



PROPIEDADES PSICOMÉTRICAS DEL PHQ-9 EN USUARIOS PERUANOS CON PROBLEMAS DE CONSUMO DE ALCOHOL DESDE EL MODELO DE RESPUESTA GRADUADA DE SAMEJIMA

*PSYCHOMETRIC PROPERTIES OF THE PHQ-9 IN PERUVIAN USERS WITH ALCOHOL USE
PROBLEMS BASED ON SAMEJIMA'S GRADED RESPONSE MODEL*

***Aaron Caycho-Caja**

*FCCTP - Universidad de San Martín de Porres, Perú
acaychoc@usmp.pe*

<https://orcid.org/0000-0003-1478-8954>

Zoraida Barrios-Sallo

*Universidad César Vallejo, Perú
zbarriosb@ucvvirtual.edu.pe*

<https://orcid.org/0000-0002-0275-2340>

Eder Jesús Orihuela-Manrique

*Universidad Privada del Norte, Perú
eder.orihuela@upn.pe*

<https://orcid.org/0000-0002-7447-1601>

Como citar: Caycho-Caja, A., Barrios-Sallo, Z., & Orihuela-Manrique, E. J. (2025). Propiedades psicométricas del PHQ-9 en usuarios peruanos con problemas de consumo de alcohol desde el modelo de respuesta graduada de Samejima. *Health and Addictions/Salud y Drogas*, 25(2), 48-64. <https://doi.org/10.21134/986>

*Autor de correspondencia

Resumen

Introducción. El PHQ-9 es un instrumento poco explorado desde la teoría de respuesta al ítem (TRI) y menos en muestra clínica con problemas de consumo de alcohol; por lo cual el objetivo del presente estudio es brindar evidencia de validez del PHQ-9 desde la TRI. **Método.** Se seleccionaron 484 usuarios con problemas de uso de alcohol de la encuesta nacional peruana ENDES, de estos el 82% son varones y el restante 17%, mujeres. El estudio tiene un diseño instrumental y transversal realizado desde la TRI, modelo de respuesta graduada de Samejima. Se procuró identificar la unidimensionalidad del constructo desde un análisis factorial confirmatorio, seguido de la evaluación de este desde índices de ajuste del factor único. Adicionalmente, se analizaron los parámetros de discriminación y dificultad, además de las curvas características de los ítems y del test, entre otros análisis. **Resultados.** El análisis factorial inicial evidenció la unidimensionalidad del modelo, lo cual también se evidenció en el ajuste inicial del modelo TRI (RMSEA =0.062, SRMSR=0.068, TLI=0.972, CFI=0.979); el ítem 4 (cansancio) se mostró como el más discriminativo, también el ítem 9 (ideas de suicidio) mostró mayor grado de dificultad. Además, se observó una baja probabilidad de elección a la tercera opción de respuesta “la mayoría de días”. Adicionalmente, el constructo de depresión logra explicar el consumo de alcohol ($\beta=0.24$, $p<.001$). **Conclusión.** El PHQ-9 es un instrumento que puede emplearse en población clínica de usuarios con problemas de consumo de alcohol.

Palabras clave

PHQ-9; Teoría de Respuesta al Ítem; Consumidores de alcohol; Modelo de Samejima; MRG.

Abstract

Introduction. The PHQ-9 is a scarcely explored instrument from the perspective of item response theory (IRT), particularly in clinical samples with alcohol use problems. Therefore, the aim of the present study is to provide validity evidence for the PHQ-9 based on IRT. **Method.** A total of 484 individuals with alcohol use problems were selected from the Peruvian national survey (ENDES); 82% were male, and the remaining 17% were female. The study employed a cross-sectional, instrumental design within the framework of IRT, specifically using Samejima’s graded response model. Unidimensionality of the construct was assessed through confirmatory factor analysis, followed by evaluation of model fit indices for a single factor. Additionally, item discrimination and difficulty parameters were examined, along with item and test characteristic curves, among other analyses. **Results.** The initial factor analysis supported the unidimensionality of the model, which was further confirmed by adequate IRT model fit indices (RMSEA = .062, SRMSR = .068, TLI = .972, CFI = .979). Item 4 (*fatigue*) showed the highest discrimination, while item 9 (*suicidal ideation*) exhibited the greatest difficulty. Moreover, the response option “more than half the days” showed a low probability of being selected. Additionally, the depression construct significantly predicted alcohol use ($\beta = 0.24$, $p < .001$). **Conclusion.** The PHQ-9 is a psychometric instrument suitable for use in clinical populations with alcohol use problems.

Key Words

PHQ-9; Item Response Theory; Alcohol users; Samejima model; GRM.

1. Introducción

La depresión es un trastorno afectivo caracterizado por un estado de ánimo que tiende a la tristeza, desinterés por las actividades cotidianas, problemas de apetito o insomnio, pérdida de energía, entre otros. Se señala que un 3.8% de personas en el mundo padecen depresión, uno de los trastornos mentales más comunes y frecuentes en mujeres (6%) a comparación de los varones (4%) (Organización Mundial de la Salud [OMS], 2023). En el Perú, la prevalencia anual de depresión oscila entre un 6 y 10% de la población (Instituto Nacional de Salud Mental [INSM], 2024); en el caso de los jóvenes, más del 30% de ellos ha experimentado síntomas depresivos en el periodo posterior a la pandemia (Ministerio de Educación [MINEDU], 2023).

La depresión puede considerarse como el factor que más contribuye a la discapacidad en el mundo (OMS, 2017), ocasiona diversas alteraciones que incluso pueden llevar a quien la padece al suicidio (Organización Panamericana de la Salud [OPS], 2021) y generó, en el 2020, un impacto en los años de vida perdidos por discapacidad (AVAD) de 34 a 69 millones en todo el mundo (Santomauro et al., 2021). También suele aparecer en comorbilidad con el alcoholismo y la asociación entre ambas puede ser peligrosa ya que estas pueden conllevar a un peor pronóstico de cualquiera de estos trastornos (McHugh and Weiss, 2019). La depresión también puede influir en los motivos por los cuales las personas consumen alcohol (Desalu et al., 2019; Keum y Cano, 2023), es un predictor del consumo del alcohol en mujeres (Villanueva-Blasco et al., 2022) y suele asociarse al afrontamiento al alcohol (Desalu et al., 2019) y a su uso severo (Keum y Cano, 2023).

El alcohol es un tipo de bebida legal ampliamente utilizada en el país andino; no obstante, bajo ciertas condiciones, un patrón de uso social puede escalar a un trastorno de abuso o dependencia al alcohol; estos cuadros psicopatológicos se asocian con diversos problemas de salud física (Grewal et al., 2022; Yeo et al., 2022) y mental (Yoshimura et al., 2022), además de ansiedad o depresión (Desalu et al., 2017; Grewal et al., 2022; Keum y Cano, 2023; Villanueva-Blasco et al., 2022). Se señala que en la actualidad tanto dependencia como abuso de alcohol (Manual Diagnóstico y Estadístico de Trastornos Mentales [DSM-IV-TR], American Psychiatric Association [APA], 2000) se han fusionado en una sola categoría denominada trastorno por consumo de alcohol (AUD, por su abreviatura en inglés), descrito en el DSM-V-TR (APA, 2022), el cual hace referencia a patrones desadaptativos con relevancia clínica que traen consigo perturbaciones personales, sociales, familiares y laborales acaecidas en un lapso de doce meses. Asimismo, para el diagnóstico de AUD es necesario cumplir con, por lo menos, dos de los once criterios establecidos por el manual citado.

En el contexto peruano las personas que padecen abuso de alcohol o dependencia suelen tener muchos problemas para acceder a los servicios de salud ya que es conocido que el actual sistema de salud peruano no tiene la capacidad para la atención de problemas de salud física o mental (INSM, 2024). La situación se hace más grave si los casos de consumo de alcohol se presentan en comorbilidad de los cuadros descritos, entre ellos la depresión (Desalu et al., 2019; Grewal et al., 2022; Keum y Cano, 2023; McHugh and Weiss, 2019; Villanueva-Blasco et al., 2022).

