



Ciencia Latina Revista Científica Multidisciplinar, Ciudad de México, México.
ISSN 2707-2207 / ISSN 2707-2215 (en línea), julio-agosto 2026,
Volumen 10, Número 4.

https://doi.org/10.37811/cl_rcm.v10i4

PROPIEDADES PSICOMÉTRICAS DEL NAQ-R EN EL CONTEXTO ACADÉMICO: ESTRUCTURA FACTORIAL, FIABILIDAD E INVARIANZA DE MEDICIÓN

**PSYCHOMETRIC PROPERTIES OF THE NAQ-R IN THE
ACADEMIC CONTEXT: FACTOR STRUCTURE, RELIABILITY
AND MEASUREMENT INVARIANCE**

Emily Karina Ordoñez Balcázar
Universidad Técnica de Machala

Alisson Pamela Moncada Sánchez
Universidad Técnica de Machala

Carlos Patricio Carpio Mosquera
Universidad Técnica de Machala

Marco Adrián Criollo Armijos
Universidad Técnica de Machala

Propiedades psicométricas del NAQ-R en el contexto académico: estructura factorial, fiabilidad e invarianza de medición

Emily Karina Ordoñez Balcázar¹

eordonez11@utmachala.edu.ec

<https://orcid.org/0009-0007-4048-8185>

Universidad Técnica de Machala, Facultad de Ciencias Químicas y de la Salud, Machala, Ecuador.

Carlos Patricio Carpio Mosquera

ccarpio@utmachala.edu.ec

<https://orcid.org/0000-0001-6335-1922>

Universidad Técnica de Machala, Facultad de Ciencias Químicas y de la Salud, Machala, Ecuador.

Alisson Pamela Moncada Sánchez

amoncada4@utmachala.edu.ec

<https://orcid.org/0009-0008-9071-6083>

Universidad Técnica de Machala, Facultad de Ciencias Químicas y de la Salud, Machala, Ecuador.

Marco Adrián Criollo Armijos

macriollo@utmachala.edu.ec

<https://orcid.org/0000-0001-9200-2203>

Universidad Técnica de Machala, Facultad de Ciencias Químicas y de la Salud, Machala, Ecuador.

RESUMEN

Este estudio tuvo como objetivo evaluar las propiedades psicométricas del Cuestionario de Actos Negativos Revisado (NAQ-R) adaptado al contexto académico universitario, en una muestra de 915 estudiantes ecuatorianos de ciencias de la salud ecuatorianos. Se examinó la estructura factorial mediante análisis factorial confirmatorio con estimador WLSMV y correlaciones policóricas, así como la dimensionalidad esencial mediante modelado bifactor. También se analizaron la consistencia interna, la validez convergente y discriminante, la invarianza de medición por género y la validez de criterio concurrente con una medida de autoetiquetado como víctima. El modelo trifactorial presentó un ajuste excelente ($CFI = .995$, $RMSEA = .039$), mientras que el modelo bifactor evidenció unidimensionalidad esencial, con una varianza común explicada de $.823$ y un omega jerárquico de $.94$ ($ECV = .823$, $\omega_H = .94$). La escala mostró una consistencia interna excelente ($\alpha = .933$). Así mismo, se estableció invarianza escalar entre hombres y mujeres, y se observaron correlaciones moderadas entre las puntuaciones del NAQ-R y el autoetiquetado como víctima ($p = .337-.490$). En conjunto, los resultados respaldan el uso de la puntuación total del NAQ-R como indicador primario de victimización académica en contextos universitarios latinoamericanos y aportan evidencia para su utilización en procesos de investigación y detección temprana.

Palabras clave: propiedades psicométricas; NAQ-R; acoso académico; análisis factorial confirmatorio; estudiantes universitarios

¹ Autor principal

Correspondencia: eordonez11@utmachala.edu.ec

Psychometric properties of the NAQ-R in the academic context: factor structure, reliability and measurement invariance

ABSTRACT

This study aimed to evaluate the psychometric properties of the Negative Acts Questionnaire-Revised (NAQ-R), adapted to the university academic context, in a sample of 915 Ecuadorian health sciences students. Factor structure was examined through confirmatory factor analysis using the WLSMV estimator and polychoric correlations, as well as essential dimensionality through bifactor modeling. Internal consistency, convergent and discriminant validity, measurement invariance by gender, and concurrent criterion validity using a self-labeling measure of victimization were also analyzed. The three-factor model showed excellent fit ($CFI = .995$, $RMSEA = .039$), whereas the bifactor model supported essential unidimensionality, with an explained common variance of .823 and a hierarchical omega of .94 ($ECV = .823$, $\omega_H = .94$). The scale demonstrated excellent internal consistency ($\alpha = .933$). Likewise, scalar invariance across men and women was established, and moderate correlations were observed between NAQ-R scores and self-labeling as a victim ($\rho = .337-.490$). Overall, the findings support the use of the NAQ-R total score as a primary indicator of academic victimization in Latin American university contexts and provide evidence for its application in research and early detection processes.

Keywords: psychometric properties; NAQ-R; academic bullying; confirmatory factor analysis; university students

*Artículo recibido 20 mayo 2026
Aceptado para publicación: 20 junio 2026*



INTRODUCCIÓN

La violencia psicológica en ámbitos educativos representa una forma de agresión que, pese a su carácter invisible, genera consecuencias sobre el bienestar emocional y el rendimiento académico de los estudiantes universitarios. En la educación superior, se manifiesta mediante dinámicas jerárquicas entre docentes y estudiantes, un problema tanto educativo como social insuficientemente documentado (Vargas et al., 2020; Hodgins et al., 2024).

La evidencia internacional documenta prevalencias entre el 18.3% y el 76.7% de estudiantes que reportan conductas de maltrato psicológico docente (Ardestani et al., 2022; Ruíz et al., 2020), incluyendo humillaciones, exclusión social, críticas destructivas y prácticas discriminatorias (Ramírez, 2024; Correa et al., 2022). En ciencias de la salud, Zhou et al. (2024) estimaron una prevalencia global del 65.6% entre estudiantes de enfermería.

Las consecuencias de la violencia psicológica docente trascienden el malestar emocional inmediato. Se asocian con síntomas de depresión y ansiedad, dificultades académicas y disminución del compromiso educativo (Aquino et al., 2022; Blanco et al., 2025; Gusfre et al., 2022; Strohmeier et al., 2024).

El constructo se ha conceptualizado como un fenómeno multidimensional que abarca distintas formas de conducta negativa (Einarsen et al., 2009; Hodgins et al., 2024). En el contexto latinoamericano y particularmente en Ecuador, la producción científica sobre violencia psicológica en el ámbito universitario permanece limitada. Aunque algunos estudios locales han reportado el acoso sexual, con aproximadamente el 11.08% de estudiantes reportando comentarios intimidatorios relacionados con su apariencia o sexualidad (Pereira et al., 2024), la violencia psicológica docente continúa siendo un fenómeno escasamente identificado desde una perspectiva empírica (Guerrero-Gonzales et al., 2025). Esta situación contrasta con la creciente visibilidad internacional del problema y limita el desarrollo de intervenciones basadas en datos contextualizados.

