



**Universidad Nacional Mayor de San Marcos**

**Universidad del Perú. Decana de América**

Dirección General de Estudios de Posgrado

Facultad de Ciencias Matemáticas

Unidad de Posgrado

**Efecto del tamaño de muestra en los índices de ajuste  
del análisis factorial confirmatorio sobre el  
cuestionario PHQ-9 aplicado a la población peruana**

**TESIS**

Para optar el Grado Académico de Magíster en Estadística

Matemática

**AUTOR**

Moises Enrique TARAZONA COCHACHIN

**ASESOR**

Dra. Zoraida Judith HUAMÁN GUTIÉRREZ

Lima, Perú

2022



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

## Referencia bibliográfica

---

Tarazona, M. (2022). *Efecto del tamaño de muestra en los índices de ajuste del análisis factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana*. [Tesis de maestría, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Matemáticas, Unidad de Posgrado]. Repositorio institucional Cybertesis UNMSM.

---

## Metadatos complementarios

| <b>Datos de autor</b>            |   |
|----------------------------------|---|
| Nombres y apellidos              | Moises Enrique Tarazona Cochachin   |
| Tipo de documento de identidad   | DNI   |
| Número de documento de identidad | 43491114  |
| URL de ORCID                     | <a href="https://orcid.org/0000-0002-0045-2491">https://orcid.org/0000-0002-0045-2491</a> |
| <b>Datos de asesor</b>           |   |
| Nombres y apellidos              | Zoraida Judith Huamán Gutiérrez   |
| Tipo de documento de identidad   | DNI   |
| Número de documento de identidad | 09890094  |
| URL de ORCID                     | <a href="https://orcid.org/0000-0002-5533-2454">https://orcid.org/0000-0002-5533-2454</a> |
| <b>Datos del jurado</b>          |   |
| <b>Presidente del jurado</b>     |   |
| Nombres y apellidos              | Roger Pedro Norabuena Figueroa  |
| Tipo de documento                | DNI   |
| Número de documento de identidad | 41493243  |
| <b>Miembro del jurado 1</b>      |   |
| Nombres y apellidos              | Helfer Joel Molina Quiñones   |
| Tipo de documento                | DNI   |
| Número de documento de identidad | 42332500  |
| <b>Miembro del jurado 2</b>      |   |
| Nombres y apellidos              | Jorge Luis Llanos Tizado  |
| Tipo de documento                | DNI   |
| Número de documento de identidad | 17848381  |
| <b>Datos de investigación</b>    |   |
| Línea de investigación           | A.3.2.1.Analisis Multivariante  |

|  |   |
|--|---|
| Grupo de investigación                                 | No aplica   |
| Agencia de financiamiento                              | Sin financiamiento.   |
| Ubicación geográfica de la investigación               | País: Perú<br>Departamento: Lima<br>Provincia: Lima<br>Distrito: Lima<br>Latitud: -12.0561558<br>Longitud: -77.0845199            |
| Año o rango de años en que se realizó la investigación | 2020 - 2021   |
| URL de disciplinas OCDE                                | Estadísticas, Probabilidad<br><a href="https://purl.org/pe-repo/ocde/ford#1.01.03">https://purl.org/pe-repo/ocde/ford#1.01.03</a> |



**UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS**

Universidad del Perú, DECANA DE AMÉRICA

**FACULTAD DE CIENCIAS MATEMÁTICAS  
VICEDECANATO DE INVESTIGACIÓN Y POSGRADO  
UNIDAD DE POSGRADO**

**ACTA DE SUSTENTACIÓN VIRTUAL DE TESIS PARA OPTAR EL GRADO  
ACADÉMICO DE MAGÍSTER EN ESTADÍSTICA MATEMÁTICA**

Siendo las, 11:00 horas del día Viernes Cuatro de noviembre de dos mil veintidós, en la sala virtual meet.google.com/xmo-ykcm-nhr, el Jurado de Tesis conformado por los siguientes docentes:

|                 |   |                                      |
|-----------------|---|--------------------------------------|
| PRESIDENTE      | : | Dr. Roger Pedro Norabuena Figueroa   |
| MIEMBRO         | : | Dr. Helfer Joel Molina Quiñones      |
| MIEMBRO EXTERNO | : | Mg. Jorge Luis Llanos Tiznado        |
| ASESORA         | : | Dra. Zoraida Judith Huamán Gutiérrez |

se reunieron para la sustentación de la tesis titulada: «**EFFECTO DEL TAMAÑO DE MUESTRA EN LOS ÍNDICES DE AJUSTE DEL ANÁLISIS FACTORIAL CONFIRMATORIO SOBRE EL CUESTIONARIO PHQ-9 APLICADO A LA POBLACIÓN PERUANA**», presentada por el Señor Bachiller Moises Enrique Tarazona Cochachin, egresado de la Maestría en Estadística Matemática.

Concluida la exposición, los miembros del Jurado de Tesis procedieron a formular sus preguntas que fueron absueltas por el graduando; acto seguido se procedió a la evaluación correspondiente, según tabla adjunta, habiendo obtenido el Señor Bachiller **Moises Enrique Tarazona Cochachin**, el calificativo de MUY BUENO (18).

Habiendo sido aprobada la sustentación de la Tesis, el Jurado Evaluador recomienda para que el Consejo de Facultad apruebe el otorgamiento del Grado Académico de **Magíster de Maestría en Estadística Matemática** al Bachiller **Moises Enrique Tarazona Cochachin**.

Siendo las 11:53 horas, se levantó la sesión, firmando para constancia la presente Acta:

Dr. Roger Pedro Norabuena Figueroa  
**PRESIDENTE**

Dr. Helfer Joel Molina Quiñones  
**MIEMBRO**

Mg. Jorge Luis Llanos Tiznado  
**MIEMBRO EXTERNO**

Dra. Zoraida Judith Huamán Gutiérrez  
**ASESOR**



"Año del Bicentenario del Perú: 200 años de independencia "

## Universidad Nacional Mayor de San Marcos

Universidad del Perú. Decana de América

### FACULTAD DE CIENCIAS MATEMÁTICAS

### VICEDECANATO DE INVESTIGACIÓN Y POSGRADO

### UNIDAD DE POSGRADO

---

## Informe de Evaluación de Originalidad

Vicedecano de Investigación y Posgrado: Dr. Eugenio Cabanillas Lapa

Operador: Mag. José Luis Acuña Guillermo

Tesis: "EFECTO DEL TAMAÑO DE MUESTRA EN LOS ÍNDICES DE AJUSTE DEL ANÁLISIS FACTORIAL CONFIRMATORIO SOBRE EL CUESTIONARIO PHQ-9 APLICADO A LA POBLACIÓN PERUANA"

Tesista: MOISES ENRIQUE TARAZONA COCHACHIN

Recepción de documento: 26/03/2022

Aplicación del software: 04/04/2022

Software utilizado: Turnitin

Configuración del software:

- Excluir material bibliográfico
- Excluir material citado
- Excluir fuentes pequeñas N° palabras: 40
- Sin depósito

Resultado del análisis con el software: diez por ciento (10%)

Fuentes de originales de las similitudes encontradas:

- Ver abajo

Documento cumple con criterios de originalidad.