Si bien no es un enfoque tan extendido como la teoría clásica de los test, la teoría de respuesta al ítem (TRI) pretende establecer un nexo entre el comportamiento que dan los sujetos a un ítem y el rasgo latente subyacente (Attorresi et al., 2009) a partir de la probabilidad de respuesta al ítem y bajo ciertos parámetros como dificultad, discriminación, pseudoazar, entre otros. Dentro de este enfoque, Samejima (1969) propuso un modelo de respuesta graduada (MRG) para ítems politómicos en el cual la probabilidad de respuesta a cada categoría del reactivo depende del rasgo latente, según sus parámetros de dificultad y discriminación. La discriminación de un ítem, análogo a la carga factorial, hace referencia a la asociación entre un ítem y el rasgo latente medido; mientras que la dificultad, dado un ítem con K categorías con K-1 umbrales (b), señala el nivel de rasgo latente θ donde hay un 50% de probabilidad de elegir la categoría k+1 o una superior (Abad et al., 2011). Este enfoque ha comenzado a generar mayor interés, y su aplicación a la medición de instrumentos cortos como el *Patient Health Questionnaire* (PHQ-9) ya se ha dado en otros contextos (Arias et al., 2024; Cumbe et al., 2020; Gao y Liu, 2024;

Kitamura et al., 2023; Ma et al., 2021; Reich et al., 2018), aunque aún es muy poco frecuente en Perú (Carranza-Esteban et al., 2024; Lingán-Huamán et al., 2023)

Respecto a la medición de la depresión, el PHQ-9 ha brindado características favorables para la medición de sus síntomas desde la TRI, siendo frecuentemente utilizado el MRG de Samejima. Desde el MRG, por ejemplo, Cumbe et al. (2020), en pacientes de Mozambique, hallaron parámetros de discriminación (a) entre su máximo (máx.) y mínimo (mín.) valor de 2.08 y 1.29 en los ítems p8 (malestar psicomotor) y p5 (cambios de apetito) del PHQ-9, respectivamente; sus parámetros de dificultad (b) se encuentran entre -0.31 (p4 [cansancio], b1, mín.) y 3.44 (p9 [pensamientos suicidas], b3, máx.). Los autores señalan que los ítems más informativos son p1 (anhedonia), p8 y aquellos menos informativos son p5, p3 (problemas de sueño); asimismo, la función de información del test se orienta al lado derecho con mayor precisión entre los valores teta (Θ) de 0 y 4.

En China, Gao y Liu (2024), en la prueba de ajuste de cada ítem identificaron RMSEA aceptables (<0.06) en las nueve preguntas y sus parámetros b se encontraron entre 1.92 y 3.45 para los ítems p9 y p2 (ánimo depresivo); de igual forma, la dificultad del ítem se extiende desde 0.69 (p6 [sentimientos de inutilidad], b1, mín.) y 2.44 (p7 [dificultades de concentración], b3, máx.). Se adiciona que, los ítems más informativos son p2 y p7, aunque se halló funcionamiento diferencial del ítem en p2, p3, p6 y p9, encontrando que las adolescentes mujeres puntúan más alto en depresión que los varones. Ma et al. (2021), en pacientes chinos con depresión mayor, hallaron parámetros a entre 2.80 (p2) y 1.49 (p5), y sus valores b de dificultad se hallan entre -0.43 (b1, p9, mín.) y 1.68 (b3, p8, máx.). Asimismo, la función de información del test brinda mayor información en el rango Θ de -3 a 1, es decir orientada a la izquierda, a los valores $\logit \Theta$ de menor rasgo. Kitamura et al. (2023), con mujeres japonesas embarazadas, contrario a los anteriores, hallaron dificultades en el PHQ-9 y tuvieron que retirar los ítems p3, p4, p5, p9. Con la versión reducida los autores encontraron valores de discriminación entre de 2.36 y 1.31 en los ítems p6 y p8; además, su dificultad b oscila en -0.093 (b1, p1, mín.) y 3.72 (b3, p8, máx.), con una función de información ubicada a la derecha del $\logit \Theta$ promedio.

Arias et al. (2024) en un estudio realizado en 27 países europeos, hallaron que los ítems más discriminantes suelen ser el p2, p1, p6 ($a > 2.85$) mientras que el menos discriminante es el p3. La investigación multinacional mostró que en casi todos los países la función de información del test se orienta al lado derecho entre los valores teta (Θ) de 0 y 4; mientras que en la función de información del test este tiene mayor precisión a su lado derecho, entre 0 y 2. Se señala que los autores no utilizaron el ítem p9 asociado a ideas suicidas por posibles implicancias éticas. Reich et al. (2018) en población general de alemanes hallaron ítems discriminativos p2, p8 de 3.97 y 2.32, y sus valores de dificultad están en su mínimo de -0.83 (p4, b1) y máximo de 2.39 (p8, b3). Reich et al. (2018) hallaron puntuaciones Θ entre -1 y 2 al lado derecho de la función de información del test.

En América, Palmer et al. (2024) en pacientes con diversos diagnósticos de salud mental de Estados Unidos, IRT-GRM, encontraron puntajes de discriminación entre 2.16 y 1.17 en los ítems p1, p9; la dificultad se extiende entre los valores de -1.31 (p4, b1) hasta su extremo más elevado de 3.03 (p9, b3). El estudio de Palmer et al. (2024) encuentra que la función de información del test se distribuye con mayor precisión entre los límites de +2 y -2 logits, encontrándose una mayor capacidad para diferenciar los diferentes niveles del rasgo.

Barroso et al. (2019) en una comunidad rural de afrodescendientes brasileños, TRI-MRG, encontraron parámetros de discriminación entre 3.24 y 1.55 para p9 y p5 respectivamente; los parámetros de dificultad alcanzan su menor valor (-0.22, b1) en el ítem p4 y obtiene su punto más alto en 2.29 (b3) en el ítem p8. No obstante, Barroso et al., (2019) encuentran en la curva característica del ítem (CCI) p1-p9 que la tercera categoría de respuesta tiene muy baja probabilidad de respuesta y se solapa entre las demás, por lo cual consideran la posibilidad de recodificar las respuestas para mejorar la comprensión de la depresión. Damiano et al. (2023), también en población general de Brasil, IRT-MRG, hallaron parámetros de discriminación entre 2.93 (p2) y 1.55 (p1), y sus parámetros de dificultad se extienden entre su mínimo valor de 0.43 (PHQ2, b1) y su valor más alto 3.13 (p9, b3).

Lingán-Huamán et al. (2023), en universitarios peruanos, encontraron que los ítems p2, p6, p9 brindan mayor información; mientras que p3, p5, p8 son los que menor información brindan. Asimismo, los autores peruanos destacan los ítems p2 y p6 como aquellos que permiten discriminar mejor la depresión entre los evaluados. Respecto al parámetro de dificultad este se extiende entre su menor valor de -0.97 (p3, b1) hasta su máximo 3.53 (p8, b3). Carranza-Esteban et al. (2024), también en Perú, emplearon la TRI-MRG para adaptar una versión breve del PHQ-9 con cuatro ítems en una muestra de quechuahablantes, hallando un ajuste favorable (RMSEA = .088; SRMSR = .036; TLI = .97; CFI = .99) del modelo inicial. También se destaca que el parámetro de discriminación estuvo entre 1.91 y 2.83 para los ítems p3, p4, p2 y p1, y la dificultad de este se extiende en su mínimo valor (-0.339, b1) en el ítem p2 hasta su valor elevado (2.58, b3) en p1.

En el contexto clínico, si bien la teoría clásica de los test ha brindado importantes aportes, la TRI puede ser un valioso enfoque para el análisis de los instrumentos cortos cuyo interés recaiga en examinar las características de los ítems. En la búsqueda realizada en diversas bases de datos, se ha notado que son escasos los estudios del PHQ-9 desde el enfoque TRI y no se ha examinado sus características con población clínica de personas con problemas de consumo perjudicial de alcohol o dependencia. En función a lo anterior, el objetivo del presente es analizar las características psicométricas del PHQ-9 desde el enfoque TRI e identificar los ítems que permiten discriminar mejor la depresión en la muestra diana.

2. Método

2.1. Tipo de estudio

El presente es un estudio de tipo instrumental (Ato et al., 2013) ya que se brinda las características psicométricas del PHQ-9 desde el enfoque TRI, además que se muestra evidencia de validez en relación a otras variables.

