudiantes. Ninguno de los estudiantes escogió no participar, y quienes lo hicieron firmaron un consentimiento informado. Asimismo, no fueron compensados económica ni académicamente. Al término, se les agradeció por su colaboración.

**Análisis estructural.** La estructura interna del TAI-E fue analizada mediante el modelamiento de ecuaciones estructurales (análisis factorial confirmatorio) utilizando el programa EQS 6.2 (Bentler & Wu, 2012) bajo las siguientes condiciones: método de máxima verosimilitud (Beauducel & Herzberg, 2006) y matrices policóricas (Lee, Poon, & Bentler,

1995) debido a que el escalamiento Likert de los ítems los ubica como medidas de carácter ordinal (Dominguez-Lara, 2014b).

Fueron evaluados los modelos de medición predominantes en la literatura. El primer modelo (M1) fue el de dos factores ortogonales (Bauermeister et al., 1983; Szafranski et al., 2012). La ortogonalidad se infiere a partir de la estrategia analítica utilizada (rotación Varimax) en estudios previos. Se considera la presencia de dos factores, *Emocionalidad* y *Preocupación*, pero independientes. Los dos siguientes modelos fueron de dos factores oblicuos. Por un lado, se modeló la correlación interfactorial entre *Emocionalidad* y *Preocupación* (M2a; Papantoniou et al., 2011), y posteriormente se evaluó otro modelo de dos factores relacionados (M2b; Villegas et al., 2015).

A continuación, se modeló una estructura bifactor (M3), que evalúa cuánta varianza de los ítems es explicada por el factor general (AE) los factores específicos (*Emocionalidad* y *Preocupación*). Finalmente, fue evaluada una estructura unidimensional (M4) donde todos los ítems son influidos por la misma variable latente (Bauermeister et al., 1983; Papantoniou et al., 2011; Spielberger, 1980b; Szafranski et al., 2012; Taylor & Deane, 2002).

Para la evaluación del ajuste de los modelos evaluados fueron utilizados diversos índices: RMSEA ( $\leq .05$ ), CFI ( $\geq .95$ ), SRMR ( $\leq .05$ ) y la prueba general  $X^2$  corregida por la falta de normalidad de las variables ( $SB-\chi^2$ ; Satorra & Bentler, 1994).

Por otro lado, la valoración del modelo bifactor fueron utilizados fue realizada tanto con los índices de ajuste, como con indicadores que proveen información sobre la fortaleza del factor general, como el coeficiente omega jerárquico ( $\omega_h$ ; Zinbarg, Yovel, Revelle, & McDonald, 2006), que indica el monto de varianza total que puede ser atribuida al factor general; el  $H_h$  (Raykov & Hancock, 2005), que brinda la confiabilidad máxima de los factores específicos y del general, esperando magnitudes de .70 o superiores; y el ECV (*Explained Common Variance*; Reise, Scheines, Widaman & Haviland, 2013) que da información sobre el monto de varianza común que es explicada por el factor general. Se espera un ECV elevado ( $> .60$ ) para concluir que hay poca varianza común entre factores más allá que la explicada por el factor general. También fue implementado el ECV-I (Stucky, Thisen, & Edelen, 2013) que brinda información sobre qué porcentaje de la varianza verdadera de cada ítem es explicada por el factor general. Los cálculos fueron ejecutados con el módulo *IndicesBifactor.xls* (Dominguez-Lara & Rodriguez, 2017).

**Validez interna.** Fueron obtenidas evidencias internas de

validez convergente a partir del análisis de la varianza promedio extraída del factor (AVE; *Average Variance Extracted*) considerando cantidades por encima de .50 como satisfactorias (Fornell & Larcker, 1981).

**Análisis de la confiabilidad.** Con el mejor modelo, fueron analizadas la tau-equivalencia y el impacto de la presencia de errores correlacionados (Dunn, Baguley, & Brunsten, 2014). En cuanto a la evaluación del modelo tau-equivalente ( $M_T$ ), este se realizó sobre la base del modelo congénico ( $M_C$ ) forzando la igualdad de las cargas factoriales, comparándolo con el modelo previo ( $M_C$ ) considerando variaciones en el CFI ( $CFI_T - CFI_C \leq -.01$ ), RMSEA ( $RMSEA_T - RMSEA_C \geq .015$ ) y SRMR ( $SRMR_T - SRMR_C \geq .030$ ) (Chen, 2007). Luego de ello, con base en  $M_C$ , se exploró la pertinencia de modelar errores correlacionados a través de los Índices Modificadores de Lagrange (IM; Sörbom, 1989) con el objetivo de evaluar si su presencia incrementa espuriamente los índices de ajuste (Dominguez-Lara, 2016b; Dunn et al., 2014).

Asimismo, se calculó la confiabilidad de los puntajes observados mediante el coeficiente  $\alpha$  (Cronbach, 1951) con intervalos de confianza (IC; Dominguez-Lara & Merino-Soto, 2015), y en las variables latentes por medio del coeficiente  $\omega$  (McDonald, 1999) y  $H$  (Dominguez-Lara, 2016c; Hancock & Mueller, 2001). Por último, fue realizada una corrección por la presencia de errores correlacionados a los coeficientes  $\alpha$  ( $\alpha_{\text{corregido}}$ ; Pascual-Ferrá & Beatty, 2015) y  $\omega$  ( $\omega_{\text{corregido}}$ ; Raykov, 2001). Para el coeficiente  $\alpha$  se esperan magnitudes elevadas ( $> .80$ ; Merino, Navarro, & García, 2014), así como para los coeficientes  $\omega$  y  $H$  (Raykov & Hancock, 2005).