En este contexto, resulta imprescindible contar con instrumentos de medición que hayan sido validados específicamente en poblaciones universitarias latinoamericanas, pues la extrapolación directa de herramientas desarrolladas en otros contextos culturales e institucionales puede comprometer la validez de las inferencias derivadas. La adaptación rigurosa de instrumentos existentes, acompañada de una

evaluación sistemática de sus propiedades psicométricas, constituye por tanto un paso metodológico necesario antes de su aplicación en investigación o práctica clínica.

La medición válida y confiable de la victimización académica constituye un requisito para documentar empíricamente la violencia psicológica docente y diseñar intervenciones preventivas. El Cuestionario de Actos Negativos Revisado (NAQ-R) fue diseñado originalmente por Einarsen et al. (2009) para identificar conductas de acoso en entornos laborales donde existen relaciones jerárquicas entre las personas involucradas. Su estructura trifactorial captura distintas manifestaciones del acoso y ha mostrado propiedades psicométricas adecuadas en diversos marcos culturales. Estas características lo posicionan como instrumento transferible al contexto docente-estudiante, donde las relaciones de poder presentan analogías estructurales con el ámbito laboral.

La literatura reciente ha documentado considerable variabilidad en la estructura factorial del NAQ-R según el marco cultural y ocupacional, con algunos estudios identificando estructuras bifactoriales mientras otros mantienen el modelo trifactorial original. Más allá del número de factores, validaciones previas han señalado que las correlaciones interfactoriales del NAQ-R tienden a ser elevadas, lo que ha llevado a considerar la posibilidad de que el instrumento opere como una medida esencialmente unidimensional, esto es, que un factor general de victimización concentre la mayor parte de la varianza común y la fiabilidad de la puntuación total, aun cuando los factores específicos conserven valor descriptivo. Esta cuestión, central para definir el modo de puntuación e interpretación clínica del instrumento, no ha sido sistemáticamente examinada en adaptaciones académicas. Paralelamente, se han desarrollado instrumentos específicos para escenarios clínicos (Maghbouli et al., 2025), lo que sugiere la necesidad de evaluar empíricamente la transferibilidad del constructo a cada entorno específico.

Este estudio evalúa las propiedades psicométricas del Cuestionario de Actos Negativos Revisado (NAQ-R) adaptado al contexto académico universitario como instrumento de medición de la victimización académica, en una muestra de estudiantes de ciencias de la salud ecuatorianos. Específicamente, se examina: (a) la estructura factorial del instrumento mediante modelos confirmatorios alternativos, (b) la dimensionalidad esencial del constructo mediante modelado bifactor, (c) la consistencia interna a nivel de escala total y factores, (d) la validez convergente y discriminante, (e) la invarianza de medición por

género, y (f) la validez de criterio concurrente mediante la correlación con una medida de autoetiquetado como víctima.

METODOLOGÍA

Participantes

En el marco de un diseño instrumental (Ato et al., 2013; Ruiz-Robledillo et al., 2025), la muestra estuvo compuesta por $N = 915$ estudiantes universitarios de la Facultad de Ciencias Químicas y de la Salud de la Universidad Técnica de Machala, Ecuador. Los participantes pertenecían a las carreras de Psicología Clínica; Bioquímica y Farmacia; Enfermería; y Alimentos, cursando entre el primero y el décimo semestre. La selección se realizó mediante muestreo no probabilístico por conveniencia, considerando la accesibilidad a la población objetivo (Otzen y Manterola, 2017).

El tamaño muestral final se obtuvo tras aplicar, de forma previa a los análisis preliminares y a los análisis psicométricos principales, dos criterios de depuración secuenciales orientados a preservar la integridad de los datos. En primer lugar, se eliminaron registros duplicados exactos ($n = 113$; 10.8% de los 1,043 registros inicialmente recibidos), definidos operacionalmente como registros con valores idénticos en la totalidad de las variables recogidas —datos sociodemográficos, 22 ítems del NAQ-R e ítem de autoetiquetado—. Es importante precisar que estos 113 registros no correspondieron a respuestas uniformes intraindividuales, sino a duplicaciones exactas entre registros distintos.

Se identificaron 59 grupos de entre dos y ocho registros idénticos; en cada grupo se conservó únicamente la primera aparición. Este patrón fue considerado compatible con la posibilidad de envíos múltiples, dado que el formulario de Google Forms no restringía técnicamente la realización de más de un envío. En segundo lugar, se identificaron patrones de straightlining, reconocidos en la literatura metodológica como indicadores potenciales de respuesta descuidada y de reducción de la calidad de los datos en cuestionarios autoadministrados (Kim et al., 2019; Meade y Craig, 2012).

Este criterio fue operacionalizado como varianza intraindividual igual a cero en los 22 ítems del NAQ-R, correspondiente al procedimiento de simple nondifferentiation descrito por Kim et al. (2019). Se distinguieron dos subtipos: (a) 15 casos con respuesta uniforme “De vez en cuando” en todos los ítems, que fueron eliminados por considerarse psicométricamente poco plausibles e indicativos de respuesta no diferenciada de baja calidad, dado que implicaban reportar los 22 actos negativos evaluados con

idéntica frecuencia no nula; y (b) 163 casos con respuesta uniforme “Nunca”, que fueron conservados dado que la ausencia total de exposición es teóricamente esperable en escalas de victimización aplicadas a poblaciones generales, especialmente en presencia de efecto suelo (Einarsen et al., 2009; Reuning y Plutzer, 2020).

La muestra resultante (N = 915) preservó una ratio caso-ítem de 41.6:1, holgadamente superior a las recomendaciones convencionales para análisis factorial confirmatorio (Kline, 2016).

Instrumento

Se utilizó la versión española del NAQ-R validada por González-Trijueque y Graña (2013), adaptando el marco de referencia del ámbito laboral al académico. La adaptación consistió en modificar el marco de referencia de los ítems, sustituyendo referencias al entorno laboral por el ámbito universitario docente-estudiante, manteniendo la estructura semántica original. El proceso de adaptación siguió tres etapas. En primer lugar, se realizó una adaptación lingüística por parte del equipo investigador, en la que los términos propios del entorno laboral (p. ej., "trabajo", "compañeros", "supervisor") fueron reemplazados por referentes del contexto universitario (p. ej., "estudios", "compañeros de clase", "docente"), sin alterar la estructura sintáctica ni la carga semántica de los ítems.

En segundo lugar, se llevó a cabo un juicio de expertos con tres especialistas en psicometría y psicología educativa, quienes evaluaron la claridad, pertinencia y adecuación cultural de los ítems adaptados mediante una escala de cuatro puntos; todos los ítems alcanzaron índices de validez de contenido (IVC) $\geq .80$, indicando acuerdo satisfactorio entre evaluadores. En tercer lugar, se realizó una prueba piloto con 30 estudiantes universitarios de características similares a la muestra principal, quienes completaron el instrumento y reportaron verbalmente posibles dificultades de comprensión; los resultados no evidenciaron ambigüedades relevantes ni necesidad de reformulación adicional de los ítems.