- **El documento cumple los criterios de originalidad. Sin observaciones.**

Fecha del informe: 05/04/2022



Firmado digitalmente por  
CABANILLAS LAPA Eugenio FAU  
20148092282 soft  
Motivo: Soy el autor del documento  
Fecha: 05.04.2022 13:08:36 -05:00

**Dr. Eugenio Cabanillas Lapa**

**VICEDECANO DE INVESTIGACIÓN Y POSGRADO**

## **Dedicatoria**

Dedicado a mis padres

Mis hijos

y a mi esposa

## **Agradecimiento**

Agradecimiento especial a mi asesora Mag. Zoraida Judith Huamán Gutiérrez, por su apoyo incondicional en la culminación del presente trabajo de tesis.

A los docentes de la Facultad de Ciencias Matemáticas y en particular a los docentes de la maestría en Estadística Matemática, por inculcarme sus conocimientos en bien de la sociedad.



# ÍNDICE GENERAL

|   |    |
|---|----|
| CAPÍTULO 1: INTRODUCCIÓN .....                                    | 11 |
| 1.1 Situación Problemática.....                                   | 11 |
| 1.2 Formulación Problemática.....                                 | 12 |
| 1.2.1 Problema general .....                                      | 12 |
| 1.2.2 Problemas específicos.....                                  | 12 |
| 1.3 Justificación.....  | 13 |
| 1.3.1 Justificación teórica .....                                 | 13 |
| 1.3.2 Justificación práctica.....                                 | 13 |
| 1.4 Objetivos.....  | 13 |
| 1.4.1 Objetivo general.....                                       | 13 |
| 1.4.2 Objetivos específicos .....                                 | 14 |
| CAPÍTULO 2: MARCO TEÓRICO .....                                   | 15 |
| 2.1 Antecedentes de la investigación.....                         | 15 |
| 2.2 Bases teóricas .....  | 27 |
| 2.2.1 Análisis factorial confirmatorio .....                      | 27 |
| 2.2.2 Índices de Ajuste del Análisis Factorial Confirmatorio..... | 33 |
| 2.2.3 Tamaño de muestra.....                                      | 36 |
| 2.3 Marcos conceptuales o glosario.....                           | 40 |
| 2.3.1 PQH-9. ....   | 40 |
| 2.3.2 SEM .....   | 40 |
| 2.3.3 Causalidad probabilística.....                              | 40 |
| 2.3.4 Distribución normal multivariante.....                      | 41 |
| CAPÍTULO 3: METODOLOGÍA .....                                     | 42 |

|   |  |    |
|---|--|----|
| 3.1                                     | Tipo y diseño de investigación.....              | 42 |
| 3.2                                     | Unidad de análisis.....                          | 42 |
| 3.3                                     | Población de estudio.....                        | 42 |
| 3.4                                     | Tamaño de muestra.....                           | 42 |
| 3.5                                     | Selección de muestra .....                       | 42 |
| 3.6                                     | Técnicas de recolección de Datos.....            | 43 |
| 3.7                                     | Análisis e interpretación de la información..... | 43 |
| CAPÍTULO 4: RESULTADOS Y DISCUSIÓN..... |  | 44 |
| 4.1                                     | Resultados.....                                  | 44 |
| 4.2                                     | Prueba de hipótesis.....                         | 66 |
| 4.3                                     | Discusión .....                                  | 66 |
| CONCLUSIONES .....                      |  | 67 |
| RECOMENDACIONES.....                    |  | 68 |
| REFERENCIAS BIBLIOGRÁFICAS.....         |  | 69 |
| ANEXOS.....                             |  | 73 |

## **LISTA DE FIGURAS**

Figura 1. Parámetros del análisis factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018

Figura 2. Chi-Cuadrado del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Figura 3. Chi-Cuadrado/gl del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018

Figura 4. CFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Figura 5. TLI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Figura 6. NNFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Figura 7. RFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Figura 8. NFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Figura 9. PNFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Figura 10. IFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Figura 11. RNI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Figura 12. AIC del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Figura 13. BIC del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Figura 14. RMSEA del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Figura 15. RMR del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Figura 16. SRMR del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Figura 17. GFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Figura 18. AGFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Figura 19. PGFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Figura 20. MFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Figura 21. ECVI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

## LISTA DE TABLAS

Tabla 1. Medidas de bondad de ajuste del modelo factorial confirmatorio paramétrico sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018

Tabla 2. Chi-Cuadrado del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Tabla 3. Chi-Cuadrado/gl del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018

Tabla 4. CFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Tabla 5. TLI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Tabla 6. NNFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Tabla 7. RFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Tabla 8. NFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Tabla 9. PNFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Tabla 10. IFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Tabla 11. RNI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Tabla 12. AIC del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Tabla 13. BIC del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Tabla 14. RMSEA del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Tabla 15. RMR del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Tabla 16. SRMR del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Tabla 17. GFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Tabla 18. AGFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Tabla 19. PGFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Tabla 20. MFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

Tabla 21. ECVI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.

## RESUMEN

### EFFECTO DEL TAMAÑO DE MUESTRA EN LOS ÍNDICES DE AJUSTE DEL ANÁLISIS FACTORIAL CONFIRMATORIO SOBRE EL CUESTIONARIO PHQ-9 APLICADO A LA POBLACIÓN PERUANA

Moises Enrique Tarazona Cochachin

Mayo – 2021

Asesora: Zoraida Judith Huamán Gutiérrez

**Objetivo:** Determinar el efecto del tamaño de muestra en los índices de ajuste del análisis factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana.

**Método:** Investigación aplicada, de diseño experimental, población conformada por 34476 registros de la encuesta ENDES-INEI 2018, y el tamaño de muestra estuvo conformado por muestra elegidas mediante el muestreo Bootstrap con 100 réplicas en cantidad de 50, 100, 250, 500, 750, 1000, 2500, 5000 y 10000 casos.

**Resultados:** El tamaño de muestra presenta un efecto directo en los índices de ajuste del análisis factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, tal como se muestran en las figuras y tablas 2 al 21, donde se evidencia que el tamaño de muestra presenta un efecto directo con concavidad hacia arriba en el  $X^2$ ,  $X^2/df$ , AIC y BIC, efecto directo con concavidad hacia abajo en el CFI, TLI, NNFI, RFI, NFI, PNFI, IFI, RNI, GFI, AGFI, PGFI Y MFI, y en algunas medidas de bondad de ajuste un efecto indirecto en el RMSEA, RMR, SRMR y ECVI. Por lo que se contrasta un efecto directo e inverso del tamaño de muestra sobre las medidas de bondad de ajuste.