**Versión breve: selección de los ítems, análisis estructural y de confiabilidad.** A partir del modelo bifactor fueron considerados los ítems que presentaron mayor carga en el factor general ( $\lambda > .70$ ), y menos con su factor específico ( $\lambda < .20$ ), a fin de garantizar una versión breve y unidimensional. La estructura interna y confiabilidad de esta nueva versión breve (TAI-EB) fue evaluada con los procedimientos ejecutados con la versión extensa (TAI-E).

**Versión breve: Equivalencia empírica entre TAI-E y TAI-EB.** La equivalencia empírica entre versiones se evaluó por medio de una correlación bivariada. Se hipotetizó una magnitud elevada entre versiones, ya que la extensa (TAI-E) contiene a la breve (TAI-EB). Sin embargo, para evitar el efecto de la varianza debido a elementos comunes y evitar obtener coeficientes de correlación espuriamente altos, esta asociación fue corregida (Levy, 1967). Este método es importante porque permite tener un mayor respaldo respecto a la versión breve evaluada (Dominguez-Lara, Merino-Soto, & Navarro-Loli, 2016; Merino-Soto & Ángulo-Ramos, 2013), ya que en ocasiones se puede concluir favorablemen-

te sin considerar la influencia de los elementos en común.

Finalmente, en cuanto a la equivalencia en la estimación del error de medición (Merino & Lautenschlager, 2003) entre versiones, fueron calculados los coeficientes para cada versión, incluyendo los IC, y se efectuó una comparación estadística mediante el método asintótico basado en la distribución  $F$  (Feldt, 1980, Feldt, Woodruff, & Salih, 1987) utilizando el programa *AlphaTest* (Lautenschlager & Meade, 2008). Como parte del cálculo ejecutado con el *AlphaTest* fue utilizada la correlación corregida por errores correlacionados (Levy, 1967) mencionada en el párrafo anterior. Asimismo, fue comparada la confiabilidad obtenida en el presente estudio con la encontrada en otro estudio (Taylor & Deane, 2002).

## RESULTADOS

### Análisis descriptivo

De acuerdo con los resultados, se observa una marcada tendencia hacia las puntuaciones más bajas en cada ítem (efecto piso; Terwee et al., 2007). A su vez, la mayor parte de los ítems presentan índices estandarizados de asimetría (*Standardized Skew Index*; SSI; Malgady, 2007) de magnitud moderada a alta ( $SSI \geq .25$ ), lo que indica la pertinencia de utilizar procedimientos que tengan en cuenta las características distribucionales de cada ítem (p.e.  $SB-\chi^2$ ).

### Análisis estructural del TAI-E y evidencias internas de validez.

El modelo base fue el de dos factores ortogonales (M1), que presentó índices adecuados:  $SB-\chi^2_{(120)} = 428.362$  ( $p < .001$ ); CFI = .971; RMSEA (IC 90%) = .091 (.082, .100); SRMR = .352; aunque los dos últimos índices brindan información que desacredita dicho modelo. Luego de ello fueron evaluados dos modelos de factores oblicuos (M2a y M2b; Tabla 2) cuyos índices de ajuste fueron superiores a M1 (M2a:  $SB-\chi^2_{(103)} = 298.277$  [ $p < .001$ ]; CFI = .982; RMSEA [IC 90%] = .071 [.062, .080]; SRMR = .057; M2b:  $SB-\chi^2_{(151)} = 353.744$  [ $p < .001$ ]; CFI = .987; RMSEA [IC 90%] = .060 [.052, .068]; SRMR = .052). No obstante, la elevada magnitud de la correlación interfactorial ( $\varphi_{ij} \geq .90$ , ver Tabla 2) en ambos modelos sugiere la presencia de un factor general.

Posteriormente fue analizado un modelo bifactor (M3), el cual presentó índices de ajuste superiores a los dos modelos previos:  $SB-\chi^2_{(88)} = 158.214$  ( $p < .001$ ); CFI = .994; RMSEA (IC 90%) = .046 (.034, .057); SRMR = .047. Además, de los índices de ajuste presentados, los  $\omega_n$  y  $H_n$  del factor general son ampliamente superiores comparado con los factores espe-

cíficos (Tabla 2), y de forma conjunta con la magnitud del ECV del factor general (.873) y los ítems ( $ECV-I_{promedio} = .891$ ) apoyan la premisa de que el factor general (AE) es más robusto que los específicos (Rodríguez, Reise, & Haviland, 2016).

Finalmente, a fin de brindar más evidencias a favor de la unidimensionalidad del TAI-E, fue evaluado el ajuste de un modelo unifactorial (M4), obteniéndose índices de magnitudes aceptables:  $SB-\chi^2_{(152)} = 354.192$  ( $p < .001$ ); CFI = .987; RMSEA (IC 90%) = .060 (.051, .068); SRMR = .052. En conclusión, el TAI-E está mejor representado por una estructura unidimensional, la cual presenta evidencias favorables de validez interna convergente ( $AVE > .50$ ).

**Confiabilidad**

En cuanto al análisis de modelos de medición, sobre la base del modelo congénico (M4) fue especificada la igualdad estadística de cargas factoriales (modelo de medición tau-equivalente;  $M_T$ ). El  $M_T$  mostró un ajuste adecuado ( $SB-\chi^2_{(152)} = 397.896$  [ $p < .001$ ]; CFI = .986; RMSEA [IC 90%] = .060 [.052, .067]; SRMR = .073), y las variaciones en el CFI ( $\Delta CFI = -.001$ ),

RMSEA ( $\Delta RMSEA = .000$ ) y SRMR ( $\Delta SRMR = .021$ ) no fueron sustanciales tomando como referencia al congénico ( $M_C$ ), por lo que las cargas factoriales son estadísticamente similares, por lo que es pertinente el uso del coeficiente  $\alpha$  (Dunn et al., 2014).