2.2. Participantes

La muestra se obtuvo de la Encuesta Demográfica y de Salud Familiar (ENDES), Perú, del Instituto Nacional de Estadística e Informática (INEI, 2020,2021,2022), el cual desde el 2014 realiza una encuesta anual a través de un cuestionario que evalúa depresión, el PHQ-9, y otro adicional asociado al consumo de alcohol (CIDI). La muestra proviene de 25 departamentos del Perú de costa, sierra y selva, y 43 distritos de Lima Metropolitana. De esta, se consideró únicamente las encuestas de los años 2020, 2021 y 2022 incluyendo únicamente los participantes que consumieron alcohol durante los últimos doce meses, añadiendo los criterios del CIDI (al menos un criterio), por lo cual fueron excluidos los participantes que no cumplieron con esta consigna. En ese sentido, se obtuvo un total de 492 sujetos, pero se retiraron a ocho menores de edad. De esta muestra final (N = 484), el 82% de ellos son varones (M = 35 años, n = 396) y el 17% son mujeres (M = 30 años, n = 88); de ellos, el 9% concluyó la educación primaria, 59%, 19%, 20% y 2% poseen educación secundaria, superior no-universitaria, universitaria y posgrado, respectivamente. Asimismo, el 21.3% respondió afirmativamente a una de las preguntas del CIDI, un 34.1% respondió afirmativamente a 2 preguntas del CIDI, y un 56% de sujetos contestó a 3 o más preguntas del CIDI. Se señala que el brindar dos o más respuestas afirmativas en el CIDI implican tener dos o más criterios de consumo problemático de alcohol; por lo cual, excluyendo del total el 21.3% con una sola respuesta, se tiene un 78.7% de sujetos que respondieron a dos o más preguntas del CIDI, lo cual implicaría dos o más criterios diagnósticos según el DSM-V-TR. Este grupo tendría un posible trastorno por consumo de alcohol (AUD), según los criterios del DSM-V-TR.

2.3. Instrumentos

Patient Health Questionnaire (PHQ-9)

Es un instrumento elaborado por Spitzer et al., (1999) a partir de los criterios diagnósticos considerados en el DMS-IV. Este consta de nueve preguntas que evalúan diferentes criterios de depresión experimentados en las dos últimas semanas bajo la consigna “¿Qué tan seguido ha tenido molestias debido a los siguientes problemas?”. Después, se sucede el ítem p1 “Poco interés o placer en hacer cosas” y los restantes ocho. El

instrumento consta de cuatro opciones de respuesta en formato Likert donde el puntaje de 0 se asigna a la categoría “para nada”, 1 a “varios días”, 2 a “la mayoría de días”, 3 a “casi todos los días”. El PHQ-9 posee puntos de corte que consideran un tipo de depresión leve, moderada, moderadamente severa y severa cuando se obtiene puntajes entre 5-9, 10-14, 15-19 y 20-27, respectivamente. Existe una versión peruana en población general con propiedades psicométricas favorables asociadas a fiabilidad elevada ($\omega = 0.87$) y una estructura unidimensional (CFI = 0.936; SRMR = 0.039; RMSEA = 0.089), aunque obtenida en un contexto previo a la pandemia del 2020 (Villarreal-Zegarra et al., 2019).

Composite International Diagnostic Interview (CIDI)

Mide el posible consumo perjudicial o dependencia alcohólica (Kessler, 2004), y considera los criterios dados por el DSM-IV. El CIDI considera una pregunta de filtro “¿en los últimos doce meses, usted tomó bebidas alcohólicas o licor 12 veces o más? - considere todas las veces a lo largo de los 12 meses” que sirve para diferenciar a los sujetos que no han empleado bebidas alcohólicas en el año en curso. Para considerar un posible consumo problemático, luego de esta pregunta de tamización, el CIDI emplea 16 preguntas adicionales que miden, en los últimos doce meses, los criterios de incumplimiento de obligaciones, problemas sociales o interpersonales, consumo de situaciones de riesgo, deseo, tolerancia, abstinencia (2 ítems), mayor tiempo de consumo previsto (3 ítems), esfuerzos sin éxito de control del consumo, tiempo excesivo, abandono de actividades, problemas psicológicos o físicos. Los 16 ítems señalados guardan correspondencia a los once criterios diagnósticos considerados en el DSM-V-TR (APA, 2022) y para que se considere un posible AUD se considera por lo menos dos respuestas afirmativas a cualquiera de sus preguntas. No obstante, de la versión de 16 ítems del CIDI, se retiraron las preguntas “Siguió bebiendo a pesar de que le provocaba problemas con estas personas” y “Le han detenido más de una vez por alterar el orden o manejar bajo los efectos del alcohol”; la primera de estas, por redundancia, la otra, porque ya no se considerada un criterio actual. Adicionalmente, desde el DSM-IV-TR (APA, 2000), el CIDI considera 3 criterios diagnósticos de abuso de alcohol, de las cuales si se contesta una indicaría abuso. Añade 10 criterios de dependencia que exigen por lo menos tres respuestas del CIDI para considerar esta categoría.

2.4. Procedimiento

Se accedió a la encuesta de dominio público del INEI y se seleccionó la Encuesta Demográfica y de Salud Familiar en los periodos 2022, 2021 y 2020 (INEI, 2020,2022,2023), de los cuales se extrajeron únicamente los casos que habían respondido de manera afirmativa a la pregunta “¿En los últimos 12 meses, usted tomó bebidas alcohólicas o licor 12 veces o más? con la cual se filtró únicamente aquellos casos de los participantes que consumieron bebidas alcohólicas en ese periodo; de este grupo no se incluyeron a los participantes que dejaron casilleros en blanco. La encuesta está a cargo de INEI, institución que capacita a sus encuestadores para el proceso de recolección de datos en los periodos mencionados mediante una entrevista directa en los hogares de los participantes. Una vez realizada la recolección de datos el INEI publica las bases obtenidas en su página web con el acceso libre a estos, por lo cual no se necesita permiso para su uso. Como se señaló, únicamente se consideraron a los sujetos que respondieron afirmativamente indicando un consumo de alcohol en los últimos doce meses y los demás criterios del CIDI, además de eliminar participantes que no lograron completar en su totalidad tanto el PHQ-9 y el CIDI. De acuerdo a lo señalado, en el año 2022 la encuesta ENDES obtuvo 34201 participantes a nivel nacional, de los cuales se halló únicamente 245 sujetos con las características señaladas para el análisis. En periodo 2021, de 34115 encuestas, se filtraron únicamente 142 sujetos; mientras que, en el año 2020, de 34403 encuestas, solo se consideró 105 participantes. Al realizar la sumatoria de estos tres periodos se obtuvo 484 sujetos (muestra final), excluyendo 8 adolescentes. Al final, se extrajeron grupos equilibrados en cantidad para análisis adicionales entre 1) consumidores no problemáticos y AUD y 2) dependencia y abuso de sustancias para explicarse estos desde los puntajes de depresión. Luego, se procedió a realizar los análisis psicométricos empleando el software libre R-Studio.

2.5. Análisis de datos

Se analizó la estadística univariante de los 9 ítems del instrumento a partir de la asimetría, curtosis, dificultad y discriminación del ítem. Posterior a esto, se inspeccionó la matriz de correlaciones policóricas empleada en análisis factorial confirmatorio (AFC) procurando evitar colinealidad elevada ($>.85$) entre los ítems; de igual manera, en la matriz estandarizada de los residuales, se espera correlaciones menores en valor absoluto a 2.58. Luego, se puso a prueba la unidimensionalidad del modelo mediante un AFC con el método de estimación WLSMV, esperando cargas factoriales satisfactorias ($\lambda >.30$), a la vez que se esperó índices de ajuste aceptables ($SRMR \leq .08$, $CFI \geq .92$, $RMSEA \leq .08$, $TLI \geq .92$; $GFI \geq .92$) (Hair et al., 2014; MacCallum et al., 1996). Adicionalmente se brindó la consistencia interna del modelo mediante el coeficiente alfa (α) y omega (ω). Luego se aplicó el GRM de Samejima para ítems politómicos con los parámetros de dificultad y discriminación. Antes, se evaluó la unidimensionalidad del modelo con las medidas de ajuste previamente señaladas, también se evaluó el RMSEA de cada ítem esperando valores cercanos a cero ($RMSEA \leq .06$). Seguidamente, se obtuvo las cargas factoriales de los ítems del PHQ-9 obtenidas en el modelo TRI. Luego, se analizaron los parámetros de discriminación (a) según los puntos de corte sugeridos por Baker (1992) considerando los valores entre .01 - .34, .35 - .64 y .65 - 1.34, a los cuales llama muy bajo, bajo y moderado; y se señalan los valores altos o muy altos aquellos mayores a 1.35 y 1.70, respectivamente. Estos valores a indican una asociación más fuerte entre el rasgo latente y el ítem. Respecto al parámetro de dificultad (b), este valor puede caer en el rango teórico de $-4 \leq b \leq +4$ y refleja el nivel de rasgo necesario teta (θ) para una determinada probabilidad de respuesta (50%), además aquellos ítems con mayor dificultad son los que necesitarían de mayor nivel del rasgo latente medido. Como pasos siguientes se analizó la curva característica del ítem (CCI), la función de información del ítem y del test y la curva de probabilidad del ítem. Al final, se obtuvo la fiabilidad marginal del instrumento, la cual revela el grado error asociado a la medición y a la precisión del test.