**Conclusiones:** El tamaño de muestra presenta efecto directo o indirecto según corresponda en los índices de ajuste del análisis factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana.

**PALABRAS CLAVE:** Medidas de bondad de ajuste, Análisis factorial confirmatorio, cuestionario PHQ-9.

## ABSTRAC

# EFFECT OF SAMPLE SIZE ON THE ADJUSTMENT INDICES OF THE CONFIRMATORY FACTORIAL ANALYSIS ON THE PHQ-9 QUESTIONNAIRE APPLIED TO THE PERUVIAN POPULATION

Moises Enrique Tarazona Cochachin

May - 2021

Advisor: Zoraida Judith Huamán Gutiérrez

**Objective:** To determine the effect of the sample size on the adjustment indices of the confirmatory factor analysis on the PHQ-9 questionnaire applied to the Peruvian population.

**Method:** Applied research, experimental design, population made up of 34,476 records from the ENDES-INEI 2018 survey, and the sample size was made up of a sample chosen through Bootstrap sampling with 100 replications in the amount of 50, 100, 250, 500, 750, 1000, 2500, 5000 and 10000 cases.

**Results:** The sample size has a direct effect on the adjustment indices of the confirmatory factor analysis on the PHQ-9 questionnaire applied to the Peruvian population, as shown in figures and tables 2 to 21, where it is evidenced that the size sample shows a direct effect with concavity upwards in the  $\chi^2$ ,  $\chi^2 / gl$ , AIC and BIC, direct effect with concavity downwards in the CFI, TLI, NNFI, RFI, NFI, PNFI, IFI, RNI, GFI, AGFI, PGFI and MFI, and in some goodness-of-fit measures an indirect effect on the RMSEA, RMR, SRMR and ECVI. Therefore, a direct and inverse effect of the sample size is contrasted on the goodness of fit measures.

**Conclusions:** The sample size has a direct or indirect effect, as appropriate, on the adjustment indices of the confirmatory factor analysis on the PHQ-9 questionnaire applied to the Peruvian population.

**KEY WORDS:** Goodness of fit measures, Confirmatory factor analysis, PHQ-9 questionnaire.



# **CAPÍTULO 1: INTRODUCCIÓN**

## **1.1 Situación Problemática**

En la actualidad, muchos métodos de modelamiento son empleados para estudiar los fenómenos que requieren la utilización de un conjunto complejo de variables. El análisis factorial confirmatorio (AFC), como caso particular del modelo de ecuaciones estructurales, estudia las construcciones teóricas de un constructo y sus indicadores, para analizarlos y conocer su viabilidad. Debido a que no existe un criterio único para la evaluación de ajuste del modelo teórico obtenido como resultado del AFC, se desarrolló una amplia gama de índices de ajuste, las cuales son afectados directamente por el tamaño de muestra y método de estimación.

Muchos estudios fueron realizados controlando el método de estimación y el tamaño de muestra, en la cual incluyen el cumplimiento de la normalidad multivariante y la complejidad del modelo estimado, para evaluar los índices de ajuste del AFC. Sin embargo, estas evaluaciones de los índices de ajuste no fueron estudiados ampliamente en aplicación a casos reales, particularmente al cuestionario PHQ-9 (Patient Health Questionnaire-9). De ello, radica el interés de fijar tamaños de muestra que muestren la invariabilidad de los índices de ajuste del AFC sobre el cuestionario PHQ-9, aplicado a la población peruana.

Los índices de ajuste proporcionan información sobre la idoneidad del modelo para explicar los datos (Keith, 2019), sin embargo éstos índices de ajuste se ven influenciados principalmente por el tamaño de muestra, así como, por la complejidad del modelo propuesto y el método de estimación.

Por otro lado, el tamaño de muestra se refiere al número de participantes que se incluye en un estudio de investigación. Si se realiza adecuadamente el cálculo, no solamente se establecerán conclusiones con sustento estadístico y se limitarán los riesgos potenciales a los sujetos incluidos en la investigación, sino que además el estudio podrá planearse adecuadamente para optimizar costos económicos y tiempo (Rendón-Macías & Villasís-Keever, 2017). Así el tamaño de muestra proporciona garantías de que el AFC refleje o se ajuste al comportamiento teórico planteado, por medio de los índices de ajuste.

Por ello, es importante evaluar los índices de ajuste del AFC aplicado a datos reales y sin aceptación de supuestos a priori, ni métodos de estimación seleccionados de antemano.

## **1.2 Formulación Problemática**

En toda investigación donde se lleve a cabo un modelamiento, es necesario saber la bondad de ajuste de esta, avalado a un tamaño de muestra óptimo que garantice dicha estimación.

### ***1.2.1 Problema general***

¿Cuál es el efecto del tamaño de muestra en los índices de ajuste del análisis factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana?

### ***1.2.2 Problemas específicos***

- ¿Cuál es el efecto del tamaño de muestra en los índices de ajuste absoluto del análisis factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana?
- ¿Cuál es el efecto del tamaño de muestra en los índices de corrección de parsimonia del análisis factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana?

- ¿Cuál es el efecto del tamaño de muestra en los índices de ajuste comparativo del análisis factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana?
- ¿Cuál es el efecto del tamaño de muestra en los índices de ajuste predictivo del análisis factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana?

### **1.3 Justificación**

#### ***1.3.1 Justificación teórica***

No existe acuerdo universal respecto al tamaño de muestra necesario para obtener índices de ajuste que avalen el AFC, por ello la necesidad de acercarse a un tamaño de muestra adecuado que garantice que los índices de ajuste presenten puntos de inflexión y no sean sensibles al tamaño de muestra, por ello es necesario conocer el tamaño de muestra que lleva a dicho punto de inflexión, y sea tomado por la comunidad científica como un referente a la hora de verificar los índices de ajuste en un AFC.

#### ***1.3.2 Justificación práctica***

Es de suma importancia para la comunidad científica, pues se cuenta con un referente de tamaño de muestra adecuado y sobre todo aplicado a un caso real tomado de la población peruana, a ello, se añaden procedimientos de submuestreo que facilitarán el análisis de los puntos de inflexión de los índices de ajuste del AFC.

### **1.4 Objetivos**

#### ***1.4.1 Objetivo general***

Determinar el efecto del tamaño de muestra en los índices de ajuste del análisis factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana

### ***1.4.2 Objetivos específicos***

- Determinar el efecto del tamaño de muestra en los índices de ajuste absoluto del análisis factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana.
- Determinar el efecto del tamaño de muestra en los índices de corrección de parsimonia del análisis factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana.
- Determinar el efecto del tamaño de muestra en los índices de ajuste comparativo del análisis factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana.
- Determinar el efecto del tamaño de muestra en los índices de ajuste predictivo del análisis factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana.