Posteriormente, por medio de los IM se obtuvo información sobre la necesidad de modelar errores correlacionados debido a su contribución al incremento de los índices de ajuste. En este sentido, fueron incluidas gradualmente las correlaciones entre los errores de los ítems 13 (*Me parece que estoy en contra de mí mismo cuando contesto los exámenes*) y 14 (*Me coge fuerte el pánico cuando rindo un examen importante*), mostrando un mejor ajuste que el M4 ( $SB-\chi^2_{(151)} = 315.270$  [ $p < .001$ ]; CFI = .990; RMSEA [IC 90%] = .054 [.045, .062]; SRMR = .051), pero sin llegar a superarlo de forma sustancial. Del mismo modo, sobre esa base se modeló la correlación entre los errores de los ítems dos (*El pensar en la calificación que pueda obtener en un curso interfiere con mi trabajo en los exámenes*) y seis (*El pensar que puede salir mal, interfiere con mi concentración en los exámenes*), incrementándose los índices de ajuste, pero no de forma

**Tabla 1**  
*Estadísticos descriptivos de los ítems del TAI-E*

	M	DE	$g_1$	SSI	$g_2$	%Mín	%Máx
Item 1 (2)	2.056	.832	.843	.608	.496	22.9	9.0
Item 2 (3)	1.963	.876	.696	.454	-.149	33.2	6.9
Item 3 (4)	1.753	.903	1.097	.672	.385	48.9	7.2
Item 4 (5)	1.537	.792	1.445	1.150	1.433	61.7	3.5
Item 5 (6)	1.726	.875	1.140	.745	.605	49.2	6.4
Item 6 (7)	1.920	.903	.792	.486	-.112	37.2	7.7
Item 7 (8)	1.561	.777	1.250	1.035	.804	59.3	2.4
Item 8 (9)	1.928	.838	.684	.487	-.051	33.5	5.3
Item 9 (10)	2.340	.924	.312	.183	-.712	17.8	13.6
Item 10 (11)	2.037	.894	.535	.335	-.474	30.9	7.2
Item 11 (12)	2.362	1.174	.241	.087	-1.427	31.1	26.1
Item 12 (13)	1.622	.907	1.401	.852	.995	59.8	7.2
Item 13 (14)	1.394	.688	1.866	1.969	3.248	70.2	2.1
Item 14 (15)	1.620	.834	1.226	.882	.702	56.9	4.0
Item 15 (16)	2.335	1.027	.272	.129	-1.051	23.9	17.6
Item 16 (17)	2.005	.982	.719	.372	-.487	36.4	11.4
Item 17 (18)	1.702	.886	1.083	.689	.238	53.2	5.3
Item 18 (19)	1.886	.886	.782	.498	-.132	39.1	6.4
Item 19 (20)	1.989	.924	.735	.431	-.236	33.8	9.3

*Nota:*  $n = 376$ ; Entre paréntesis la numeración original; *M*: media aritmética. *DE*: desviación estándar;  $g_1$ : asimetría de Fisher;  $g_2$ : curtosis de Fisher. *SSI*: Índice estandarizado de Asimetría; %Mín: porcentaje de personas con el puntaje mínimo; %Máx: porcentaje de personas con el puntaje máximo

sustancial ( $SB-\chi^2_{[150]} = 293.777 [p < .001]$ ; CFI = .991; RMSEA [IC 90%] = .051 [.042, .059]; SRMR = .050). Finalmente, fue modelada la correlación entre los errores de los ítems 14 y 17 (Siento que el corazón me late muy rápidamente durante los exámenes muy importantes), incrementándose el ajuste, pero no de forma significativa ( $SB-\chi^2_{[149]} = 278.271 [p < .001]$ ; CFI = .992; RMSEA [IC 90%] = .048 [.039, .057]; SRMR = .049). Las magnitudes de las correlaciones fueron:  $\phi_{13,14} =$

.152 ( $p < .05$ ),  $\phi_{2,6} = .153 (p < .05)$ , y  $\phi_{14,17} = .083 (p < .05)$ . No obstante, la presencia de errores correlacionados aumenta espuriamente los coeficientes  $\alpha$  y  $\omega$  del TAI-E.

En base a estos hallazgos, los coeficientes  $\alpha$  y  $\omega$  fueron corregidos por presencia de errores correlacionados (Tabla 2), disminuyendo su magnitud, pero manteniéndose en un nivel aceptable (Merino et al., 2014).

**Tabla 2**

*Parámetros de los ítems, y confiabilidad en TAI-E (modelos oblicuos, bifactor y unidimensional) y TAE-EB*