El instrumento consta de 22 ítems distribuidos en tres factores teóricos: (a) Acoso relacionado con el trabajo académico (ítems 1-7), que evalúa conductas negativas vinculadas a las tareas académicas; (b) Acoso personal (ítems 8-12), que mide ataques dirigidos a la persona; y (c) Intimidación física (ítems 13-22), que evalúa conductas de amenaza o intimidación directa. Las opciones de respuesta se presentan en una escala tipo Likert de cinco puntos: 1 = Nunca, 2 = De vez en cuando, 3 = Mensualmente, 4 = Semanalmente, 5 = Diariamente.

Autoetiquetado. Adicionalmente, se incluyó un ítem de autoetiquetado donde los participantes indicaban si se consideraban víctimas de comportamientos negativos por parte de docentes durante el último semestre, utilizando la misma escala de frecuencia de cinco puntos. Este ítem sirvió como criterio externo para evaluar la validez concurrente del instrumento.

Procedimiento

Los datos se recolectaron mediante Google Forms entre octubre y diciembre de 2025. El cuestionario autoadministrado incluyó variables sociodemográficas y el NAQ-R adaptado. La participación fue voluntaria y anónima, con un tiempo de respuesta estimado de diez minutos.

Análisis de datos

Los análisis se realizaron con el software R versión 4.5.2 (R Core Team, 2025), utilizando los paquetes lavaan versión 0.6-20 (Rosseel, 2012) para el análisis factorial confirmatorio, semTools versión 0.5-7 (Jorgensen et al., 2022) para indicadores de fiabilidad compuesta, y psych versión 2.5.6 (Revelle, 2024) para estadísticos descriptivos y análisis de consistencia interna.

Análisis factorial confirmatorio. Dada la naturaleza ordinal de los datos y la violación del supuesto de normalidad (asimetría > 2 y curtosis > 7 en la mayoría de los ítems), se empleó el estimador de mínimos cuadrados ponderados con media y varianza ajustadas (WLSMV) con matriz de correlaciones policóricas, siguiendo las recomendaciones de Li (2016) y Rhemtulla et al. (2012) para datos categóricos. Se evaluaron tres modelos: (a) unifactorial, (b) bifactorial (Trabajo académico vs. Personal/Físico combinados), y (c) trifactorial (estructura teórica original). Asimismo, se estimó un modelo bifactor donde todos los reactivos convergen en un factor general de victimización académica (G) y tres factores específicos independientes (ortogonales) basados en la teoría. Para evaluar la dimensionalidad esencial, se utilizaron la Varianza Común Explicada (ECV) y el coeficiente omega jerárquico (ω_H), donde valores superiores a .70 y .75, respectivamente, indican unidimensionalidad esencial (Rodríguez et al., 2016).

La bondad de ajuste se evaluó mediante: chi-cuadrado (χ^2), Índice de Ajuste Comparativo (CFI), Índice de Tucker-Lewis (TLI), Error Cuadrático Medio de Aproximación (RMSEA) con su intervalo de confianza al 90%, y Raíz del Residuo Cuadrático Medio Estandarizado (SRMR). Se consideraron como indicadores de buen ajuste: CFI y TLI $\geq .95$, RMSEA $\leq .06$ y SRMR $\leq .08$ (Hu y Bentler, 1999). Las

cargas factoriales estandarizadas se interpretaron como aceptables ($\lambda \geq .40$), buenas ($\lambda \geq .50$) o excelentes ($\lambda \geq .70$).

Fiabilidad. La consistencia interna se evaluó mediante el coeficiente alfa de Cronbach (α) y el coeficiente omega de McDonald (ω), considerando valores $\geq .70$ como aceptables, $\geq .80$ como buenos y $\geq .90$ como excelentes (Viladrich et al., 2017). El coeficiente omega se considera más apropiado para modelos de AFC al no asumir tau-equivalencia.

Validez convergente y discriminante. La validez convergente se evaluó mediante la Varianza Media Extraída (AVE), considerando valores $\geq .50$ como adecuados (Fornell y Larcker, 1981). La validez discriminante se evaluó mediante los criterios de Fornell-Larcker y HTMT (Henseler et al., 2015), considerando umbrales de $<.85$ (estricto) y $<.90$ (liberal).

Se evaluó la invarianza de medición por género mediante una secuencia de modelos progresivamente restrictivos (configural, métrico, escalar), aplicando los criterios de Chen (2007): $\Delta CFI \leq .010$ y $\Delta RMSEA \leq .015$. Dado que cuatro ítems del Factor 3 (NAQ18, NAQ20, NAQ21, NAQ22) presentaron celdas vacías en categorías de mayor severidad dentro del grupo masculino, las categorías 3, 4 y 5 fueron colapsadas en una sola ("Mensualmente o más frecuente") para los 22 ítems en ambos grupos, a fin de preservar la estimación estable de umbrales bajo WLSMV (Sass et al., 2014). Esta recodificación se restringió a los análisis de invarianza; el resto de los análisis psicométricos conservó la escala original de cinco puntos.

Validez de criterio. La validez concurrente se evaluó mediante correlaciones de Spearman entre las puntuaciones del NAQ-R (total y por factores) y el ítem de autoetiquetado como víctima de acoso. La magnitud de las correlaciones se interpretó según los criterios de Cohen (1988): pequeña ($r = .10-.29$), moderada ($r = .30-.49$) y grande ($r \geq .50$).

Consideraciones éticas

La investigación se rigió por los principios de la Declaración de Helsinki (World Medical Association, 2013) y los lineamientos éticos de la APA (2017). El estudio fue aprobado por el comité de ética institucional (Resolución N°359/2020), instancia que cumple funciones de supervisión ética de la investigación en la institución. La participación fue voluntaria y anónima. Antes de acceder al cuestionario, cada participante leyó una hoja informativa que describía el propósito del estudio, la

naturaleza voluntaria de su participación, la ausencia de riesgos previsibles, el tratamiento anónimo de los datos y el derecho a abandonar la encuesta en cualquier momento sin consecuencias. El consentimiento se registró mediante una pregunta obligatoria de aceptación en Google Forms; solo quienes aceptaron pudieron continuar con el cuestionario. Se obtuvo autorización de los autores originales del NAQ-R (Guy Notelaers y Ståle V. Einarsen) para su adaptación al contexto académico ecuatoriano.

RESULTADOS Y DISCUSIÓN

Análisis preliminares

La adecuación muestral resultó excelente según el índice Kaiser-Meyer-Olkin ($KMO = .96$), superando el umbral de $.90$ considerado óptimo. La prueba de esfericidad de Bartlett fue estadísticamente significativa, $\chi^2 (231) = 11018.70$, $p < .001$, confirmando que la matriz de correlaciones difiere significativamente de una matriz identidad y que los datos son apropiados para el análisis factorial.