Los análisis se complementaron brindando evidencia en relación a otras variables considerando a la depresión como variable explicativa de los puntajes en consumo de alcohol desde un modelo de ecuaciones estructurales (SEM). El modelo empleó el método de estimación WLSMV y se esperó un coeficiente beta (β) desde una interpretación de la magnitud del efecto (Domínguez-Lara, 2018). Asimismo, se propuso dos modelos de regresión logística para determinar si depresión predice 1) la pertenencia a la categoría AUD o consumidores no problemáticos (DSM-V-TR) y 2) la pertenencia a la categoría de abuso o de dependencia (DSM-IV). Para tal fin se consideró los ítems del CIDI respectivos a abuso, dependencia o AUD, y se equilibraron los grupos, hallándose 72 sujetos en la categoría AUD y 72 en la de no problemáticos. Para el grupo restante, se consideraron 35 en cada uno. Se añadió una comparación de medias entre 1) consumidores no problemáticos y los de posible AUD y 2) las categorías de abuso y dependencia.

2.6. Consideraciones éticas

La encuesta ENDES, al ser un estudio de relevancia nacional considera como elementos importantes los consentimientos de los participantes, además de proteger la confidencial de los datos individuales de los participantes los cuales no son divulgados, lo cual se indica en el Art. 97, del Decreto Supremo 043-2001-PCM que resguarda el anonimato de los datos tratados por el INEI; de igual manera, la investigación se realizó considerando los lineamientos brindados en la Declaración de Helsinki. En tal sentido, este se realizó respetando los diferentes lineamientos nacionales e internacionales para la recolección de datos. Asimismo, el presente cuenta con la aprobación del Vicerrectorado de Investigación e Innovación de la Universidad de San Martín de Porres, con el código N ° 3308-2023.

3. Resultados

Se analizó la estadística descriptiva univariante de los ítems del PHQ-9, aunque los valores de asimetría no se encontraron dentro de lo esperado, en especial en el ítem p9. Asimismo, se realizó un análisis preliminar al enfoque TRI encaminado a identificar la unidimensionalidad del modelo mediante un análisis factorial

confirmatorio del PHQ-9 a partir de la matriz policórica, obteniéndose cargas factoriales elevadas en los ítems ($\lambda > .70$). No obstante, la correlación de los residuales en la diagonal superior (Tabla 1) indican valores elevados (> 2.58) entre los pares p1-p2, p1-p6, p3-p5, p4-p9, p6-p9 por lo que se consideró ítems problemáticos al p1 y p9. Además, la correlación entre los ítems p6-p9 debajo de la diagonal, alertan de una moderada colinealidad entre las variables. Adicionalmente se obtuvo los índices de ajuste encontrándose valores aceptables (RMSEA= 0.079 [IC 90% 0.064-0.095], SRMR = 0.074, CFI = 0.965, GFI = 0.993) y consistencia interna elevada ($\alpha = .84$; $\omega = .85$). Se señala que los pares de ítems que tuvieron elevadas correlaciones residuales forman parte de la dimensión cognitivo-afectivo del PHQ-9, lo cual reflejaría un factor latente; no obstante, se procedió a asumir el instrumento como esencialmente unidimensional.

Tabla 1

Estadística descriptiva del PHQ-9, matriz de correlación policórica y de residuales del PHQ-9

Indicador		p1	p2	p3	p4	p5	p6	p7	p8	p9
Anhedonia	p1	1	4.46*	-1.064	0.528	-1.202	-4.349*	0.166	-2.305	-1.81
Ánimo depresivo	p2	0.71	1	-1.421	-1.522	-1.728	-0.632	-1.269	-1.362	-0.198
Problemas de sueño	p3	0.53	0.54	1	2.184	2.847*	-1.842	-1.001	0.003	-1.37
Cansancio	p4	0.61	0.6	0.61	1	0.219	-1.309	0.633	1.068	-4.067*
Cambios de apetito	p5	0.52	0.54	0.59	0.56	1	-1.89	1.06	0.11	0.173
Sentimientos de inutilidad	p6	0.43	0.58	0.46	0.53	0.46	1	0.383	1.408	6.112*
Dificultades de concentración	p7	0.55	0.53	0.47	0.56	0.54	0.54	1	1.41	-2.013
Malestar psicomotor	p8	0.5	0.55	0.52	0.6	0.53	0.6	0.56	1	-1.11
Pensamientos suicidas	p9	0.38	0.51	0.36	0.24	0.47	0.74*	0.36	0.41	1
	<i>M</i>	0.49	0.5	0.52	0.36	0.42	0.25	0.28	0.24	0.1
	<i>SD</i>	0.75	0.74	0.87	0.71	0.82	0.64	0.63	0.6	0.39
	<i>g1</i>	1.6	1.57	1.7	2.28	2.01	2.96	2.6	2.84	5.02
	<i>g2</i>	2.19	2.26	1.93	5.05	3.11	8.69	6.93	8.05	28.56
	<i>Di</i>	0.16	0.17	0.17	0.12	0.14	0.08	0.09	0.08	0.03
	<i>D</i>	0.59	0.63	0.56	0.6	0.58	0.54	0.54	0.53	0.39
	λ	0.777	0.808	0.715	0.774	0.717	0.748	0.702	0.731	0.645

Nota. *M*, media, *SD*, desviación estándar; *g1*, asimetría; *g2*, curtosis; λ , lambda; *p_n*, ítem *n* del PHQ-9; *, correlación excesiva; *Di*, dificultad; *D*, discriminación

Teoría de respuesta al ítem – modelo de Samejima

Se puso a prueba la unidimensionalidad del modelo inicial, obteniéndose índices de ajuste satisfactorios ($M2=77.68$, $df=27$, $p<0.001$, $RMSEA=0.062[0.046/0.078]$, $SRMSR=0.068$, $TLI=0.972$, $CFI=0.979$). Posteriormente, se hizo la prueba de ajuste de cada ítem de manera individual, hallándose valores aceptables (Ver Tabla 2). También se obtuvo los parámetros de discriminación (*a*) y dificultad (*b*) de todos los ítems. Los parámetros *a* son elevados y todos son superiores 1.48. De estos el ítem más discriminativo es p4 y el menos discriminativo el p9. Los parámetros de dificultad se orientan en su totalidad hacia la derecha de los valores más altos del rasgo estudiado, siendo el b1 de menor rasgo en el ítem p2. No obstante, el ítem p9 evidencia que los examinados

PROPIEDADES PSICOMÉTRICAS DEL PHQ-9 EN USUARIOS PERUANOS CON PROBLEMAS DE CONSUMO

necesitan mayores niveles de depresión para responder a la categoría “varios días” o una superior (b_1), de igual manera con los restantes parámetros b_2 y b_3 . Asimismo, todos los ítems tienen un b_3 superior a 2.079 logits (p_3) que indica mayor presencia de rasgo para responder a la última categoría “casi todos los días”.