	Modelo 2a			Modelo 2b			Modelo Bifactor				Modelo unidimensional		Versión breve	
	F <sub>1</sub>	F <sub>2</sub>	h <sup>2</sup>	F <sub>I</sub>	F <sub>II</sub>	h <sup>2</sup>	F <sub>E</sub>	F <sub>G</sub>	h <sup>2</sup>	ECV-I	F <sub>1</sub>	h <sup>2</sup>	F <sub>1</sub>	h <sup>2</sup>
Item 1 (2)	.709	-	.503	.708	-	.501	.137	.694	.500	.962	.708	.501	-	-
Item 2 (3)	--	.661	.437	.652	-	.425	.307	.615	.473	.801	.651	.424	-	-
Item 3 (4)	-	.768	.589	.761	-	.580	<b>.005</b>	<b>.777</b>	<b>.604</b>	<b>1.000</b>	.761	.580	.786	.617
Item 4 (5)	-	.592	.350	.576	-	.331	-.081	.590	.355	.982	.576	.332	-	-
Item 5 (6)	-	.699	.488	.673	-	.453	.295	.653	.513	.831	.673	.453	-	-
Item 6 (7)	-	.777	.604	.769	-	.592	.663	.725	.965	.545	.769	.591	-	-
Item 7 (8)	.712	-	.508	.698	-	.487	<b>-.250</b>	<b>.720</b>	<b>.581</b>	<b>.892</b>	.698	.487	.743	.551
Item 8 (9)	.694	-	.481	.697	-	.486	.249	.686	.533	.884	.697	.486	-	-
Item 9 (10)	.628	-	.395	.630	-	.397	.495	.630	.642	.618	.630	.397	-	-
Item 10 (11)	.777	-	.604	-	.784	.615	<b>.145</b>	<b>.770</b>	<b>.615</b>	<b>.966</b>	.784	.615	.720	.519
Item 11 (12)	-	-	-	.760	-	.577	-	-	-	-	.759	.577	-	-
Item 12 (13)	-	-	-	-	.684	.468	-	-	-	-	.685	.469	-	-
Item 13 (14)	-	.754	.569	.746	-	.557	-.086	.773	.606	.988	.747	.558	-	-
Item 14 (15)	.820	-	.673	.808	-	.654	<b>-.176</b>	<b>.839</b>	<b>.736</b>	<b>.958</b>	.809	.654	.857	.735
Item 15 (16)	.685	-	.469	.684	-	.467	.108	.680	.475	.975	.684	.467	-	-
Item 16 (17)	-	.748	.560	-	.742	.550	.112	.728	.543	.977	.742	.550	-	-
Item 17 (18)	.767	-	.588	-	.768	.589	<b>-.258</b>	<b>.789</b>	<b>.690</b>	<b>.903</b>	.768	.590	.841	.707
Item 18 (19)	-	-	-	.750	-	.563	-	-	-	-	.750	.563	-	-
Item 19 (20)	-	.715	.511	-	.704	.496	.121	.687	.487	.970	.704	.496	-	-
F1	1			1										
F2	.973	1		1.00	1									
AVE	.527	.514	-	.505	.544	-	-	-	-	-	.515	-	.626	-
$\omega$	.899	.894	-	.934	.856	-	-	-	-	-	.953	-	.893	-
$\omega_h$	.005	.050	-	-	-	-	-	.937	-	-	-	-	-	-
H	.905	.897	-	.938	.860	-	-	-	-	-	.955	-	.901	-
H <sub>b</sub>	.380	.506	-	-	-	-	-	.947	-	-	-	-	-	-
$\alpha$	.861	.854	-	.906	.801	-	-	-	-	-	.933	-	.844	-
IC95% Inf	.832	.824	-	.886	.762	-	-	-	-	-	.919	-	.812	-
IC95% Sup	.885	.879	-	.923	.835	-	-	-	-	-	.945	-	.871	-
$\alpha_{\text{corregido}}$	-	-	-	-	-	-	-	-	-	-	.866	-	.813	-
$\omega_{\text{corregido}}$	-	-	-	-	-	-	-	-	-	-	.951	-	.880	-

**Nota:** n = 376; Entre paréntesis la numeración original; F<sub>1</sub>: Emocionalidad; F<sub>2</sub>: Preocupación; F<sub>I</sub>: Ansiedad durante el examen; F<sub>II</sub>: Ansiedad ante la situación de examen; F<sub>E</sub>: factor específico; F<sub>G</sub>: Factor general; ECV-I: ECV de los ítems; AVE: Varianza promedio extraída;  $\omega$ : Coeficiente omega;  $\omega_h$ : Coeficiente omega jerárquico; H: coeficiente H; H<sub>b</sub>: coeficiente H Bifactor; En negrita: ítems seleccionados para la versión breve;  $\alpha_{\text{corregido}}$  y  $\omega_{\text{corregido}}$ : coeficientes corregidos por presencia de errores correlacionados.

### **Análisis estructural y de confiabilidad del TAI-EB**

A partir de los resultados obtenidos con el modelo bifactor (M3), fueron considerados para conformar la forma breve los cinco ítems con mayor carga en el factor general ( $\lambda > .70$ ), y menos en el factor específico que les correspondía ( $\lambda < .20$ ) (ítems 3, 7, 10, 14 y 17, de la nueva numeración).

El modelo unidimensional mostró índices de ajuste bastante buenos:  $SB-\chi^2_{(152)} = 3.202$  ( $p = .669$ ); CFI = 1.000; RMSEA (IC 90%) = .000 (.000, .056); SRMR = .015. además de validez interna convergente (AVE > .60). Sobre esta base, se modeló la igualdad estadística de las cargas factoriales ( $M_7$ ), mostrando un ajuste adecuado ( $SB-\chi^2_{(152)} = 15.698$  [ $p = .073$ ]; CFI = .996; RMSEA [IC 90%] = .045 [.000, .080]; SRMR = .047), así como variaciones no significativas respecto al modelo previo con relación al CFI ( $\Delta CFI = -.004$ ), RMSEA ( $\Delta RMSEA = -.045$ ) y SRMR ( $\Delta SRMR = .035$ ).

En cuanto a la confiabilidad, en vista del cumplimiento del supuesto de tau-equivalencia es adecuado calcular el coeficiente  $\alpha$  como estimador de la confiabilidad de los puntajes; y los coeficientes  $\omega$  y  $H$ , del constructo. Con los tres coeficientes fueron obtenidas magnitudes aceptables (Tabla 2).