## **CAPÍTULO 2: MARCO TEÓRICO**

### **2.1 Antecedentes de la investigación**

Vargas Halabí & Mora Esquivel (2017), en su trabajo de investigación titulado: “Tamaño de la muestra en modelos de ecuaciones estructurales con constructos latentes: Un método práctico”, tuvo como objetivo ofrecer una guía práctica a investigadores educativos para utilizar la diversidad de criterios y justificar el cálculo apropiado del tamaño de muestra en ecuaciones estructurales, mediante la planificación a priori de modelos de medida y modelos estructurales. Se recurre a un método cuantitativo para determinar los insumos con base en diagramas de sendero de modelos a priori de una investigación, así como interfaces disponibles en internet para calcular los tamaños de muestra. La aplicación del método constata su utilidad para que el investigador defina un umbral de casos que permita satisfacer distintos criterios, con respecto a depender de una única regla o criterio para justificar el tamaño de la muestra apropiado. Asimismo, revela la ventaja que podría representar el uso de parcelas en la planificación a priori de modelos con esta técnica. (Mora-Esquivel & Halabí, 2017)

Cangur & Ercan (2015), en su trabajo de investigación titulado: “Comparison of Model Fit Indices Used in Structural Equation Modeling Under Multivariate Normality”, tuvo como objetivo investigar el impacto de las técnicas de estimación y los tamaños de muestra en los índices de ajuste del modelo en modelos de ecuaciones estructurales contruidos de acuerdo con el número de variables latentes exógenas bajo normalidad multivariada. El estudio se diseñó como 4 x 3 para examinar los efectos de 4 técnicas de estimación diferentes (Maxima verosimilitud, Mínimos cuadrados generalizados, Distribución libre asintótica y Chi-cuadrado escalado de Satorra-Bentler) y 3 tamaños de

muestra diferentes (100, 500 y 1000) en condiciones de distribución normal multivariadas tomando ambos modelos estructuras en consideración. Concluyen que, la estadística de prueba de Chi-cuadrado de relación al grado de libertad, el error cuadrático medio de aproximación de la raíz y el índice de ajuste comparativo son los índices menos afectados por la técnica de estimación y el tamaño de la muestra bajo normalidad multivariada, especialmente con un tamaño de muestra grande.(Cangur & Ercan, 2015)

Ramírez, Barrera, & Correa (2013), en su trabajo de investigación titulado: “Efecto del tamaño de muestra y el número de réplicas bootstrap”, tuvo como objetivo evaluar el efecto del tamaño de muestra y el número de réplicas bootstrap en las estimaciones de la media, varianza y algunos percentiles para una distribución de probabilidad bajo diferentes medidas de asimetría. Para analizar el efecto del tamaño de muestra y el número de muestras bootstrap en las distribuciones de algunos estimadores, se simula por el método de Monte Carlo una muestra aleatoria de tamaño  $n$  de una distribución gamma  $(\alpha, \beta)$ , de esta muestra se toman  $m$  muestras de tamaño  $n$  con reemplazo y a cada una de estas  $m$  muestras se les calcula el estadístico de interés, por lo tanto se van a tener  $m$  valores de dicho estadístico con los cuales se estima su distribución muestral. Se realizan 1000 simulaciones de este procedimiento obteniendo 1000 distribuciones bootstrap del estadístico de interés. Luego se compara la verdadera distribución muestral del estadístico con cada una de las 1000 distribuciones bootstrap. La comparación se hace mediante el uso de la distancia máxima de Kolmogorov-Smirnov. Finalmente se hace un análisis del promedio de las distancia máximas para cada uno de los casos. Los tamaños de muestras  $n$ , que se evalúan son 5, 10, 20, 40, 80, 150 y 300. Para cada uno de estos tamaños de muestras se extraen  $m$  réplicas bootstrap de tamaños 50, 100, 250, 500 y 1000. Concluyen que se encontró que el tamaño de muestra tiene un efecto en las estimaciones de las distribuciones de los estadísticos de interés. En todos los casos la DMP (distancia máxima promedio) disminuye a medida que se aumenta el tamaño de muestra hasta un valor dado, pero en el caso del percentil 95, este comportamiento se puede apreciar. Por otro lado, se notó que la DMP disminuye dependiendo del número de muestras bootstrap utilizado, pero los valores recomendados en nuestros casos estudiados son de 100 o 200 muestras bootstrap.(Ramírez et al., 1969)

Thoemmes, Rosseel, & Textor (2018), en su trabajo de investigación titulado: “Local Fit Evaluation of Structural Equation Models Using Graphical Criteria”, tuvo como objetivo presentar un enfoque alternativo en el que la prueba global se divide en pruebas locales utilizando criterios gráficos simples. En contraste, los criterios gráficos como la d-separación o la trek-separación permiten derivar implicaciones que pueden usarse para la evaluación de ajuste local, un enfoque que casi nunca se aplica. Concluyen que, las pruebas locales son una valiosa adición a la caja de herramientas del investigador aplicado. Así, la evaluación de ajuste local como un suplemento que puede fomentar una explicación más exhaustiva del ajuste del modelo, ya que la evaluación de ajuste local puede proporcionar información de diagnóstico útil, especialmente cuando los modelos no convergen o no se ajustan.(Thoemmes et al., 2018)

Shevlin & Miles (1998), en su trabajo de investigación titulado: “Effects of sample size, model specification and factor loadings on the GFI in confirmatory factor analysis”, tuvo por objetivo examinar el rendimiento del GFI en condiciones de diferentes tamaños de muestra, especificación del modelo y magnitud de las cargas factoriales. Para ello, el estudio empleó un diseño de tres factores: 3 (Modelo) x 5 (Muestra) x 3 (Cargas). Los tres niveles del factor Modelo se refieren a tres tipos de modelo factorial. Cada uno de los tres modelos consta de seis variables observadas. El primer modelo es un modelo de factor único. Este es el modelo "correcto". El segundo es un modelo de factor único con tres errores correlacionados de 0.1 entre las variaciones únicas correspondientes a los ítems 1 y 2, 3 y 4, y 5 y 6. Este es el modelo "aproximado". El tercer modelo fue un modelo de dos factores correlacionado. Los elementos 1 a 3 se cargaron en el primer factor y los elementos 4 a 6 se cargaron en el segundo factor. Se especificó una correlación de 0.7 entre los dos factores. Este es el modelo mal especificado. Los cinco niveles del Factor de muestra se refieren a tamaños de muestra de 50, 100, 200,400 y 800. Los tres niveles del Factor de carga se refieren a las cargas de factor; bajo (0.3), medio (0.5) y alto (0.7), para cada condición, se calculó la matriz de covarianza de la población y se generaron 200 matrices de covarianza de muestra utilizando los procedimientos de Monte Carlo. Concluyen que, los resultados de este estudio pueden usarse para un proceso de toma de decisiones más informado en términos de la GFI. En primer lugar, está claro que un punto de corte ómnibus de 0.9 no es apropiado en todas las condiciones, en particular, cuando

las cargas de factor y los tamaños de muestra son bajos. Al igual que otros índices de ajuste, el rendimiento del GFI mejora al aumentar el tamaño de la muestra, sin embargo, lo que se ha demostrado es que el rendimiento del GFI se basa en el tamaño de la muestra y la magnitud de los parámetros. Los investigadores deben preocuparse cuando se enfrentan a un GFI inesperadamente alto, cuando las estimaciones de los parámetros son bajas.(Miles & Shevlin, 1998)