El análisis de los estadísticos descriptivos reveló una distribución marcadamente asimétrica en la mayoría de los ítems, con predominio de respuestas en la categoría "Nunca". Los datos presentaron un efecto suelo esperable en escalas de victimización no clínicas (Ma et al., 2017; Soler-Sánchez et al., 2025). El tratamiento analítico correspondiente se detalla en la sección de Análisis de datos.

Análisis factorial confirmatorio

Se evaluaron tres modelos factoriales alternativos: unifactorial, bifactorial (Trabajo académico vs. Personal/Físico combinados) y trifactorial (estructura teórica original). La Tabla 1 presenta los índices de ajuste comparativos, incluyendo el modelo bifactor. Los cuatro modelos evaluados mostraron índices de ajuste adecuados según los criterios de Hu y Bentler (1999). Entre los modelos de factores correlacionados, el trifactorial presentó el mejor ajuste relativo (Tabla 1), aunque las diferencias entre modelos fueron marginales dado que todos superaron los umbrales convencionales.

Tabla 1
Índices de Ajuste de los Modelos Factoriales Evaluados

Modelo	χ^2	gl	CFI	TLI	RMSEA [IC 90%]	SRMR
Unifactorial	731.22	209	.992	.991	.052 [.048, .056]	.062
Bifactorial	527.77	208	.995	.994	.041 [.037, .045]	.053
Trifactorial	491.62	206	.995	.995	.039 [.035, .043]	.051
Bifactor	305.41	187	.998	.998	.026 [.020, .032]	.044

Nota. χ^2 = chi-cuadrado; gl = grados de libertad; CFI = Índice de Ajuste Comparativo; TLI = Índice de Tucker-Lewis; RMSEA = Error Cuadrático Medio de Aproximación (intervalo de confianza 90% entre corchetes); SRMR = Raíz del Residuo Cuadrático Medio Estandarizado. Estimador WLSMV con correlaciones policóricas. N = 915.

El modelo bifactor presentó el mejor ajuste absoluto (Tabla 1). No obstante, más allá de los índices de ajuste, el análisis de dimensionalidad reveló que los factores específicos retuvieron una proporción mínima de la varianza común total ($S1 = 1.5\%$, $S2 = 0.2\%$, $S3 = 1.2\%$), mientras que el factor general concentró el 82.3% de dicha varianza ($ECV = .823$; Tabla 2). Estos valores superan los umbrales propuestos por Rodríguez et al. (2016) para inferir unidimensionalidad esencial, y son consistentes con un coeficiente omega jerárquico ($\omega_H = .94$) que atribuye al factor general la práctica totalidad de la fiabilidad de la puntuación total (Tabla 2). En consecuencia, el NAQ-R opera como una medida esencialmente unidimensional en esta muestra, donde la distinción trifactorial conserva valor descriptivo para la taxonomía de conductas negativas, pero no representa dimensiones de medida empíricamente independientes.

Tabla 2
Índices del Modelo Bifactor y Dimensionalidad

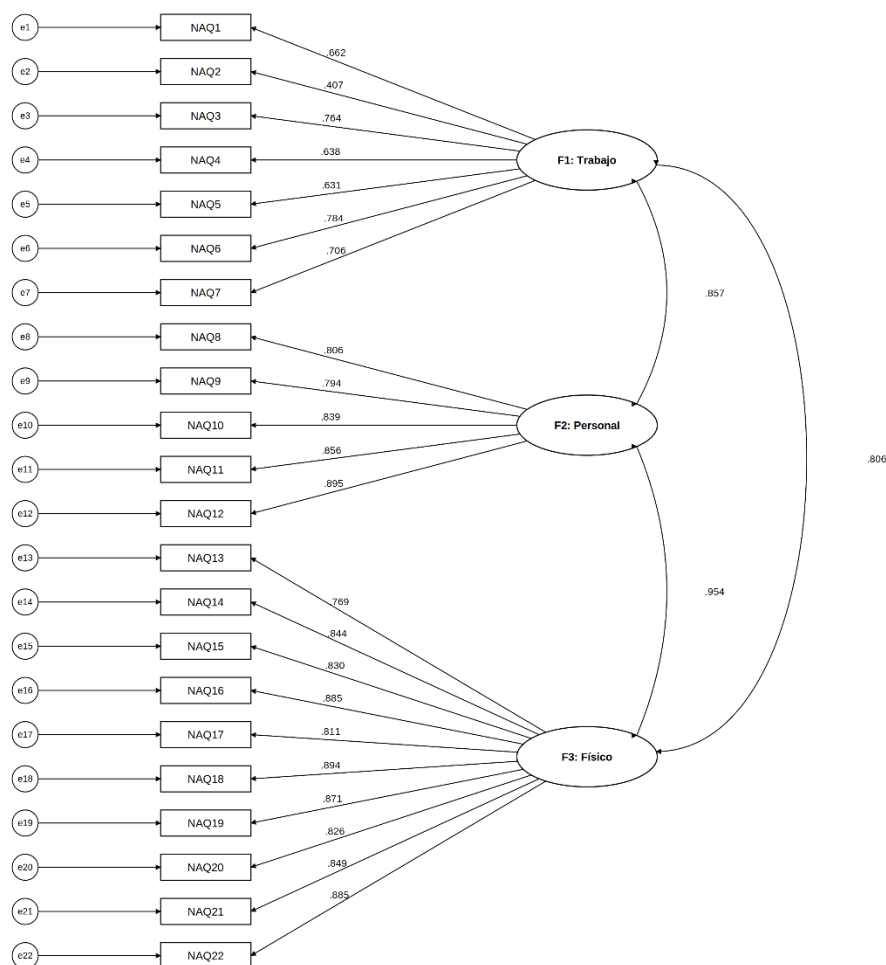
Índice	Valor	Interpretación
Varianza Factor G	12.49	—
Varianza Factor S1 (Trabajo)	1.03	—
Varianza Factor S2 (Personal)	1.10	—
Varianza Factor S3 (Físico)	0.56	—
ECV	.823	Unidimensionalidad esencial ($> .70$)
ωH	.94	Fiabilidad atribuible a G ($> .75$)
ω Total	.985	Fiabilidad total
$\omega H/\omega$	.956	Proporción de fiabilidad de G ($> .80$)

Nota. ECV = Varianza Común Explicada; ωH = omega jerárquico; G = factor general de victimización académica; S1, S2, S3 = factores específicos. Valores de ECV $> .70$ y $\omega H > .75$ indican unidimensionalidad esencial según Rodríguez et al. (2016).

Las cargas factoriales estandarizadas fueron significativas en todos los ítems ($p < .001$). El Factor 1 (Trabajo) presentó cargas entre .407 y .784 (media = .656); el Factor 2 (Personal) entre .794 y .895 (media = .838); y el Factor 3 (Físico) entre .769 y .894 (media = .846). El ítem NAQ2 mostró la carga más baja ($\lambda = .407$), dentro del rango aceptable. La estructura completa del modelo, incluyendo cargas factoriales y correlaciones interfactoriales, se presenta en la Figura 1.