Fue analizada la pertinencia de correlacionar errores tomando como referencia los IM. En base a los resultados, es conveniente relacionar los errores de los ítems 7 (*Me siento muy agitado mientras contesto un examen importante*) y 17 (*Siento que el corazón me late muy rápidamente durante los exámenes muy importantes*). La presencia de la correlación entre errores no mejora sustancialmente los índices de ajuste ( $SB-\chi^2_{(151)} = 1.864$  [ $p = .761$ ]; CFI = 1.000; RMSEA [IC 90%] = .000 [.000, .053]; SRMR = .011), aunque debido a su magnitud ( $\phi_{7,17} = .126$ ;  $p < .05$ ) fue considerada para efectuar las correcciones a los coeficientes  $\alpha$  y  $\omega$  (Tabla 2).

### **Equivalencia entre TAI-E y TAI-EB**

La correlación bivariada entre el TAI-E y TAI-EB fue  $r = .901$  (IC95% .880, .918;  $p < .001$ ), magnitud que puede considerarse adecuada entre versiones equivalente de un instrumento (Petrides, Jackson, Furnham & Levine, 2003), y la correlación corregida por espuriedad fue de .855 (IC95% .825, .880;  $p < .001$ ), magnitud que puede considerarse elevada. La estimación de la confiabilidad de los puntajes y variables latentes brinda indicadores favorables ( $> .80$  en todos los coeficientes).

Existieron diferencias significativas entre los coeficientes  $\alpha$  de ambas versiones ( $\chi^2_{(1)} = 169.135$ ;  $p < .001$ ), es decir, el error de medición es menor se incrementa al reducir la cantidad de ítems. Finalmente, no se evidenciaron diferencias

estadísticas ( $\chi^2_{(1)} = 1.938$ ;  $p = .164$ ) con respecto a otra versión breve disponible (Taylor & Deane, 2002).

## **DISCUSIÓN**

El objetivo inicial fue realizar un análisis estructural del TAI-E en universitarios peruanos desde un enfoque confirmatorio debido a que la información no permite concluir si se trata de una estructura de dos factores o de un solo factor. Inicialmente fue evaluada la ortogonalidad de los constructos *Emocionalidad* y *Preocupación*, pero los valores elevados en el RMSEA ( $> .080$ ) y el SRMR ( $> .100$ ) indican falta de parsimonia y ajuste pésimo (Brown, 2006; Chen, 2007). Esto implica que la relación entre factores no puede ser omitida en el modelo analizado, ya que de ese modo se comete una mala especificación (Hu & Bentler, 1998) y podría perderse información importante que facilitaría una interpretación más certera de la estructura del TAI-E. Por otro lado, las elevadas correlaciones interfactoriales observadas tanto en estudios precedentes (Chapell et al., 2005; Oktedalen & Hagtvet, 2011; Papantoniou et al., 2011; Ware et al., 1990) como en los modelos oblicuos analizados en el presente estudio revelan que *Emocionalidad* y *Preocupación* no pueden ser considerados factores independientes, es decir, resulta difícil diferenciar ambos componentes en una situación de evaluación. Este parece ser una característica no solo del TAI, ya que otros instrumentos que tienen como sustento el modelo dual también lo presentan (ver Ferrando, Varea, & Lorenzo, 1999) donde los componentes de *Preocupación* y *Emocionalidad* no se diferencian, o cuya correlación es de magnitud elevada (Harpell & Andrews, 2012; Heredia et al., 2008; Mowbray et al., 2015; Ringeisen et al., 2010); aunque también pudo influir otros factores, como el efecto piso, derivando en un factor de dificultad (Dominguez-Lara, 2013). Por ello, y con el respaldo del modelo bifactor analizado, un solo factor explica mejor la variabilidad de los ítems.

Considerar el TAI-E como unidimensional parece ser la opción más parsimoniosa, sin embargo si se desea lograr la diferenciación teórica postulada inicialmente (*Emocionalidad* y *Preocupación*; Spielberger, 1980b) o multidimensional (Hodapp, 1991) deben realizarse cambios profundos en la estructura del instrumento.

El reporte de confiabilidad es favorable tanto al analizar puntajes observados como variables latentes, inclusive al corregir dichos coeficientes ante la presencia de errores correlacionados. Esto es relevante considerando la potencial utilidad que pueda tener el TAI-E en contextos aplicados



(e.g., evaluación e intervención), dado que su magnitud supera ampliamente lo recomendado ( $> .90$ ; Merino et al., 2014), así como en investigación, con el objetivo de conocer la red conceptual que permitan comprender la AE en un contexto cultural distinto al que dio origen al TAI.

Con relación al segundo objetivo, el TAI-EB cuenta con respaldo empírico tanto con respecto a su estructura interna (estructura factorial robusta y evidencias internas de validez convergente), como en su relación con la versión extensa (TAI-E). Asimismo, la confiabilidad general (puntajes y variables latentes) no deja de ser meritoria ( $> .80$ ), y aunque estime con menor precisión la AE en comparación con el TAI-E, la reducción parece ser trivial ( $-9.54\%$ ).

La presente investigación tuvo algunas limitaciones. Es conveniente destacar que la muestra estuvo formada por estudiantes de psicología, y por ello es probable que los resultados estén sesgados a esa carrera. Asimismo, en Perú la carrera de psicología posee una población predominantemente femenina, por lo cual es conveniente ampliar la muestra a otros grupos con el fin de replicar estos hallazgos. Por otro lado, a pesar que el TAI-E enfoca el constructo AE como un rasgo específico situacional, es necesario destacar que es probable que los estudiantes no muestren el mismo patrón de respuesta ansiosa en las evaluaciones de todos los cursos (Dominguez-Lara, Calderón-De la Cruz, Alarcón-Parco & Navarro-Loli, en prensa), por lo que sería conveniente focalizar la evaluación a determinadas asignaturas que son las que más ansiedad podrían generar en los estudiantes (e.g., estadística).