MacCallum, Browne, & Sugawara (1996), en su trabajo de investigación titulado: “Power Analysis and Determination of Sample Size for Covariance Structure Modeling”, tuvo como objetivo presentar un marco para la prueba de hipótesis y el análisis de potencia en la evaluación del ajuste de modelos de estructura de covarianza. Para ello, se utiliza el índice RMSEA ( $\epsilon$ ). Las características críticas del enfoque implican la capacidad de especificar valores nulos y alternativos sensibles de  $\epsilon$  y definir los valores de los parámetros de no centralidad para las distribuciones  $\chi^2$  relevantes como una función de  $\epsilon$ . Esas distribuciones luego forman la base para potencia y cálculos de tamaño de muestra. Concluyen que los resultados indican que la evaluación del modelo en tales casos (grados de libertad bajos) es muy problemática y probablemente no debería llevarse a cabo a menos que estén disponibles muestras muy grandes. En conjunto con esta visión, desaconsejamos la introducción de un número sustancial de parámetros en los modelos para mejorar su ajuste. Se ha demostrado que dichos procedimientos son susceptibles del azar.(MacCallum et al., 1996)

Iacobucci (2010), en su trabajo de investigación titulado: “Structural equations modeling: Fit Indices, sample size, and advanced topics”, tuvo como objetivo discutir los dos problemas técnicos más importantes y generalizados: la elección desconcertante entre las estadísticas de ajuste y la cuestión del tamaño de muestra requerido. Para ello, primero, hay cierto acuerdo en que los investigadores deben informar el siguiente perfil de índices: el  $\chi^2$  (y sus grados de libertad y valor p), el residuo cuadrático medio estandarizado (SRMR) y el índice de ajuste comparativo (CFI), segundo, se examina la cuestión del tamaño de la muestra necesario para tener un buen modelo SEM. Para ello, se utilizó la simulación Monte Carlo con variación del tamaño de muestra (tres categorías), y tres tipos de cargas factoriales. Luego, para un tamaño de muestra dado, se generaron seis



desviaciones aleatorias normales y se transformaron mediante la matriz de correlación de la población. Los datos resultantes se analizaron mediante SEM y se anotaron las estadísticas de ajuste. En cada celda, se crearon 2000 réplicas de este tipo. El modelo lineal general, específicamente ANOVA, se utilizó para analizar el efecto del tamaño de la muestra en los índices de ajuste. Los resultados encontrados indican que a medida que aumenta el tamaño de la muestra, aumenta el  $\chi^2$ , y su valor  $p$  correspondiente disminuye. El SRMR disminuye y el CFI mejora. El efecto del tamaño de la muestra en  $\chi^2$  no es monótono, explotando para N grande (500 o 1000). El efecto para SRMR es casi lineal: cada nuevo punto de datos contribuye a ayudar a SRMR. El efecto sobre el CFI no es lineal y los datos sugieren que una muestra mínima de 50 puede ser beneficiosa, después de lo cual el impulso disminuye. Finalmente, presenta las siguientes conclusiones: (1) Los SEM no dan miedo: son progresiones naturales del análisis factorial y la regresión. (2) Como tal, tenga cuidado de no interpretar en exceso los coeficientes de ruta como si fueran causales, más que si los resultados se hubieran obtenido por regresión. (3) Aspire para un tamaño de muestra de al menos 50. (4) Idealmente, cada construcción se mediría por al menos tres variables indicadoras. Si algunas construcciones son elementos individuales, probablemente eso esté bien. Las construcciones medidas con cuatro o más variables son probablemente excesivas. (5) Utilice la estimación de máxima verosimilitud. (Por lo general, es el valor predeterminado de todos modos). (6) Verifique las estadísticas de ajuste: No tome las reglas generales demasiado en serio. No se preocupe demasiado por  $\chi^2$ , simplemente no se ajustará si el tamaño de la muestra es 50 o más. En cambio, vea si  $\chi^2/df$  es aproximadamente 3 o menos. No sea demasiado crítico si el CFI no es del todo .95, o el SRMR no es del todo .09. (7) Por otro lado, haga buenas preguntas teóricas: ¿se respalda lógicamente cada vínculo hipotético, y hay una historia teórica sólida, completa pero parsimoniosa para todo el modelo? (8) Ajuste al menos un modelo competitivo no trivial, presumiblemente representando la literatura existente sobre la cual se está construyendo el modelo focal, para ver una mejora demostrable. (Iacobucci, 2010)

Fan, Thompson & Wang (1999), en su trabajo de investigación titulado: “Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes”, tuvieron como objetivo investigar los efectos en los índices de ajuste del modelo de ecuaciones estructurales del tamaño de muestra, método de

estimación y especificación del modelo. Para ello, realizaron simulaciones Monte Carlo basado sobre un diseño experimental balanceado de tres factores (5 niveles de tamaño de muestra, 3 niveles de especificación del modelo, y 2 métodos de estimación) con 200 repeticiones para cada condición. Fueron evaluados diez índices de ajuste. Los resultados muestran que la condición del tamaño de muestra influyó fuertemente en GFI y AGFI, contabilizando un 10.27% y 11.37% de la varianza total respectivamente para estos dos índices. RMSEA, CENTRA, CFI y NNFI fueron los menos sensibles al tamaño de muestra, con menos del 1% de la variación total contabilizada para esta condición. Porque el valor de Chi-cuadrado está en función directa del tamaño de muestra, no es sorpresa observar grandes porciones de variabilidad contabilizados por el tamaño de muestra. Los otros tres índices de ajuste han tenido alrededor del 4% del total de varianza debido al tamaño de muestra. CENTRA, RMSEA, CFI y NNFI fueron los menos influenciados por el tamaño de muestra, mientras que GFI y AGFI mostraron ser demasiados afectados por el tamaño de muestra. Las principales conclusiones son: (1) Algunos índices de ajuste parecen ser no comparables en términos de información que proporcionan acerca del ajuste del modelo para modelos mal especificados, y (2) el método de estimación influyó casi a todos los índices de ajuste estimados, específicamente para modelos mal especificados. (Fan et al., 1999)