Tabla 2

Índices de ajuste de los ítems, parámetros de discriminación y dificultad del PHQ-9

Indicador	ítem	S_{X2}	$df.S_{X2}$	$RMSEA.S_{X2}$	$p.S_{X2}$	a (e.s)	b_1 (e.s)	b_2 (e.s)	b_3 (e.s)
Anhedonia	p1	23.13	20.00	0.02	0.28	2.179(.235)	0.438(.075)	1.711(.136)	2.393(.198)
Ánimo depresivo	p2	19.62	19.00	0.01	0.42	2.411(.263)	0.383(.071)	1.717(.13)	2.336(.186)
Problemas de sueño	p3	29.58	27.00	0.01	0.33	1.834(.203)	0.597(.084)	1.634(.141)	2.079(.181)
Cansancio	p4	34.67	19.00	0.04	0.02	2.442(.287)	0.802(.08)	1.879(.147)	2.194(.178)
Cambios de apetito	p5	38.81	28.00	0.03	0.08	1.978(.232)	0.859(.09)	1.65(.142)	2.143(.189)
Sentimientos de inutilidad	p6	17.27	19.00	0.00	0.57	1.995(.268)	1.272(.116)	2.136(.199)	2.533(.247)
Dificultades de concentración	p7	16.63	20.00	0.00	0.68	1.946(.243)	1.11(.104)	2.186(.197)	2.738(.266)
Malestar psicomotor	p8	35.01	21.00	0.04	0.03	2.144(.279)	1.25(.108)	2.071(.181)	2.731(.264)
Pensamientos suicidas	p9	12.59	11.00	0.02	0.32	1.493(.286)	2.245(.296)	3.344(.51)	3.861(.629)

Nota. p_n , ítem n del PHQ-9; a , discriminación del ítem; b_n , dificultad del ítem en la categoría de respuesta k_{n+1} o una superior; RMSEA, Root mean squared error of approximation; e.s, error estándar; en negrita, valores máximos y mínimos.

Curva característica del ítem

Respecto a las curvas características de los ítems del PHQ-9, se encuentra que las dos categorías intermedias (“varios días”, “la mayoría de días”) tienen una menor probabilidad de ser contestadas; en especial la categoría “la mayoría de días” que incluso tiene menor probabilidad de ser respondida a comparación de la categoría anterior “varios días” bajo los mismos niveles Θ . En todos los casos la probabilidad de elegir las categorías mayores a “para nada” se incrementan cuando sus valores Θ se encuentran al lado derecho de la curva, lo cual indica mayores niveles de depresión para ser elegidas (Figura 1).

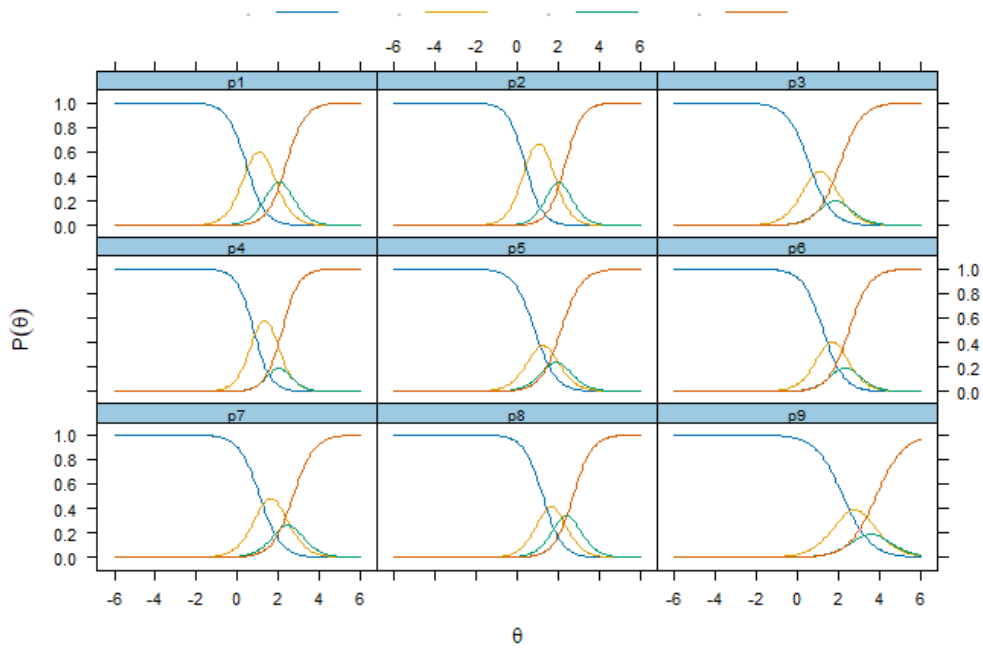


Figura 1. Curva característica de los ítems del PHQ-9

Curva de información del ítem

La curva de información del ítem hace referencia a la precisión que tienen los ítems para obtener la mayor cantidad de información del rasgo latente. En el presente caso, los ítems tienden a ser más precisos en el rango θ que se extiende entre 0 y 2 logits para, posteriormente, ir decayendo. De los nueve ítems, el que brinda menor información es el ítem p9 y el que brinda más es el ítem p4 (Figura 2).

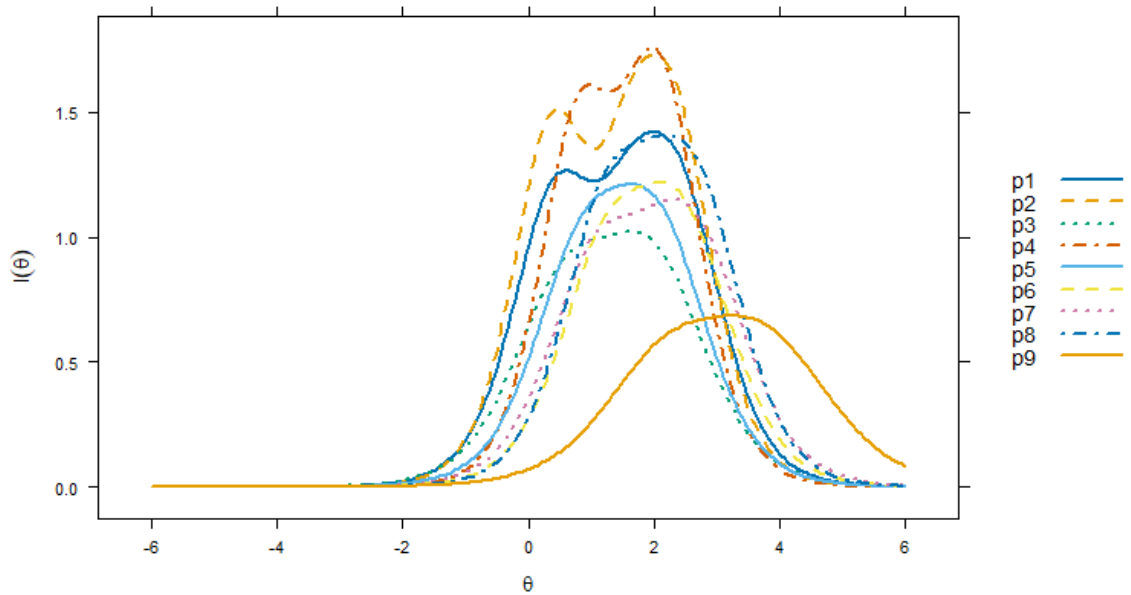


Figura 2. Curva de información de los ítems del PHQ-9

Curva de probabilidad acumulada o de umbral

Se observa que entre la tercera (“la mayoría de días”) y cuarta categoría (“casi todos los días”) del ítem p9 (Pensamientos de que estaría mejor muerto(a) o de lastimarse de alguna manera) hay una diferencia de probabilidad menor (Figura 3); es decir, los examinados tienen similares probabilidades de respuesta a tales categorías, lo cual también se ha reflejado en el solapamiento de la CCI del ítem p9; a la vez, sucede de la misma manera con los ítems restantes. En el caso del b1 del p9 se necesita un 2.245 de Θ para tener un .50 de probabilidad de responder esa categoría, pero en el caso de b2, y b3 (3.344, 3.861) estos logits son bastante cercanos. Asimismo, la curva de información del test muestra que el instrumento es más preciso para aquellos casos en los cuales se obtiene mayores niveles de depresión; de igual manera, a mayor rasgo latente, disminuye el error estándar, lo cual se refleja en la confiabilidad marginal de 0.72 que se obtuvo (Figura 4).