En cuanto a las recomendaciones, y teniendo en consideración la robustez del TAI-E, podrían estudiarse modelos explicativos del rendimiento en exámenes teniendo como base la literatura que aboga por la influencia que ejerce la AE, así como otros aspectos de personalidad (e.g., autoeficacia académica), sobre todo en los estudiantes de primer año, para quienes el proceso de adaptación suele conllevar más dificultades. En este manuscrito no fueron abordadas las diferencias entre sexos ya que su estudio amerita un análisis de invarianza (Dimitrov, 2010), el cual supera los objetivos de la presente investigación debido, fundamentalmente, a la diferencia de proporciones entre varones y mujeres. No obstante, es necesario llevarlo a cabo a fin de lograr una mejor comprensión de la AE teniendo como referencia el sexo de la persona. Del mismo modo, es importante elaborar baremos para determinar cifras de prevalencia y tener un panorama real de la magnitud del problema.

## CONFLICTO DE INTERESES

Los autores expresan que no hubo conflictos de interés al redactar el manuscrito.

## REFERENCIAS

- Ali, M.S., & Mohsin, M.N. (2013). Test Anxiety Inventory (TAI): factor analysis and psychometric properties. *Journal of Humanities and Social Sciences*, 8(1), 73-81. doi: 10.9790/0837-0817381
- Aliaga, J., Ponce, C., Bernaola, E., & Pecho, J. (2001). Características psicométricas del inventario de autoevaluación de la ansiedad ante exámenes (IDASE). *Paradigmas, Revista de Actualización Profesional*, 2(3), 11-29.
- Arias, B. (2008). *Desarrollo de un ejemplo de análisis factorial confirmatorio con Lisrel, AMOS y SAS. Seminario de Actualización en Investigación sobre Discapacidad SAID 2008*. Universidad de Valladolid: España.
- Ato, M., López, J., & Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en psicología. *Anales de Psicología*, 29(3), 1038-1059. doi: 10.6018/analesps.29.3.178511
- Bauermeister, J., Collazos, J., & Spielberger, C. (1983). The construction and validation of the Spanish form of the Test Anxiety Inventory: Inventario de Autoevaluación sobre Exámenes (IDASE). En: Spielberger, C. & Diaz-Guerrero (Eds.). *Cross-cultural Anxiety* (pp. 67-85). Washintong: McGraw-Hill.
- Beaducel, A., & Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling*, 13, 186-203. doi: 10.1207/s15328007sem1302\_2
- Bentler, P. M., & Wu, E. J. C. (2012). *EQS 6.2 for windows* [Statistical Program]. Encino, CA: Multivariate Software, Inc.
- Brown, T. (2006). *Confirmatory Factor Analysis for Applied Research*. New York: The Guilford Press.
- Canivez, G. L. (2016). Bifactor modeling in construct validation of multifaceted tests: Implications for multidimensionality and test interpretation. In K. Schweizer & C. DiStefano (Eds.), *Principles and methods of test construction: Standards and recent advancements* (pp. 247-271). Gottingen, Germany: Hogrefe.
- Chapell, M. S., Blanding, Z. B., Silverstein, M. E., Takahashi, M., Newman, B., Gubi, A., & McCann, N. (2005). Test anxiety and academic performance in undergraduate and graduate students. *Journal of Educational Psychology*, 97(2), 268-274. doi: 10.1037/0022-0663.97.2.268
- Chen, F.F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14, 464-504. doi: 10.1080/10705510701301834
- Cronbach L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334. doi: 10.1007/BF02310555
- Dimitrov, D. M. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development*, 43(2), 121-149. doi: 10.1177/0748175610373459
- Dominguez-Lara, S. (2013). ¿Ítems Politémicos o Dicotómicos? Un estudio empírico con una escala unidimensional. *Revista Argentina de Ciencias del Comportamiento*, 5(3), 30-37.
- 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. (2016a). Evaluación de modelos estructurales, más allá de los índices de ajuste. *Enfermería Intensiva*, 27, 84-85. doi: 10.1016/j.enfi.2016.03.003
- Dominguez-Lara, S. (2016b). Errores correlacionados y estimación de la