Igundunasse (2016), en su trabajo de investigación titulado: “The impact of small samples sizes in factor analytical studies”, tuvo como objetivo comprender la estructura factorial de la escala de Pettigrew y Meertens sobre prejuicios evidentes y sutiles utilizando un análisis factorial confirmatorio (CFA). Pero resultó que la naturaleza y el carácter de los problemas de tamaño de la muestra en el estudio tuvieron una influencia significativa en el resultado del estudio. En un diseño de sección transversal, utilizando un enfoque de muestreo de disponibilidad, este estudio utilizó una variedad de muestras que van desde tamaños de muestra medianos ( $n = 113$ ) a pequeños ( $n = 19$ ) de siete grupos para explorar el impacto de los tamaños de muestra pequeños en factores de estudios analíticos. Los datos fueron imputados en SPSS 18.0 y analizados en AMOS 18.0. Los resultados revelaron que los tamaños de las muestras tuvieron un impacto variable en los resultados de la investigación que los investigadores a menudo dan por sentado. Los tamaños de muestra son muy importantes para los estudios de análisis factorial. El tamaño

de la muestra también podría hacer que las cargas de factores sean inestables. A menudo es difícil determinar con precisión la causa exacta de estos términos de error en un modelo, es evidente en este caso que los tamaños de muestra contribuyeron en gran medida a este desarrollo. Por lo tanto, podemos sugerir que las violaciones de la normalidad multivariada, la falta de un tamaño de muestra suficiente y el método utilizado para recopilar datos pueden haber contribuido a los términos de error correlacionados observados en la presente investigación. Se sugiere que los investigadores deberían establecer claramente estrategias para tratar tamaños de muestra insuficientes para no oscurecer la base de la interpretación de resultados como la prueba de normalidad y el carácter de las cargas de factores. Por otro lado, los resultados deben interpretarse minuciosamente para evitar conclusiones engañosas debido a problemas de tamaño de la muestra. Este estudio atestigua el hecho de que los problemas de tamaño de la muestra podrían ser una razón importante para la distorsión en los resultados. Por lo tanto, sería pertinente sugerir que otros intentos de investigación deberían explorar el uso de los tamaños de muestra deseados y adoptar un mecanismo apropiado para tratar tales problemas cuando surjan problemas de este tipo. Se recomienda que, en cualquier tipo de estudio de análisis factorial confirmatorio, el investigador se esfuerce por establecer primero el tamaño de muestra deseado y restringir el análisis estadístico a pruebas estadísticas compatibles o buscar un equivalente no paramétrico del análisis deseado. Con estos enfoques, un investigador puede llegar a una interpretación y conclusión equitativas en estudios de análisis factorial confirmatorio que pueden tener problemas de tamaño de la muestra. (Igundunasse, 2016)

Hooper, Coughlan & Mullen (2008), en su trabajo de investigación titulado: "Structural Equation Modelling: Guidelines for Determining Model Fit Structural equation modelling: guidelines for determining model fit", tuvo por objetivo describir una selección de índices de ajuste que son ampliamente considerados como los índices más informativos disponibles para los investigadores. Consideran que tener una colección de índices de ajuste incita al investigador a seleccionar aquellos que indican un buen ajuste del modelo. Esta práctica debe desistirse a toda costa, ya que oculta los problemas subyacentes que sugieren posibles especificaciones erróneas dentro del modelo. Así, los índices de ajuste absoluto determinan qué tan bien un modelo a priori se ajusta a los datos

de la muestra y demuestra qué modelo propuesto tiene el ajuste más superior. Estas medidas proporcionan la indicación más fundamental de cuán bien la teoría propuesta se ajusta a los datos. A diferencia de los índices de ajuste incremental, su cálculo no se basa en la comparación con un modelo de referencia, sino que es una medida de qué tan bien se ajusta el modelo en comparación con ningún modelo. En esta categoría se incluyen la prueba Chi-Squared, RMSEA, GFI, AGFI, RMR y SRMR. Los índices de ajuste incremental, también conocidos como comparativos o índices de ajuste relativo, son un grupo de índices que no utilizan el chi-cuadrado en su forma cruda, sino que comparan el chi-cuadrado valor para un modelo de referencia. Para estos modelos, la hipótesis nula es que todas las variables no están correlacionadas, entre estos índices se tiene al NFI y CFI. Los índices de ajuste incremental, consideran que tener un modelo complejo casi saturado significa que el proceso de estimación depende de los datos de la muestra. Esto da como resultado un modelo teórico menos riguroso que paradójicamente produce índices de mejor ajuste. Para superar este problema, se han desarrollado dos parsimonias de índices de ajuste; el Índice de bondad de ajuste de la parsimonia (PGFI) y el Índice de ajuste normalizado de la parsimonia (PNFI). El PGFI se basa en el GFI ajustando por la pérdida de grados de libertad. El PNFI también se ajusta a los grados de libertad, sin embargo, se basa en el NFI. Ambos índices penalizan seriamente la complejidad del modelo que resulta en valores de índice de ajuste de parsimonia que son considerablemente más bajos que otros índices de bondad de ajuste. Si bien no se han recomendado niveles umbral para estos índices, señalan que es posible obtener índices de ajuste de parsimonia dentro de la región .50, mientras que otros índices de bondad de ajuste alcanzan valores superiores a .90. Sin embargo, se recomienda encarecidamente el uso de índices de ajuste de parsimonia junto con otras medidas de bondad de ajuste, ya que no se han recomendado niveles umbral para estas estadísticas, esto los ha hecho más difíciles de interpretar. Una segunda forma de índices de parsimonia es aquellos que también se conocen como índices de "criterios de información". Probablemente el más conocido de estos índices es el Criterio de Información de Akaike (AIC) o la Versión Consistente de AIC (CAIC) que se ajusta al tamaño de la muestra. Estas estadísticas se usan generalmente cuando se comparan modelos no anidados o no jerárquicos estimados con los mismos datos e indican al investigador cuál de los modelos es el más parsimonioso. Los valores más pequeños