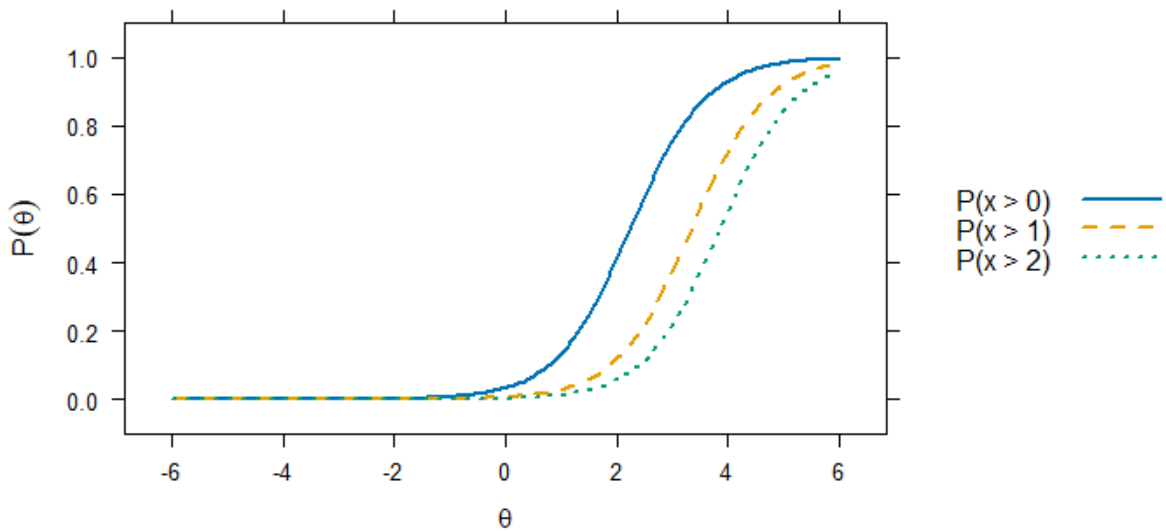


Figura 3. Curva de probabilidad del ítem p9

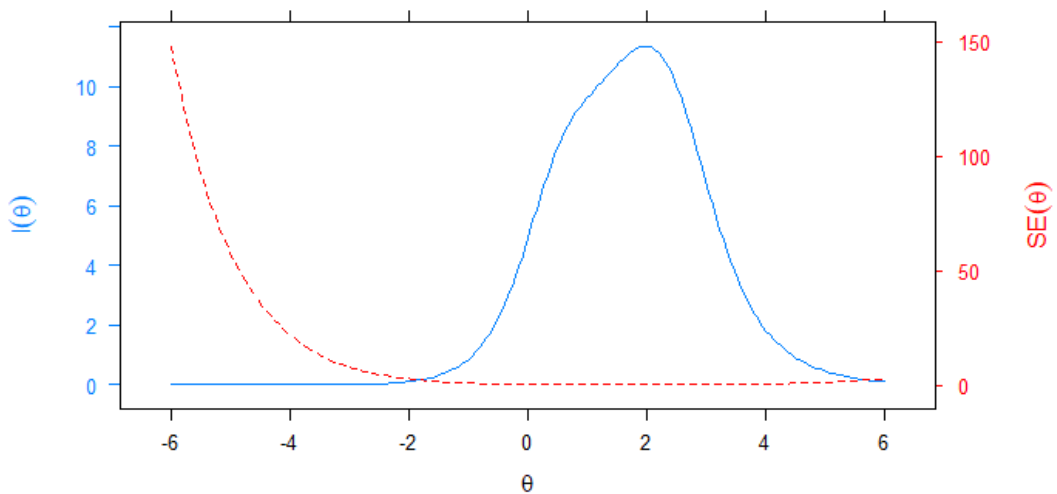


Figura 4. Curva de información del test y error estándar

Evidencia de validez en relación al consumo de alcohol

Adicionalmente, desde un modelo SEM, se propuso explicar el consumo de alcohol en la muestra en referencia a partir del constructo latente de depresión. Los resultados evidencian un efecto directo de magnitud mínimo necesaria ($\beta = 0.24$, $p < 0.001$) que implica que a mayores niveles de depresión se encontrará mayores niveles de consumo de alcohol. Adicionalmente, los índices de ajuste del modelo fueron adecuados (RMSEA= 0.069 [IC 90% 0.055-0.083], SRMR = 0.063, CFI = 0.966, GFI = 0.994).

Análisis de regresión logística según el predictor de depresión

Como análisis final, a partir de la dicotomización del consumo según las categorías del DSM-IV-TR, se halló que los mayores niveles depresión incrementan la probabilidad de pertenencia al grupo de dependencia ($\beta = .37$, $p < .001$), mientras que menores niveles, con abuso de sustancias. En un segundo modelo de regresión, según las categorías del DSM-V-TR, la probabilidad también se incrementa para el grupo con AUD, mientras que los niveles menores implican consumo no problemático ($\beta = .47$, $p < .001$). Adicionalmente, se comparó las medias de depresión de los consumidores no problemáticos ($M = 1.54$) con los que tienen AUD ($M = 4.32$) mediante la U de Mann-Whitney ($U = 1784$, $p < 0.001$) encontrándose diferencias significativas. De igual manera, se comparó abuso ($M = 2.34$) y dependencia ($M = 5.63$), aunque sin diferencias significativas ($U = 491$, $p = 0.14$).

4. Discusión

Los resultados hallados en el presente estudio reflejan similitudes a los estudios revisados en párrafos anteriores al evidenciar que el PHQ-9 tiene propiedades psicométricas apropiadas para su uso en muestra de usuarios con problemas de abuso o dependencia alcohólica.

Se identificó correlaciones elevadas entre los residuales estandarizados de los pares de ítems más vinculados a la dimensión afectivo- cognitivo del PHQ-9, aunque esta no fue considerada ya que diversos estudios consideran al PHQ-9 como medida unidimensional (Arias et al., 2024; Barroso et al., 2019; Cumbe et al., 2020; Gao y Liu, 2024; Kitamura et al., 2023, Palmer et al., 2024), además el AFC realizado con la propuesta unidimensional mostró un ajuste adecuado, de igual manera la prueba de unidimensionalidad del modelo TRI resultó favorable a esta.

Respecto al modelo TRI planteado, se destaca que el ítem p9 (suicidio) obtuvo un parámetro a de 1.49, que indica que las ideas de suicidio no necesariamente son un requisito obligatorio del cuadro afectivo ya que este alterna con la presencia de otros criterios adicionales establecidos en el DSM-V-TR (APA, 2022). Adicionalmente, la tasa del suicidio es del 15% en los pacientes con depresión (Hernández -Vásquez et al., 2016) y del 18% para los consumidores de alcohol, lo cual se estaría reflejando en ese menor nivel discriminativo (OPS, 2021). Llama la atención que el p4 (fatiga) haya obtenido el nivel de discriminación más alto, lo cual indicaría un ítem relevante para la evaluación de la depresión en la muestra estudiada, incluso si se considera que los estudios peruanos, realizados con diferentes muestras, no encontraron a este como el más discriminante (Carranza-Esteban et al., 2024; Lingán- Huamán et al., 2023), tampoco en otros (Arias et al., 2024; Barroso et al., 2019; Cumbe et al., 2020; Gao y Liu, 2024; Kitamura et al., 2023, Palmer et al., 2024). En cuanto al parámetro de dificultad, el p9 resultó el ítem más difícil, lo cual hace referencia a su baja prevalencia (Hernández -Vásquez et al., 2016; OPS, 2021) y sobre todo porque es indicador de los episodios de mayor gravedad. Asimismo, el ítem de menor dificultad fue el p2 (ánimo depresivo), relevante para captar el dominio alusivo al rasgo en discusión en los evaluados con menor presencia del rasgo. Asimismo, el parámetro b_3 de todos los ítems necesita un nivel de rasgo importante ($\Theta > 2$) para elegir la opción “casi todos los días”, asegurándose con esto una mejor medición de los casos de los sujetos con mayor presencia del rasgo.

Algunas observaciones adicionales aparecieron al hallar que, en la CCI de todos los ítems, estos mostraron una baja probabilidad de respuesta a la categoría “la mayoría de días”. Probablemente, su denominación no logra

diferenciarse de la última “casi todos de días”, esta última con mayor probabilidad de elección que la anterior. Barroso et al. (2019) en Brasil, hallaron algo similar, por lo cual se sugeriría eliminar esta categoría o mantener la denominación más cercana a la utilizada originalmente “More than half the days” que, al parecer, se aleja de la denominación brindada en la encuesta realizada por el INEI (2021,2022,2023). Respecto a la curva de información del ítem, se halló que el p2 es más informativo, lo cual también se ha encontrado en otros estudios (Lingan-Huamán, 2023; Gao y Liu, 2023), también lo fue el ítem p4. No obstante, el menos informativo fue el p9, probablemente por su menor variabilidad.