Figura 1

Modelo trifactorial del NAQ-R adaptado al contexto académico (N = 915)



Nota. Las elipses representan los factores latentes y los rectángulos los ítems observados. Los círculos (e1–e22) corresponden a las varianzas residuales. Los valores sobre las flechas factor → ítem son cargas factoriales estandarizadas (λ); todas resultaron significativas ($p < .001$). Los valores asociados a las flechas curvas bidireccionales son correlaciones latentes entre factores (ϕ). Estimación realizada con WLSMV y correlaciones policóricas. F1: Acoso relacionado con el trabajo académico; F2: Acoso personal; F3: Intimidación física.

La Tabla 3 presenta las cargas factoriales estandarizadas y la varianza explicada de cada ítem.

Tabla 3

Cargas Factoriales Estandarizadas del Modelo Trifactorial

Ítem	F1: Trabajo	F2: Personal	F3: Físico	R ²
NAQ1	.662	—	—	.438
NAQ2	.407	—	—	.166
NAQ3	.764	—	—	.584
NAQ4	.638	—	—	.407
NAQ5	.631	—	—	.398
NAQ6	.784	—	—	.615
NAQ7	.706	—	—	.498
NAQ8	—	.806	—	.650
NAQ9	—	.794	—	.630
NAQ10	—	.839	—	.704
NAQ11	—	.856	—	.733
NAQ12	—	.895	—	.801
NAQ13	—	—	.769	.591
NAQ14	—	—	.844	.712
NAQ15	—	—	.830	.689
NAQ16	—	—	.885	.783
NAQ17	—	—	.811	.658
NAQ18	—	—	.894	.799
NAQ19	—	—	.871	.759
NAQ20	—	—	.826	.682
NAQ21	—	—	.849	.721
NAQ22	—	—	.885	.783

Nota. F1 = Acoso relacionado con el trabajo académico; F2 = Acoso personal; F3 = Intimidación física; R² = varianza explicada del ítem. Todas las cargas fueron significativas con $p < .001$.

Fiabilidad

Los indicadores de fiabilidad se resumen en la Tabla 4. La consistencia interna de la escala total fue excelente ($\alpha = .933$). A nivel de factores, la fiabilidad osciló entre aceptable y excelente (Tabla 4), con valores de omega ligeramente superiores a alfa, recurrencia esperable cuando las cargas no son homogéneas.

Tabla 4
Indicadores de Fiabilidad y Validez Convergente por Factor

Factor	Ítems	α	ω	AVE
F1: Acoso trabajo académico	7	.778	.786	.444
F2: Acoso personal	5	.852	.857	.703
F3: Intimidación física	10	.919	.931	.718
Escala total	22	.933	—	—

Nota. α = alfa de Cronbach; ω = omega de McDonald; AVE = Varianza Media Extraída. Criterio $AVE \geq .50$ para validez convergente adecuada. Un análisis de sensibilidad excluyendo el ítem NAQ2 arrojó un AVE de .489 para el Factor Trabajo, valor que permanece por debajo del umbral de .50.

Validez convergente y discriminante

La validez convergente, evaluada mediante la Varianza Media Extraída (AVE; ver Tabla 4), mostró resultados mixtos. Los factores Personal (AVE = .703) y Físico (AVE = .718) superaron el criterio de .50, indicando que más del 50% de la varianza de sus indicadores es explicada por el constructo latente. Sin embargo, el Factor Trabajo (AVE = .444) no alcanzó este umbral, principalmente debido a la carga más baja del ítem NAQ2 ($\lambda = .407$).

El criterio de Fornell-Larcker no se cumplió en ninguna comparación interfactorial: las correlaciones al cuadrado superaron consistentemente los valores de AVE. El análisis HTMT (Tabla 5) confirmó que únicamente el par Personal-Físico (.956) excedió el umbral liberal, convergiendo con la evidencia bifactor de indistinguibilidad empírica.

Tabla 5*Validez Discriminante: Ratio HTMT entre Factores*

Comparación	HTMT	Criterio estricto (< .85)	Criterio liberal (< .90)
Trabajo-Personal	.868	No cumple	Cumple
Trabajo-Físico	.812	Cumple	Cumple
Personal-Físico	.956	No cumple	No cumple

Nota. HTMT = Heterotrait-Monotrait Ratio (Henseler et al., 2015). Valores inferiores a .85 (criterio estricto) o a .90 (criterio liberal) respaldan la validez discriminante entre factores; valores que igualan o superan estos umbrales indican que los constructos no se distinguen empíricamente. En la presente muestra, el par Personal-Físico excede ambos criterios, lo que sustenta la interpretación esencialmente unidimensional desarrollada en la Discusión.

Invarianza de medición por género

La Tabla 6 resume los resultados del análisis de invarianza multigrupo, demostrando la equivalencia del instrumento entre mujeres ($n = 635$) y hombres ($n = 280$). El modelo configural presentó ajuste adecuado ($CFI = .997$, $RMSEA = .029$). La transición al modelo métrico ($\Delta CFI = .001$, $\Delta RMSEA = .005$) y posteriormente al modelo escalar ($\Delta CFI = -.001$, $\Delta RMSEA = -.007$) satisfizo los criterios de Chen (2007), estableciendo invarianza escalar completa. Estos resultados indican que el NAQ-R funciona de manera equivalente en ambos géneros, permitiendo comparaciones válidas de puntuaciones entre grupos.

Tabla 6*Invarianza de Medición por Género*

Nivel	CFI	RMSEA	ΔCFI	$\Delta RMSEA$	Decisión
Configural	.997	.029	—	—	—
Métrica	.996	.034	.001	.005	Invarianza establecida
Escalar	.997	.028	-.001	-.007	Invarianza establecida

Nota. Análisis multigrupo entre mujeres ($n = 635$) y hombres ($n = 280$). Criterios de Chen (2007): $\Delta CFI \leq .010$ y $\Delta RMSEA \leq .015$. Categorías de respuesta 3, 4 y 5 colapsadas en una sola exclusivamente para los análisis de invarianza.

Validez de criterio

La Tabla 7 presenta las correlaciones entre las puntuaciones del NAQ-R y el autoetiquetado como víctima de acoso. Todas las correlaciones fueron estadísticamente significativas ($p < .001$) y de magnitud moderada según los criterios de Cohen (1988), con el factor Intimidación Física mostrando la asociación más elevada. Las conductas de intimidación física y acoso personal resultan más notorias en la autopercepción de victimización que las conductas relacionadas con el trabajo académico.

Tabla 7

Correlaciones de Spearman entre Puntuaciones del NAQ-R y Autoetiquetado

Variable	ρ	IC 95%	p
F1: Acoso trabajo académico	.337	[.28, .39]	< .001
F2: Acoso personal	.441	[.39, .49]	< .001
F3: Intimidación física	.490	[.44, .54]	< .001
NAQ-R Total	.440	[.39, .49]	< .001

Nota. ρ = rho de Spearman; IC = intervalo de confianza. Criterio externo: autoetiquetado como víctima de comportamientos negativos por parte de docentes. N = 905 (casos con datos completos en autoetiquetado).