- fiabilidad en estudios de validación. Comentarios al trabajo validación de la escala Ehealth Literacy (EHEALS) en población universitaria española. *Revista Española de Salud Pública*, 9, e1–e2.
- Dominguez-Lara, S. (2016c). Evaluación de la confiabilidad del constructo mediante el Coeficiente H: breve revisión conceptual y aplicaciones. *Psychologia: Avances en la disciplina*, 10(2), 87-94. doi: 10.21500/19002386.2134
- Dominguez-Lara, S., & Merino-Soto, C. (2015). ¿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., & Rodriguez, A. (2017). Índices estadísticos de modelos bifactor. *Interacciones*. Publicación anticipada en línea. doi: 10.24016/2017.v3n2.51
- Dominguez-Lara, S., Merino-Soto, C., & Navarro-Loli, J. (2016). Re-análisis de la validez de constructo de la Escala Breve para la Evaluación de la Capacidad Funcional (Sp-UPSA-Brief) de García-Portilla y colaboradores. *Revista de Psiquiatría y Salud Mental*, 9, 127-128. doi: 10.1016/j.rpsm.2015.11.001
- Dominguez-Lara, S., Calderón-De la Cruz, G., Alarcón-Parco, D., & Navarro-Loli, J. (en prensa). Relación entre Ansiedad ante exámenes y rendimiento en exámenes en universitarios: análisis preliminar de la diferencia según asignatura. *Revista Digital de Investigación en Docencia Universitaria*.
- Dominguez-Lara, S., Villegas, G., Cabezas, M. Aravena, S., & De la Cruz, M. (2013). Autoeficacia académica y ansiedad ante los exámenes en estudiantes de psicología de una universidad privada. *Revista de Psicología-Universidad Católica San Pablo*, 3(3), 13-23.
- Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, 105(3), 399-412. doi: 10.1111/bjop.12046
- Enright, M., Baldo, T. D., & Wykes, S. D. (2000). The efficacy of eye movement desensitization and reprocessing therapy technique in the treatment of test anxiety of college students. *Journal of College Counseling*, 3(1), 36-48. doi: 10.1002/j.2161-1882.2000.tb00162.x
- Farooqui, Y.N., Ghani, R., & Spielberger, C.D. (2012). Gender differences in test anxiety and academic performance of medical students. *International Journal of Psychology and Behavioral Sciences*, 2(2), 38-43. doi: 10.5923/j.ijpbs.20120202.06
- Feldt, L. S. (1980). A test of the hypothesis that Cronbach's alpha reliability coefficient is the same for two tests administered to the same sample. *Psychometrika*, 45, 99-105. doi: 10.1007/BF02293600
- Feldt, L. S., Woodruff, D. J., & Salih, F. A. (1987). Statistical inference for coefficient alpha. *Applied Psychological Measurement*, 11, 93-103. doi: 10.1177/014662168701100107
- Ferrando, P. J., Varea, M. D., & Lorenzo, U. (1999). Evaluación psicométrica del cuestionario de ansiedad y rendimiento (CAR) en una muestra de escolares. *Psicothema*, 11(1), 225-236.
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18, 39-50. doi: 10.2307/3151312
- Furlán, L. (2006). Ansiedad ante los exámenes. Qué se evalúa y como? *Evaluar*, 6, 32-51.
- Furlán, L., Sánchez, J., & Heredia, D. (2009). Estrategias de aprendizaje y ansiedad ante los exámenes en estudiantes universitarios. *Pensamiento Psicológico*, 5(12), 117-124.
- Hancock, G. R., & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. En R. Cudeck, S. H. C. du Toit & D. Sörbom (Eds.), *Structural equation modeling: Past and present. A Festschrift in honor of Karl G. Jöreskog* (pp. 195–261). Chicago: Scientific Software International.
- Harpell, J. V., & Andrews, J. J. (2012). Multi-informant test anxiety assessment of adolescents. *Psychology*, 3(7), 518-524. doi: 10.4236/psych.2012.37075
- Heredia, D., Piemontesi, S., Furlán, L., & Hodapp, V. (2008). Adaptación del Inventario Alemán de Ansiedad frente a los Exámenes: GTAI-A. *Evaluar*, 8, 46-60.
- Hodapp, V. (1991). Das Prüfungsängstlichkeitsinventar TAI-G: Eine erweiterte und modifizierte Version mit vier Komponenten [The Test Anxiety Inventory TAI-G: An expanded and modified version with four components]. *Zeitschrift für Pädagogische Psychologie*, 5, 121-130.
- Hu, L., & Bentler, P.M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model specification. *Psychological Methods*, 3, 424-453. doi: 10.1037/1082-989X.3.4.424
- Iroegbu, M. (2013). Effect of test anxiety, gender and perceived self-concept on academic performance of Nigerian students. *International Journal of Psychology and Counselling*, 5(7), 143-146. doi: 10.5897/IJPC2013.0218
- Keith, N., Hodapp, V., Schermelleh-Engel, K., & Moosbrugger, H. (2003). Cross-sectional and longitudinal confirmatory factor models for the German Test Anxiety Inventory: A construct validation. *Anxiety, Stress & Coping*, 16(3), 251-270. doi: 10.1080/1061580031000095416
- Lautenschlager, G. J., & Meade, A. W. (2008). AlphaTest: A windows program for tests of hypotheses about coefficient alpha. *Applied Psychological Measurement*, 23, 502-503. doi: 10.1177/0146621607312307
- 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(2), 339-358. doi: 10.1007/BF02294660
- Levy, P. (1967). The correction for spurious correlation in the evaluation of short-form tests. *Journal of Clinical Psychology*, 23, 84-86.
- Liebert, R. M., & Morris, L. W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychological Reports*, 20(3), 975-978. doi: 10.2466/pr0.1967.20.3.975
- 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
- Mandler, G., & Sarason, S.B. (1952). A study of anxiety and learning. *Journal of Abnormal and Social Psychology*, 47, 166-173.
- McDonald R.P. (1999). *Test theory: A unified treatment*. Mahwah, N.J.: L. Erlbaum Associates.
- Merino-Soto, C., & Angulo-Ramos, M. (2013). Validación de la escala de sobrecarga del cuidador de Zarit en sus versiones original y abreviada: corrección. *Revista Médica de Chile*, 141, 1083-1084. doi: 10.4067/S0034-98872013000800019
- Merino, C., & Lautenschlager, G. (2003). Comparación Estadística de la Confiabilidad Alfa de Cronbach: Aplicaciones en la Medición Educativa y Psicológica. *Revista de Psicología de la Universidad de Chile*, 12(2), 127-136.
- Merino, C., Navarro, J., & García, W. (2014). Revisión de la consistencia interna del Inventario de Inteligencia Emocional de Bar-On, EQ-I: YV. *Revista Peruana de Psicología y Trabajo Social*, 3(1), 141-154.
- Mowbray, T., Jacobs, K., & Boyle, C. (2015). Validity of the German Test Anxiety Inventory (TAI-G) in an Australian sample. *Australian Journal of Psychology*, 67(2), 121-129. doi: 10.1111/ajpy.12058
- Oktedalen, T., & Hagtvet, K.A. (2011). A revised version of the Norwegian adaptation of the Test Anxiety Inventory in a heterogeneous population. *Scandinavian Journal of Educational Research*, 55(5), 475-487. doi: 10.1080/00313831.2010.537688
- Onyeizugbo, E. U. (2010). Self-efficacy, gender and trait anxiety as moderators of test anxiety. *Electronic Journal of Research in Educational Psychology*, 8(1), 299-312.
- Papantoniou, G., Moraitou, D., & Filippidou, D. (2011). Psychometric properties of the Greek version of the Test Anxiety Inventory. *Psychology*, 2(3), 240-246. doi: 10.4236/psych.2011.23038
- Pascual-Ferrá, P., & Beatty, M.J. (2015). Correcting internal consistency estimates inflated by correlated item errors. *Communication Research Reports*, 32(4), 347-352. doi: 10.1080/08824096.2015.1089858
- Petrides, K. V., Jackson, C. J., Furnham, A., & Levine, S. Z. (2003). Exploring issues of personality measurement and structure through the development of a short form of the Eysenck personality profiler. *Journal of Personality Assessment*, 81, 271-280. doi: 10.1207/S15327752JPA8103\_10
- Piemontesi, S., & Heredia, D. (2009). Relaciones entre la ansiedad ante exámenes, estrategias de afrontamiento, autoeficacia para el aprendizaje autorregulado y rendimiento académico. *Revista Tesis*, 2, 74-86.