sugieren un modelo adecuado y parsimonioso, sin embargo, debido a que estos índices no están normalizados a una escala de 0-1, es difícil sugerir un punto de corte que no sea que el modelo que produce el valor más bajo es el más superior. También vale la pena señalar que estas estadísticas necesitan un tamaño de muestra de 200 para que su uso sea confiable. Con respecto a los índices que deben informarse, no es necesario ni realista incluir todos los índices incluidos en el resultado del programa, ya que será una carga tanto para el lector como para el revisor. Dada la gran cantidad de índices de ajuste, se convierte en una tentación elegir aquellos índices de ajuste que indiquen el mejor ajuste. Esto debe evitarse a toda costa, ya que esencialmente está barriendo información importante debajo de la alfombra. En una revisión se encontró que los índices de ajuste más comúnmente reportados son el CFI, GFI, NFI y el NNFI. Al decidir qué índices reportar, seguir lo que se usa con más frecuencia no es necesariamente una buena práctica, ya que algunas de estas estadísticas a menudo se basan únicamente en razones históricas, en lugar de su sofisticación. Si bien no existen reglas de oro para la evaluación del ajuste del modelo, es necesario informar una variedad de índices porque los diferentes índices reflejan un aspecto diferente del ajuste del modelo. Aunque el modelo Chi-Square tiene muchos problemas asociados, aún es esencial que esta estadística, junto con sus grados de libertad y el valor p asociado, se informe en todo momento. El nivel de umbral sugeridos consta de un formato de presentación de dos índices. Esto siempre incluye el SRMR con el NNFI (TLI), RMSEA o el CFI. Se habla enérgicamente sobre qué índices incluir y aboga por el uso de la prueba de Chi-Square, el RMSEA, el CFI y el SRMR. Se aconseja que se informen las correlaciones múltiples al cuadrado de cada ecuación. Es razonable incluir la estadística Chi-Square, sus grados de libertad y valor p, el RMSEA y su intervalo de confianza asociado, el SRMR, el CFI y un índice de ajuste de parsimonia como el PNFI. Estos índices se han elegido sobre otros índices, ya que se ha encontrado que son los más insensibles al tamaño de la muestra, la especificación errónea del modelo y las estimaciones de los parámetros. (Hooper et al., 2008)

Rožman, Tominc & Milfelner (2020), en su trabajo de investigación titulado: “A Comparative Study Using Two SEM Techniques on Different Samples Sizes for Determining Factors of Older Employee’s Motivation and Satisfaction”, tuvieron como objetivo es presentar las relaciones entre el estrés, la motivación, la satisfacción y las

relaciones de los empleados mayores en el lugar de trabajo utilizando dos enfoques diferentes y tamaños de muestra diferentes. La investigación se implementó en una muestra inicial de 1013 empleados mayores. En el siguiente paso, se calcularon seis muestras más pequeñas utilizando la selección aleatoria de casos, muestras con 25, 50, 100, 250, 400 y 500 empleados mayores. Los resultados en las muestras más grandes han demostrado ser bastante sólidos ya que se confirmaron con ambos enfoques. Con respecto a los índices de ajuste del modelo, ambos métodos demostraron ser sensibles a la muestra. El índice de ajuste normalizado (NFI) no estaba en los intervalos sugeridos para muestras por debajo de 100, independientemente del método SEM utilizado. Para el CB-SEM (Basado en la covarianza), el GFI estaba por debajo del umbral sugerido de 0.9, excepto para la muestra más grande, pero se sabe que este índice es altamente sensible a tamaños de muestra pequeños. Además, el RMSEA demostró estar afectado por el mismo problema para muestras menores de 400, ya que era más alto que el umbral recomendado de 0.08 o 0.07, que se reconoce como indicativo de un buen ajuste, lo que significa que estrictamente por la regla, en algunos casos, el modelo debe ser rechazado. (Rožman et al., 2020)

Koyuncu & Kılıç (2019), es su trabajo de investigación titulado: “The use of exploratory and confirmatory factor analyses: A document analysis”, tuvo por objetivo revisar la investigación de desarrollo de escala publicada en Turquía entre 2006 y 2016 con respecto a sus procesos de análisis exploratorio (EPT) y factor confirmatorio (CFA). Dentro de este alcance, se investigó la distribución de los estudios de acuerdo con los años y los análisis factoriales, el grado en que sus hipótesis surgieron para ambos métodos de análisis y la distribución de los resultados de EFA y CFA según los documentos publicados. De esta manera, nuestro objetivo fue inferir resultados significativos en la aplicación práctica de EFA y CFA que se refieren con frecuencia en teoría. Por lo tanto, la presente investigación incorpora un análisis de documento descriptivo sobre 131 estudios de desarrollo a escala publicados en los campos de la educación y las ciencias sociales en las revistas indexadas en la base de datos TR Dizin del Consejo Científico y Tecnológico de Turquía (TÜBİTAK) entre 2006 y 2016. Solo se usó EFA en 59 artículos, solo CFA en 4 artículos, y ambos Se emplearon métodos de análisis en 58 artículos. Los 10 documentos restantes no utilizaron ninguno de los análisis e incluyeron solo análisis de ítems. El tamaño promedio de la muestra fue de 395 participantes para EFA y 529

participantes para CFA. Los tamaños de muestra fueron, en consecuencia, adecuados en términos de análisis factoriales. EFA y CFA se realizaron sobre la misma muestra en 36 artículos, mientras que los análisis se basaron en diferentes muestras en 22. En términos de estimación, el 65% de 117 EFA utilizando documentos emplearon el Análisis de componentes principales (PCA), el 2% utilizó la Estimación de máxima verosimilitud (MLE), el 1% utilizó Análisis paralelo (PA), y el 3% utilizó Factorización del eje principal (PAF), mientras que el 29% no especificó su método de estimación. Las hipótesis de los análisis de EPT se examinaron en el 93% de los artículos. Se tomaron en cuenta las parcelas al determinar el número de factores en el 50% de los documentos. Se usó el método de rotación Varimax en el 67% de los EFA utilizando papeles. La varianza total explicada fue del 53,05%. Las hipótesis analíticas no se examinaron en el 87% de los 62 estudios realizados con CFA; 51 de ellos incluyeron diagramas de seguimiento y 52 incluyeron cargas factoriales. El 90% de los documentos no dieron ninguna información sobre el método de estimación de parámetros. Todos los índices, excepto el Índice de bondad de ajuste de la parsimonia (PGFI), fueron aceptables como índices de ajuste modelo en los documentos. (Koyuncu & Kılıç, 2019)

Tanaka (1987), en su trabajo de investigación titulado: “How Big Is Big Enough?: Sample Size and Goodness of Fit in Structural Equation Models with Latent Variables”, tuvo como objetivo proponer alternativas para el tamaño de muestra y medidas de bondad de ajuste de modelos de ecuaciones estructurales con variables latentes. Para ello, realizó una aproximación Monte Carlo, con diversos tamaños de muestra y métodos de estimación. Concluye que, la idoneidad del tamaño de la muestra está íntimamente relacionada con el tamaño del modelo a estimar. Cincuenta observaciones pueden ser suficientes para un modelo que hipotetiza una sola variable latente subyacente a cuatro indicadores medidos. El mismo número de observaciones será inadecuado para un modelo con 20 variables medidas y cuatro variables latentes. En particular, la adecuación del tamaño de la muestra está vinculada a la cantidad de parámetros estimados en el modelo. La información sobre el número de parámetros estimados en un modelo es la salida estándar en los programas de modelado de variables latentes. La complejidad del método de estimación determina la idoneidad del tamaño de la muestra. Los desarrollos recientes en modelos de variables latentes que hacen menos suposiciones sobre la distribución de

los datos y permiten la no normalidad de los datos requerirán más sujetos que métodos más estándar, como ML y GLS. Por lo tanto, el costo de hacer menos suposiciones de distribución sobre los datos es la necesidad de un gran tamaño de muestra. Un criterio importante para seleccionar un índice de ajuste es su generalización entre los métodos de estimación. (Tanaka, 1987)