Discusión

Este estudio evaluó las propiedades psicométricas del NAQ-R adaptado al contexto académico en estudiantes ecuatorianos de ciencias de la salud. Los resultados confirman un ajuste excelente del modelo trifactorial, alta consistencia interna y validez de criterio concurrente adecuada.

Estructura factorial

El modelo trifactorial del NAQ-R mostró ajuste adecuado en el contexto académico (CFI = .995, RMSEA = .039), en línea con la validación original (Einarsen et al., 2009) y adaptaciones hispanohablantes (Dujó López et al., 2020). Sin embargo, el hallazgo central de este estudio es que los indicadores bifactor (Tabla 2) confirmaron unidimensionalidad esencial, indicando que la estructura trifactorial, aunque estadísticamente viable, no representa dimensiones de medida empíricamente independientes. Este patrón es coherente con la conceptualización de la violencia psicológica docente como un síndrome unitario donde las manifestaciones conductuales covarían estrechamente (Einarsen

et al., 2009), y respalda el uso de la puntuación total del NAQ-R como indicador primario de victimización académica.

La variabilidad factorial del NAQ-R documentada internacionalmente (Sá Cunha et al., 2024) sugiere que la convergencia hacia un constructo unitario en el entorno académico puede explicarse por la concentración del poder en una figura de autoridad única. La consistencia interna (Tabla 4) es comparable a validaciones internacionales del NAQ-R (Einarsen et al., 2009; Dujo López et al., 2020; Sá Cunha et al., 2024).

Correlaciones interfactoriales y validez discriminante

Las correlaciones interfactoriales elevadas (.806–.954) no constituyen una deficiencia psicométrica; representan, más bien, evidencia empírica de la estructura esencialmente unidimensional del constructo. En el escenario docente-estudiante, la indistinguibilidad empírica entre acoso personal e intimidación física puede explicarse por las características específicas de la asimetría de poder académica. La autoridad concentrada del docente hace que las conductas intimidatorias o de acoso personal se solapen fenomenológicamente, constituyendo un mecanismo unitario de victimización directa donde la diferenciación perceptual resulta difícil para las víctimas (Einarsen et al., 2009). Por lo tanto, se desaconseja la interpretación diferencial de estas subescalas para fines de cribado en ámbitos académicos.

El AVE subóptimo del Factor Trabajo (Tabla 4) es consistente con Dujo López et al. (2020), quienes reportaron valores de .35–.37 en el NAQ-R español. Dado que las cargas y la fiabilidad son aceptables, este resultado no invalida el factor.

La insuficiencia de validez discriminante según el criterio de Fornell-Larcker concuerda con la estructura esencialmente unidimensional revelada por el modelo bifactor. Cuando las correlaciones latentes superan .80, este criterio resulta algebraicamente insatisfactorio (Henseler et al., 2015), limitación que la literatura sobre el NAQ-R ha identificado recurrente (Aoki et al., 2024). El análisis HTMT (Tabla 5) confirmó que únicamente el par Personal-Físico excede el umbral liberal, reforzando que estas dimensiones no operan como constructos empíricamente diferenciados en el contexto académico. Este hallazgo sugiere que, en escenarios universitarios caracterizados por relaciones jerárquicas asimétricas entre docentes y estudiantes, la distinción entre acoso personal e intimidación

física podría carecer de relevancia empírica sustantiva, en la medida en que ambas formas de victimización parecen manifestarse de manera entrelazada desde la experiencia subjetiva del estudiante. En este sentido, aunque la estructura trifactorial mantuvo un ajuste estadístico excelente, los resultados del modelo bifactor respaldan la conveniencia de interpretar el instrumento desde una perspectiva esencialmente unidimensional. Futuras investigaciones podrían contrastar formalmente la estructura trifactorial con modelos bifactoriales como alternativa más parsimoniosa, particularmente en contextos académicos donde la diferenciación entre dimensiones específicas parece limitada.

Validez de criterio concurrente

Las correlaciones entre las puntuaciones del NAQ-R y el autoetiquetado como víctima fueron moderadas ($\rho = .337-.490$), siguiendo la tendencia teóricamente esperada. Investigaciones previas han establecido que las medidas conductuales (como el NAQ-R) y las medidas de autoetiquetado capturan constructos relacionados, pero parcialmente distintos (Nielsen et al., 2021). Esta discrepancia refleja que muchos individuos experimentan conductas objetivamente negativas, pero no se identifican como víctimas, un fenómeno atribuible a mecanismos de negación defensiva, normalización cultural del maltrato o desconocimiento del término "acoso" (Nielsen et al., 2021).

El factor Intimidación Física presentó la correlación más alta con el autoetiquetado ($\rho = .490$), próxima al umbral de magnitud grande. Este resultado sugiere que las conductas de intimidación directa son reconocidas con mayor facilidad como acoso, mientras que las relacionadas con el trabajo académico ($\rho = .337$) tienden a confundirse con exigencia pedagógica. En la academia latinoamericana, el maltrato verbal se disfraza con frecuencia de rigor académico (Obermeier Pérez, 2025), lo que explicaría esta diferencia.

Efecto suelo

Los datos presentaron un efecto suelo severo, con más del 80% de respuestas en la categoría "Nunca" para la mayoría de los ítems, particularmente aquellos relacionados con intimidación física. Soler-Sánchez et al. (2025) y Ma et al. (2017) documentaron patrones equivalentes, lo que sitúa este resultado dentro del rango esperable para el constructo.

Implicaciones prácticas

Los hallazgos respaldan el uso del NAQ-R adaptado como herramienta de investigación en contextos académicos de ciencias de la salud latinoamericanos. Dado que las subescalas no aportan información incremental sustantiva (varianza específica mínima), se recomienda priorizar la puntuación total como indicador primario. La correlación moderada con el autoetiquetado sugiere que ambos métodos de evaluación son complementarios: el NAQ-R captura la exposición objetiva a conductas negativas, mientras que el autoetiquetado refleja el procesamiento psicológico de dichas experiencias. No obstante, el autoetiquetado operó como ítem único de criterio, práctica habitual en la literatura sobre acoso (Nielsen et al., 2021), aunque su interpretación debe considerarse indicativa más que confirmatoria de validez de criterio.

Para las instituciones de educación superior, estos hallazgos proporcionan una base para futuras validaciones en marcos universitarios latinoamericanos. Dado que las conductas negativas tienden a co-ocurrir, las intervenciones deben abordar el fenómeno de manera integral. La discrepancia entre exposición conductual y autopercepción de victimización indica que los programas de sensibilización deben incluir componentes educativos que ayuden a los estudiantes a reconocer las conductas negativas como formas de acoso, superando la normalización cultural que puede operar en entornos académicos jerárquicos.