La curva de información del test (Figura 4) señala que el PHQ-9 tiene menor error de medición cuando evalúa a la muestra de estudio con mayores niveles de depresión, aunque la probabilidad de error se incrementa en los casos de los sujetos con menores niveles del rasgo. Lo anterior tiene sentido ya que, al tener una muestra clínica de consumidores problemáticos o dependientes de alcohol, estos tienen mayor probabilidad de poseer el rasgo latente, además el PHQ-9 cumple su cometido de diferenciar a los casos que tienen niveles más elevados de depresión de los que tienen menor presencia del rasgo, por lo cual se espera que, al evaluarse un trastorno de menor prevalencia en la población, el instrumento se muestre más preciso en los niveles más elevados de este y no lo contrario. Asimismo, este tipo de curva se ha hallado en otros estudios similares en los cuales el PHQ-9 tiene mayor precisión de captar la variabilidad de los niveles de depresión en sus niveles más elevados (Arias et al., 2024; Barroso et al., 2019; Cumbe et al., 2020; Gao y Liu, 2024; Kitamura et al., 2023, Palmer et al., 2024).

Finalmente, como se esperaba, el modelo SEM, indica la asociación entre la depresión y el mayor consumo de alcohol en la muestra (Desalu et al., 2017; Grewal et al., 2022; Keum y Cano, 2023; Villanueva-Blasco et al., 2022). De igual manera, la regresión logística indica que los mayores niveles de depresión logran asociarse a una mayor probabilidad de dependencia o de AUD. Adicionalmente, al compararse las medias entre consumidores no problemáticos y los que tienen AUD, también se observó diferencias de sus medias, más altas para el grupo con AUD, lo cual afirma la asociación entre el consumo y depresión o cuadros de salud mental (Desalu et al., 2017; Grewal et al., 2022; Keum y Cano, 2023; Villanueva-Blasco et al., 2022).

La depresión es uno de los problemas más importantes en el país andino (INSM, 2024); además su sistema de salud aqueja de una sobrecarga de atenciones de salud mental. Esta situación se agrava en la atención de los usuarios con problemas de uso de alcohol, por lo cual el contar con un instrumento de tamización rápido de depresión en estos será de mucha utilidad dada la demanda, más aún si este se ha desarrollado desde un enfoque riguroso como la TRI. De igual manera, a partir de la asociación entre la depresión y el uso de alcohol, los profesionales de salud mental deben tomar acciones preventivas de la comorbilidad de estos trastornos. En los casos en los cuales se ha identificado la comorbilidad, lo recomendable sería considerar tratamientos de depresión basados en evidencia.

4.1. Limitaciones

La muestra evaluada tiene una proporción mayor de hombres sobre las mujeres, por lo cual no se pudo realizar análisis asociados al funcionamiento diferencial del ítem entre ambos grupos. A pesar de realizar una búsqueda exhaustiva de mujeres con consumo de alcohol, no se logró identificar este tipo de patrón en la encuesta del INEI de los años 2019, 2018 y 2017. Tampoco se pudo obtener análisis adicionales del grupo con puntajes elevados en el PHQ-9 que los ubique en la categoría de depresión, lo cual debería realizarse en futuros análisis. Asimismo, a la evidencia de la elevada correlación de residuales, no se tomó en cuenta la naturaleza bifactorial del constructo, aunque lo anterior rebasa los objetivos del estudio. Adicionalmente, al ser la encuesta ENDES un tipo de medición en un contexto no clínico, no se pudo acceder a mayor información que pueda dar seguridad respecto al posible AUD de los entrevistados. Si bien se consideraron los criterios diagnósticos del DSM-V-TR (o del DSM-IV) expresados en el CIDI, estos no necesariamente involucran un tipo de diagnóstico, aunque el grupo analizado no deja de ser una población de elevado riesgo. Tampoco se cuenta con alguna “gold standard” asociada a un diagnóstico de depresión o AUD.

4.2. Conclusiones

El PHQ-9 es un instrumento con propiedades psicométricas satisfactorias para evaluar la depresión y es útil para su medida en los niveles más altos de este tipo de trastorno en la muestra de estudio. Asimismo, a pesar de que se evalúa este rasgo en comorbilidad con alcohol, el PHQ-9 logra ser una medida exhaustiva que permite discriminar a los sujetos con mayor presencia del rasgo. Adicionalmente, como es de esperarse, el constructo latente de depresión tiene una asociación con el consumo de alcohol y sus mayores niveles predicen la pertenencia a los grupos con mayor consumo o diagnóstico más severo.

4.3. Agradecimientos

Al Director JMFA, SADL y equipo de investigación de la FCCTP de la Universidad de San Martín de Porres.

4.4. Conflicto de interés

Los autores declaran no tener ningún conflicto de interés.

4.5. Financiación

El presente estudio fue autofinanciado por los investigadores..

4.6. Contribución de los autores

ACC: Conceptualización, metodología, análisis formal, curación, redacción y supervisión. ZBS: Metodología, curación, redacción del borrador original – revisión y edición y visualización. EOM: Metodología, curación, redacción del borrador original – revisión y edición y visualización.

4.7. Declaración de disponibilidad de los datos

La base de datos empleada para el estudio se encuentra disponible en repositorio OSF, accesible a través del siguiente enlace:<https://doi.org/10.17605/OSF.IO/PGYTS>

5. Referencias

- American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders*, 4th Edition, Text Revision. American Psychiatric Association
- American Psychiatric Association (2022). *Diagnostic and Statistical Manual of Mental Disorders*, 5th Edition, Text Revision. American Psychiatric Association
- Abad, F. J., Olea, J., Ponsoda, J., & García, C. (2011). *Medición en ciencias sociales y de la salud*. Síntesis.
- Arias de la Torre, J., Vilagut, G., Ronaldson, A., Valderas, J. M., Bakolis, I., Dregan, A., Molina, A. J., Navarro-Mateu, F., Pérez, K., Bartoll-Roca, X., Elices, M., Pérez-Sola, V., Serrano-Blanco, A., Martín, V., & Alonso, J. (2023). Reliability and cross-country equivalence of the 8-item version of the Patient Health Questionnaire (PHQ-8) for the assessment of depression: results from 27 countries in Europe. *The Lancet regional health. Europe*, 31, 100659. <https://doi.org/10.1016/j.lanepe.2023.100659>
- 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. <https://doi.org/10.6018/analesps.29.3.178511>
- Attorresi, H., Lozzia, G.S., Abal, F.J.P, Galibert M.S., & Aguerri, M.E. (2009). Teoría de Respuesta al ítem. Conceptos Básicos y aplicaciones para La medición de constructos psicológicos. *Revista Argentina de Clínica Psicológica* 18, 179-188. https://www.cienciarred.com.ar/ra/usr/35/825/racp_xviii_2_pp179_188.pdf
- Baker, F.B. (1992). *Item Response Theory: Parameter Estimation Techniques*. Marcel Dekker.