- Piemontesi, S., Heredia, D., & Furlán, L. (2009). Correlatos de la ansiedad ante los exámenes: una aproximación a la teoría de la reducción en la eficiencia. *Revista Tesis*, 2, 74-86.
- Raykov, T. (2001). Bias in coefficient alpha for fixed congeneric measures with correlated errors. *Applied Psychological Measurement*, 25(1), 69-76. doi: 10.1177/01466216010251005
- Raykov, T., & Hancock, G. R. (2005). Examining change in maximal reliability for multiple-component measuring instruments. *British Journal of Mathematical and Statistical Psychology*, 58(1), 65-82. doi: 10.1348/000711005X38753
- Reise, S.P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, 47(5), 667-696. doi: 10.1080/00273171.2012.715555
- Reise, S.P., Scheines, R., Widaman, K.F., & Haviland, M.G. (2013). Multidimensionality and structural coefficient bias in structural equation modeling: A bifactor perspective. *Educational and Psychological Measurement*, 73(1), 5-26. doi: 10.1177/0013164412449831
- Ringeisen, T., Buchwald, P., & Hodapp, V. (2010). Capturing the multidimensionality of test anxiety in cross-cultural research: An English adaptation of the German Test Anxiety Inventory. *Cognition, Brain, Behavior. an Interdisciplinary Journal*, 14(4), 347-364.
- Rodriguez, A., Reise, S.P., & Haviland, M.G. (2016). Evaluating bifactor models: calculating and interpreting statistical indices. *Psychological Methods*, 21(2), 137-150. doi: 10.1037/met0000045
- Sarason, I. G., & Stoops, R. (1978). Test anxiety and the passage of time. *Journal of Consulting and Clinical Psychology*, 46(1), 102-109.
- 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.
- Seipp, B. (1991). Anxiety and academic performance: a meta-analysis of finding. *Anxiety Research*, 4(1), 27-41. doi: 10.1080/08917779108248762
- Sörbom, D. (1989). Model modification. *Psychometrika*, 54(3), 371-384. doi: 10.1007/BF02294623
- Spielberger, C.D. (1980a). *Tensión y Ansiedad*. México: Harla.
- Spielberger, C.D. (1980b). *Test Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.
- Spielberger, C.D., González, H., Taylor, C., Algaze, B., & Anton, W. (1978). Examination stress and test anxiety. In: C. Spielberger & I. Sarason (Eds.), *Stress and Anxiety*. Washintong: Hemisphere.
- Spielberger, C.D., & Vagg, P. (1995). Test Anxiety. A transactional process. En: C. Spielberger & P. Vagg (Eds.). *Test Anxiety: Theory, assessment and treatment* (pp. 3-14). Washintong: Taylor & Francis.
- Stucky, B.D., Thissen, D., & Edelen, M.O. (2013). Using logistic approximations of marginal trace lines to develop short assessments. *Applied Psychological Measurement*, 37(1), 41-57. doi: 10.1177/0146621612462759
- Szafranski, D.D., Barrera, T.L., & Norton, P.J. (2012). Test anxiety inventory: 30 years later. *Anxiety, Stress, & Coping: An international Journal*, 25(6), 667-677. doi: 10.1080/10615806.2012.663490
- Sub, A., & Phabha, C. (2003). Academic performance in relation with perfectionism, test procrastination and test anxiety of high school children. *Psychological Studies*, 48, 7-81.
- Taylor, J., & Deane, F.P. (2002). Development a short form of the Test Anxiety Inventory. *The Journal of General Psychology*, 129(2), 127-136. doi: 10.1080/00221300209603133
- 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
- Valero, L. (1999). Evaluación de ansiedad ante exámenes: Datos de aplicación y fiabilidad del cuestionario CAEX. *Anales de Psicología*, 15(2), 223-231.
- Villegas, G., Dominguez-Lara, S., Sotelo, N., & Sotelo, L. (2015). Propiedades psicométricas del Inventario de Autoevaluación de la Ansiedad ante Exámenes (IDASE) en universitarios de Lima. *Revista Mexicana de Psicología Educativa*, 3(1), 15-21.
- Ware, W. B., Galassi, J. P., & Dew, K. M. H. (1990). The test anxiety inventory: A confirmatory factor analysis. *Anxiety Research*, 3(3), 205-212. doi: 10.1080/08917779008248753
- Zinbarg, R. E., Yovel, I., Revelle, W., & McDonald, R. P. (2006). Estimating generalizability to a latent variable common to all of a scale's indicators: A comparison of estimators for  $\omega_h$ . *Applied Psychological Measurement*, 30(2), 121-144. doi: 10.1177/0146621605278814