Desde el ámbito de la gestión institucional, la existencia de un instrumento psicométricamente validado abre la posibilidad de implementar sistemas de monitoreo periódico del clima académico. Estos sistemas permitirían detectar de manera temprana situaciones de riesgo, orientar recursos hacia las carreras o semestres con mayor prevalencia de conductas negativas y evaluar el impacto de las intervenciones implementadas, dotando así a las instituciones de educación superior de una herramienta concreta para la toma de decisiones basada en datos.

Limitaciones y direcciones futuras

El muestreo por conveniencia y el diseño transversal restringen la generalización a otras poblaciones universitarias e impiden establecer relaciones causales o evaluar estabilidad temporal. Aunque la muestra incluyó cuatro carreras de ciencias de la salud, no se examinó la invarianza de medición entre ellas, lo cual constituye una prioridad para investigaciones futuras. Esta limitación adquiere particular

relevancia considerando que las carreras evaluadas presentan diferencias curriculares e institucionales potencialmente significativas —por ejemplo, en la intensidad de prácticas clínicas, modalidades de supervisión, formas de evaluación académica y grado de interacción jerárquica docente-estudiante—, factores que podrían modular tanto la percepción como la manifestación del acoso académico. Asimismo, la restricción a una sola institución limita la generalización de los hallazgos a otros contextos universitarios ecuatorianos y latinoamericanos. Futuros estudios deberían emplear muestras probabilísticas de múltiples instituciones, examinar la equivalencia psicométrica del instrumento entre carreras y explorar la posible influencia de las características formativas específicas sobre la experiencia de victimización en contextos universitarios.

CONCLUSIONES

El hallazgo principal del presente estudio reside en la unidimensionalidad esencial del instrumento: aunque la estructura trifactorial conserva utilidad descriptiva para caracterizar las distintas manifestaciones del fenómeno, el NAQ-R opera fundamentalmente como una medida unitaria de victimización en contextos universitarios. Este resultado indica que las conductas hostiles registradas, independientemente de su naturaleza específica, convergen hacia un factor general de acoso percibido, lo cual otorga solidez conceptual y parsimonia psicométrica al instrumento. Por tanto, se recomienda emplear la puntuación total como indicador primario en investigaciones empíricas y procesos de detección temprana, reservando el análisis por subescalas para aquellas situaciones que demanden una caracterización más específica del fenómeno.

Desde una perspectiva aplicada, los indicadores satisfactorios de validez y confiabilidad obtenidos posicionan al NAQ-R como un recurso de utilidad concreta para profesionales vinculados a la salud mental, la orientación educativa y la gestión institucional. Su implementación sistemática en entornos universitarios podría facilitar el diseño de programas de intervención, estrategias preventivas y políticas institucionales dirigidas a erradicar el acoso académico. La invarianza escalar por género representa además un valor añadido del instrumento, ya que garantiza que las comparaciones entre hombres y mujeres reflejan diferencias reales en el constructo y no artefactos de medición, lo que amplía su utilidad en estudios que busquen identificar patrones diferenciales de victimización.

La disponibilidad de un instrumento psicométricamente validado en el contexto universitario latinoamericano constituye una condición para que la investigación sobre conductas negativas en la relación docente-estudiante avance hacia la documentación empírica sistemática del fenómeno y el desarrollo de líneas de investigación en la región.

DECLARACIÓN DE CONFLICTO DE INTERESES

Los autores declaran que no existe ningún conflicto de intereses.

FINANCIACIÓN

Esta investigación no recibió financiación específica de agencias de financiación en los sectores público, comercial o sin fines de lucro.

REFERENCIAS BIBLIOGRÁFICAS

- Aoki, P. R., Ferreira, M. C., Vilar, G., y Sousa, N. F. (2024). Cuestionario de Actos Negativos Revisado: Adaptación y validación al contexto educativo brasileiro. *Psico-USF*, 29(1), 1-14. <https://doi.org/10.1590/1413-82712024290101>
- Aquino, C. D., Galdino, M. J. Q., Martins, J. T., Ribeiro, R. P., y Robazzi, M. L. C. C. (2022). Workplace bullying and burnout in nursing workers. *Acta Paulista de Enfermagem*, 35, eAPE02436. <https://doi.org/10.37689/acta-ape/2022AO02436>
- Ardestani, S. M. S., Mosadeghrad, A. M., y Khosravi, S. (2022). Workplace bullying among Iranian nurses: A systematic review. *Nursing Ethics*, 29(3), 631-647. <https://doi.org/10.1177/09697330211051104>
- Ato, M., López-García, J. J., y Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en psicología [A classification system of research designs in psychology]. *Anales de Psicología*, 29(3), 1038-1059. <https://doi.org/10.6018/analesps.29.3.178511>
- Blanco, V., Sanmartín, R., y Otero, C. (2025). Cybervictimization and psychological outcomes in university students. *Computers in Human Behavior*, 163, 108477. <https://doi.org/10.1016/j.chb.2024.108477>
- Chen, F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464-504. <https://doi.org/10.1080/10705510701301834>



- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2.^a ed.). Lawrence Erlbaum Associates.
- Correa, L., Pedraza, L., y Vargas, C. (2022). Violencia psicológica en contextos educativos universitarios de salud. *Revista de Salud Pública*, 24(2), 1-8. <https://doi.org/10.15446/rsap.v24n2.98045>
- Dujo López, V., Horcajo-Gil, P. J., Nuevo, R., Baca-García, E., y Perez-Rodriguez, M. M. (2020). Psychometric validation of the Spanish version of the Negative Acts Questionnaire-Revised (NAQ-R) in the workplace. *Psicothema*, 32(4), 584-591. <https://doi.org/10.7334/psicothema2020.174>
- Einarsen, S., Hoel, H., y Notelaers, G. (2009). Measuring exposure to bullying and harassment at work: Validity, factor structure and psychometric properties of the Negative Acts Questionnaire-Revised. *Work & Stress*, 23(1), 24-44. <https://doi.org/10.1080/02678370902815673>
- Fornell, C., y Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39-50. <https://doi.org/10.2307/3151312>
- González-Trijueque, D., y Graña, J. L. (2013). El acoso psicológico en el lugar de trabajo: prevalencia y análisis descriptivo en una muestra multiocupacional [Psychological harassment in the workplace: prevalence and descriptive analysis in a multi-occupational sample]. *Psicopatología Clínica, Legal y Forense*, 9, 37-62.
- Guerrero-Gonzales, A., Torres, M., y Ponce, R. (2025). Violencia docente en universidades latinoamericanas: Una revisión sistemática [Teacher violence in Latin American universities: A systematic review]. *Educación XX1*, 28(1), 245-268. <https://doi.org/10.5944/educXX1.35842>
- Gusfre, K. S., Lohne, V., Severinsson, E., y Lindstrom, I. (2022). A systematic review of interventions to prevent and reduce workplace bullying in healthcare. *Journal of Nursing Management*, 30(5), 1359-1375. <https://doi.org/10.1111/jonm.13607>
- Henseler, J., Ringle, C. M., y Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, 43(1), 115-135. <https://doi.org/10.1007/s11747-014-0403-8>