- Barroso, S. M., Melo, A. P. S., da Silva, M. A., & Guimarães, M. D. C. (2019). Evaluation of the Brazilian version of Patient Health Questionnaire (PHQ-9) in Quilombola population using the Item Response Theory. *Salud Mental*, 42(1), 43–50. <https://doi.org/10.17711/SM.0185-3325.2019.006>
- Carranza-Esteban, R. F., Mamani-Benito, O., Cjuno, J., Tito-Betancur, M., Caycho-Rodríguez, T., Vilca, L. W., Torales, J., & Barrios, I. (2024). Adaptation and Validation of the Patient Health Questionnaire for Depression and Anxiety (PHQ-4) in a sample of Quechua-speaking Peruvians. *Medicina Clínica y Social*, 8(1), 63–74.
- Cumbe, V. F. J., Muanido, A., Manaca, M. N., Fumo, H., Chiruca, P., Hicks, L., De Jesus Mari, J., & Wagenaar, B. H. (2020). Validity and item response theory properties of the Patient Health Questionnaire-9 for primary care depression screening in Mozambique (PHQ-9-MZ). *BMC Psychiatry*, 20(1). <https://doi.org/10.1186/s12888-020-02772-0>
- Damiano, R. F., Hoffmann, M. S., Gosmann, N. P., Pan, P. M., Miguel, E. C., & Salum, G. A. (2023). Translating measurement into practice: Brazilian norms for depressive symptom assessment with the Patient Health Questionnaire (PHQ-9). *Brazilian Journal of Psychiatry*, 45(4), 310–317. <https://doi.org/10.47626/1516-4446-2022-2945>
- Desalu, J. M., Kim, J., Zaso, M. J., Corriders, S. R., Loury, J. A., Minter, M. L., & Park, A. (2019). Racial discrimination, binge drinking, and negative drinking consequences among black college students: serial mediation by depressive symptoms and coping motives. *Ethnicity and Health*, 24(8), 874–888. <https://doi.org/10.1080/13557858.2017.1380170>
- Dominguez-Lara, S. (2018). Magnitud del efecto, una guía rápida. *Educación Médica*, 19(4), 251-254. <https://doi.org/10.1016/j.edumed.2017.07.002>
- Gao, X., & Liu, Z. (2024). Analyzing the psychometric properties of the PHQ-9 using item response theory in a Chinese adolescent population. *Annals of General Psychiatry*, 23(1). <https://doi.org/10.1186/s12991-024-00492-3>
- Grewal, G., Polisetty, T., Cannon, D., Ardeljan, A., Vakharia, R. M., Rodriguez, H. C., & Levy, J. C. (2022). Alcohol Abuse, Morbid Obesity, Depression, Congestive Heart Failure, and Chronic Pulmonary Disease are Risk Factors for 90-Day Readmission After Arthroscopic Rotator Cuff Repair. *Arthroscopy, Sports Medicine, and Rehabilitation*, 4(5), e1683–e1691. <https://doi.org/10.1016/j.asmr.2022.06.015>
- Hair, J., Black, W., Babin, B., & Anderson, R. (2014). *Multivariate data analysis. (7th Edition)*. Pearson.
- Hernández-Vásquez, M., Azañedo, D., Rubilar-González, J., Huarez, B., & Grendas, L. (2016). Evolución y diferencias regionales de la mortalidad por suicidios en el Perú, 2004-2013. *Revista Peruana de Medicina Experimental y Salud Pública*, 33(4), 751-757. <https://dx.doi.org/10.17843/rpmesp.2016.334.2562>
- Instituto Nacional de Salud Mental (2024). *La depresión y el suicidio: mitos y creencias*. INSM. <https://www.insm.gob.pe/oficinas/comunicaciones/notasdeprensa/2012/006.html>
- Instituto Nacional de Estadística e Informática (2022). *Microdatos: Base de datos de la Encuesta Demográfica y de Salud Familiar—ENDES Perú*. Instituto Nacional de Estadística e Informática. <http://iinei.inei.gob.pe/microdatos/>
- Kessler, R. C., & Üstün, T. B. (2004). The World Mental Health (WMH) Survey Initiative Version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *International journal of methods in psychiatric research*, 13(2), 93–121. <https://doi.org/10.1002/mpr.168>
- Keum, B. T. H., & Cano, M. A. (2023). Online racism, depressive and anxiety symptoms, coping-related drinking motives, and alcohol use severity among Black, Latina/o/x, and Asian emerging adults. *Addictive Behaviors*, 136. <https://doi.org/10.1016/j.addbeh.2022.107468>
- Kitamura, T., Usui, Y., Wakamatsu, M., Minatani, M., & Hada, A. (2023). What Are The Core Symptoms of Antenatal Depression? A Study Using Patient Health Questionnaire-9 among Japanese Pregnant Women in the First Trimester. *Healthcare (Switzerland)*, 11(10). <https://doi.org/10.3390/healthcare11101494>

- Lingán-Huamán, S. K., Dominguez-Lara, S., Alarcón-Almeyda, M. S., de Moura, G. B., & Paiva, T. T. (2023). Psychometric Properties of the Patient Health Questionnaire-9 in Peruvian University Students. *Psihologijske Teme*, 32(3), 451–470. <https://doi.org/10.31820/pt.32.3.3>
- Ma, S., Yang, J., Yang, B., Kang, L., Wang, P., Zhang, N., Wang, W., Zong, X., Wang, Y., Bai, H., Guo, Q., Yao, L., Fang, L., & Liu, Z. (2021). The Patient Health Questionnaire-9 vs. the Hamilton Rating Scale for Depression in Assessing Major Depressive Disorder. *Frontiers in Psychiatry*, 12. <https://doi.org/10.3389/fpsyg.2021.747139>
- MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, 1(2), 130–149. <https://doi.org/10.1037/1082-989X.1.2.130>
- McHugh, R. K., & Weiss, R. D. (2019). Alcohol Use Disorder and Depressive Disorders. *Alcohol research: current reviews*, 40(1), arcr.v40.1.01. <https://doi.org/10.35946/arcr.v40.1.01>
- Ministerio de Educación (2023). *Noticias*. MINEDU. <https://juventud.gob.pe/2023/06/el-32-3-de-jovenes-de-15-a-29-anos-presentaron-algun-problema-de-salud-mental/>
- Organización Mundial de la Salud (2023). *Depresión*. WHO. <https://www.who.int/es/news-room/fact-sheets/detail/depression>
- Organización Mundial de la Salud (2017). *Depression and Other Common Mental Disorders*. WHO. <https://www.who.int/publications/i/item/depression-global-health-estimates>
- Organización Panamericana de la Salud (2021). *Serie – alcohol y suicidio*. OPS. <https://www.paho.org/es/documentos/serie-alcohol-alcohol-suicidio>
- Palmer, E. O. C., Ker, S., Rentería, M. E., Carmody, T., & Rush, A. J. (2024). Psychometric evaluation and linking of the PHQ-9, QIDS-C, and VQIDS-C in a real-world population with major depressive disorder. *Neuropsychiatric Disease and Treatment*, 20, 671–687. <https://doi.org/10.2147/NDT.S444223>
- Reich, H., Rief, W., Brähler, E., & Mewes, R. (2018). Cross-cultural validation of the German and Turkish versions of the PHQ-9: An IRT approach. *BMC Psychology*, 6(1). <https://doi.org/10.1186/s40359-018-0238-z>
- Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph Supplement*, 34(4, Pt. 2), 100.
- Santomauro, D.F.; Herrera, A.M.M.; Shadid, J.; Zheng, P.; Ashbaugh, C.; Pigott, D.M.; Abbafati, C.; Adolph, C.; Amlag, J.O.; Aravkin, A.Y.; et al. (2021). Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. *Lancet*, 398, 1700–1712. <https://doi.org/10.3390/jcm13185423>
- Spitzer, R. L., Kroenke, K., & Williams, J. B. (1999). Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. *JAMA*, 282(18), 1737–1744. <https://doi.org/10.1001/jama.282.18.1737>
- Villanueva-Blasco, V. J., Mateu-Mollá, J., Villanueva-Silvestre, V., & Vázquez-Martínez, A. (2022). Relationship Between Depression and Risky Alcohol Consumption in Women: the Mediating Role of Coping Styles and Age. *International Journal of Mental Health and Addiction*. <https://doi.org/10.1007/s11469-022-00931-w>
- Villarreal-Zegarra, D., Copez-Lonzoy, A., Bernabé-Ortiz, A., Melendez-Torres, G. J., & Bazo-Alvarez, J. C. (2019). Valid group comparisons can be made with the Patient Health Questionnaire (PHQ-9): A measurement invariance study across groups by demographic characteristics. *PLoS one*, 14(9), e0221717. <https://doi.org/10.1371/journal.pone.0221717>
- Wang, H., Zhu, Y., Shi, J., Huang, X., & Zhu, X. (2022). Time perspective and family history of alcohol dependence moderate the effect of depression on alcohol dependence: A study in Chinese psychiatric clinics. *Frontiers in psychology*, 13, 903535. <https://doi.org/10.3389/fpsyg.2022.903535>

Yeo, Y., Jeong, S. M., Shin, D. W., Han, K., Yoo, J., Yoo, J. E., & Lee, S. P. (2022). Changes in Alcohol Consumption and Risk of Heart Failure: A Nationwide Population-Based Study in Korea. *International Journal of Environmental Research and Public Health*, 19(23). <https://doi.org/10.3390/ijerph192316265>

Yoshimura, A., Matsushita, S., Kimura, M., Yoneda, J. ichi, Maesato, H., Yokoyama, A., & Higuchi, S. (2022). Influence of ADHD, especially attention-deficit characteristics, on the course of alcohol-dependent individuals. *BMC Psychiatry*, 22(1). <https://doi.org/10.1186/s12888-022-04455-4>