- Hodgins, M., Mac Curtain, S., y Mannix-McNamara, P. (2024). Workplace bullying and harassment in health and social care: A systematic review. *Trauma, Violence, & Abuse*, 25(2), 1357-1371. <https://doi.org/10.1177/15248380231173479>
- Hu, L., y Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55. <https://doi.org/10.1080/10705519909540118>
- Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., y Rosseel, Y. (2022). semTools: Useful tools for structural equation modeling [Software de computadora]. <https://CRAN.R-project.org/package=semTools>
- Kim, Y., Dykema, J., Stevenson, J., Black, P., y Moberg, D. P. (2019). Straightlining: Overview of measurement, comparison of indicators, and effects in mail–web mixed-mode surveys. *Social Science Computer Review*, 37(2), 214–233. <https://doi.org/10.1177/0894439317752406>
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4ª ed.). The Guilford Press.
- Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, 48(3), 936-949. <https://doi.org/10.3758/s13428-015-0619-7>
- Ma, C., Shang, J., y Bott, M. J. (2017). Linking unit collaboration and nursing leadership to nurse outcomes and quality of care. *Journal of Nursing Administration*, 47(7-8), 396-405. <https://doi.org/10.1097/NNA.0000000000000497>
- Maghbouli, N., Rezaee, R., y Soltani, M. (2025). Development and validation of an academic bullying questionnaire for medical students. *BMC Medical Education*, 25, 112. <https://doi.org/10.1186/s12909-025-06112-3>
- Meade, A. W., y Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods*, 17(3), 437–455. <https://doi.org/10.1037/a0028085>
- Nielsen, M. B., Tangen, T., Idsoe, T., Matthiesen, S. B., y Magerøy, N. (2021). Post-traumatic stress disorder as a consequence of bullying at work and at school. A literature review and meta-analysis. *Aggression and Violent Behavior*, 21, 17-24. <https://doi.org/10.1016/j.avb.2015.01.001>

- Obermeier Pérez, C. (2025). El disfraz del rigor: Violencia psicológica normalizada en universidades latinoamericanas [The disguise of rigor: Normalized psychological violence in Latin American universities]. *Revista Iberoamericana de Educación Superior*, 16(45), 1-18.
<https://doi.org/10.22201/issue.20072872e.2025.45.1534>
- Otzen, T., y Manterola, C. (2017). Técnicas de muestreo sobre una población a estudio [Sampling techniques on a population study]. *International Journal of Morphology*, 35(1), 227-232.
<https://doi.org/10.4067/S0717-95022017000100037>
- Pereira, C., Vera, M., y Torres, K. (2024). Acoso sexual en la universidad ecuatoriana: Prevalencia y factores asociados [Sexual harassment in Ecuadorian university: Prevalence and associated factors]. *Revista Ecuatoriana de Medicina y Ciencias Biológicas*, 45(1), 23-35.
<https://doi.org/10.26807/remcb.v45i1.1089>
- R Core Team. (2025). R: A language and environment for statistical computing [Software de computadora]. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Ramírez, M. A. (2024). Violencia docente en ciencias de la salud: Una revisión sistemática [Teacher violence in health sciences: A systematic review]. *Educación Médica*, 25(1), e100828.
<https://doi.org/10.1016/j.edumed.2024.100828>
- Reuning, K., y Plutzer, E. (2020). Valid vs. Invalid Straightlining: The Complex Relationship Between Straightlining and Data Quality. *Survey Research Methods*, 14(5), 439–459.
<https://doi.org/10.18148/srm/2020.v14i5.7641>
- Revelle, W. (2024). psych: Procedures for psychological, psychometric, and personality research [Software de computadora]. <https://CRAN.R-project.org/package=psych>
- Rhemtulla, M., Brosseau-Liard, P. É., y Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods*, 17(3), 354-373.
<https://doi.org/10.1037/a0029315>
- Rodriguez, A., Reise, S. P., y Haviland, M. G. (2016). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment*, 98(3), 223-237.
<https://doi.org/10.1080/00223891.2015.1089249>



- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1-36. <https://doi.org/10.18637/jss.v048.i02>
- Ruíz, C., Sánchez, M., y Romero, A. (2020). Workplace bullying in healthcare: Systematic review and meta-analysis. *Journal of Advanced Nursing*, 76(4), 956-972. <https://doi.org/10.1111/jan.14284>
- Ruiz-Robledillo, N., Ferrer-Cascales, R., Albaladejo-Blázquez, N., Sánchez-SanSegundo, M., y Zaragoza-Martí, A. (2025). Validation of an instrument to measure aggressive behavior in university students. *Frontiers in Psychology*, 16, 1548932. <https://doi.org/10.3389/fpsyg.2025.1548932>
- Sass, D. A., Schmitt, T. A., y Marsh, H. W. (2014). Evaluating model fit with ordered categorical data within a measurement invariance framework: A comparison of estimators. *Structural Equation Modeling: A Multidisciplinary Journal*, 21(2), 167–180. <https://doi.org/10.1080/10705511.2014.882658>
- Sá Cunha, F., Costa, E., Ferreira, M. C., y Brandão, C. (2024). Propriedades psicométricas do NAQ-R em estudantes universitários de ciências da saúde brasileiros [Psychometric properties of the NAQ-R in Brazilian health sciences university students]. *Psicologia: Reflexão e Crítica*, 37, 12. <https://doi.org/10.1186/s41155-024-00312-7>
- Soler-Sánchez, M. I., Meseguer-de-Pedro, M., y García-Izquierdo, M. (2025). Sexual harassment at work: Psychometric properties of a measurement scale in a Spanish sample. *International Journal of Environmental Research and Public Health*, 22(1), 68. <https://doi.org/10.3390/ijerph22010068>
- Strohmeier, D., Gradinger, P., y Spiel, C. (2024). Cyberbullying and traditional bullying: Conceptualization, measurement, and overlap. *Zeitschrift für Psychologie*, 232(1), 3-14. <https://doi.org/10.1027/2151-2604/a000538>
- Vargas, A., Soto, M., y Quijada, F. (2020). Violencia en la educación superior: Revisión de literatura latinoamericana [Violence in higher education: Latin American literature review]. *Actualidades Investigativas en Educación*, 20(3), 1-27. <https://doi.org/10.15517/aie.v20i3.43171>

- Viladrich, C., Angulo-Brunet, A., y Doval, E. (2017). A journey around alpha and omega to estimate internal consistency reliability. *Anales de Psicología*, 33(3), 755-782.
<https://doi.org/10.6018/analesps.33.3.268401>
- World Medical Association. (2013). World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. *JAMA*, 310(20), 2191-2194.
<https://doi.org/10.1001/jama.2013.281053>
- Zhou, H., Peng, J., Wang, D., Xu, X., y Feng, X. (2024). Global prevalence of bullying among nursing students: A meta-analysis and systematic review. *Nurse Education Today*, 133, 106090.
<https://doi.org/10.1016/j.nedt.2023.106090>