Gomer, Jiang & Yuan (2019), en su trabajo de investigación titulado: “New Effect Size Measures for Structural Equation Modeling”, tuvo como objetivo desarrollar dos nuevas clases de medidas de tamaño del efecto para SEM generalizando el “*d*” de Cohen. La primera clase consiste en definiciones que son teóricamente equivalentes a  $F_{ML}(\Sigma_0, \Sigma(\theta))$ , la contraparte de la población de la función de discrepancia basada en la distribución normal. La segunda clase de medidas de tamaño del efecto tiene una semejanza más estricta con la “*d*” de Cohen en su forma original. Se investigan varias versiones de estas generalizaciones para identificar la que está menos afectada por el tamaño de la muestra y la distribución de la población, pero más sensible a la especificación errónea del modelo. Su desempeño bajo supuestos de distribución violados, la gravedad de la especificación errónea del modelo y varios tamaños de muestra normales se examinan utilizando tanto la estimación de máxima verosimilitud como la estimación M-robusta. Los resultados de Monte Carlo indican que una medida en la primera clase de medidas de tamaño del efecto se ve poco afectada por el tamaño y la distribución de la muestra al tiempo que conserva la sensibilidad a la especificación errónea del modelo y, por lo tanto, se recomienda que los investigadores informen en publicaciones. Si bien el propósito de “*ε*” no es reemplazar las estadísticas de prueba o los índices de ajuste, sí complementa el análisis estadístico en las aplicaciones. Proporciona información adicional que no se puede obtener de las medidas existentes en SEM. La información complementaria a las pruebas de hipótesis que capturan el tamaño de la especificación errónea sin estar sujetos a la necesidad de controlar los errores de Tipo I es extremadamente valiosa. Un modelo con una pequeña especificación errónea puede ser rechazado en una prueba de hipótesis, pero puede ser recompensado por el tamaño de efecto suplementario “*ε*”. (Gomer et al., 2019)



Huba & Harlow (1987), en su trabajo de investigación titulado: “Robust Structural Equation Models: Implications for Developmental Psychology”, tuvo como objetivo comparar resultados de varias técnicas alternativas en modelos de ecuaciones estructurales. Concluyen que la "bondad de ajuste" generalizada o el índice GFI parece más apropiado por dos razones. Primero, se puede demostrar que una forma general del índice GFI tiene un valor óptimo cuando las funciones de ajuste de modelado causal alcanzan su mínimo. Es decir, las técnicas de modelado causal que minimizan los coeficientes de Chi-cuadrado o de tipo Chi-cuadrado maximizarán los coeficientes GFI. En segundo lugar, a partir de este resultado general, se puede demostrar que se pueden derivar coeficientes específicos para máxima verosimilitud, mínimos cuadrados generalizados, distribución asintóticamente libre y mínimos cuadrados no ponderados. Por lo tanto, el coeficiente general es apropiado para estos estimadores diferentes, y aproximadamente la misma métrica para el coeficiente se aplica independientemente del método de estimación de parámetros. Es decir, los coeficientes GFI obtenidos de diferentes métodos de estimación de parámetros se pueden comparar entre sí. (Huba & Harlow, 1987)

## **2.2 Bases teóricas**

### ***2.2.1 Análisis factorial confirmatorio***

***Introducción.*** El análisis factorial es un nombre genérico que se da a una clase de métodos estadísticos multivariantes cuyo propósito principal es definir la estructura subyacente en una matriz de datos. Generalmente hablando Abordar el problema de cómo analizar la estructura de las interrelaciones (correlaciones) entre un gran número de variables Con la definición de una serie de dimensiones subyacentes comunes, conocidas como factores (Hair et al., 1999).

El modelo de ecuaciones estructurales (SEM) abarca una familia entera de modelos conocidos con muchos nombres entre ellos análisis de la estructura de la covarianza, análisis de variable latente y análisis de factor confirmatorio. Producto de una evolución de la modelización multiecuacional desarrollada principalmente en la

econometría y fusionada con los principios de medición de la psicología y la sociología el SEM se ha convertido en una herramienta integral tanto en la investigación académica como en la práctica. Como puede esperarse de una técnica de uso tan extendido y con tantas variantes y aplicaciones muchos investigadores no están seguros de lo que constituye el modelo de ecuaciones estructurales. Sin embargo, todas las técnicas SEM se distinguen por dos características: estimación de relación de dependencia múltiples y cruzadas, y la capacidad de representar conceptos no observados en estas relaciones y tener en cuenta el error de medida en el proceso de estimación (Hair et al., 1999). SEM, puede jugar un papel confirmatorio porque el investigador tiene un control completo sobre la especificación de los indicadores de cada constructo. Además, SEM permite un test estadístico de calidad del ajuste para la solución confirmatoria del factor propuesto. El análisis de factor confirmatorio (CFA) es particularmente útil en la validación de las escalas para la medida de los constructos específicos.

En muchas ocasiones el investigador tiene ideas bastante claras sobre cuáles pueden ser las variables latentes y sobre qué relaciones puede haber entre ellas y las variables observadas. En estos casos, dispone de dos técnicas muy relacionadas entre sí, el factorial confirmatorio (CFA) y los modelos de ecuaciones estructurales (SEM). El CFA es un caso particular de SEM. Estas dos técnicas tienen un objetivo fundamental, contrastar un modelo teórico sobre las relaciones entre variables independientes y variables dependientes. Tanto unas como otras pueden ser variables latentes o variables observadas, medidas. Para lograr este objetivo el investigador dispone, como en el análisis factorial exploratorio, de un conjunto de datos observados que se resumen en una matriz de, varianzas-covarianzas. Así, el objetivo se reduce a estimar un conjunto de parámetros que indiquen la relación (causal o correlacional) entre las variables, de modo que pueda estimarse una matriz de varianzas-covarianzas poblacional que sea lo más parecida posible a la matriz observada, muestral. La matriz estimada dependerá lógicamente del modelo que el investigador quiera contrastar, esto es, del conjunto de parámetros que suponga necesarios para explicar sus datos (Catena et al., 2003).

El análisis factorial es un método estadístico utilizado para descubrir estructuras latentes que hipotéticamente pueden subyacer a la covarianza o correlación entre las

variables típicamente observadas continuamente. Más técnicamente, podemos decir que la tarea del análisis factorial es aproximar una matriz, ya sea la matriz de covarianza o correlación, por una de rango inferior (Denis, 2016).

Los modelos de análisis factorial confirmatorio (CFA), también llamados modelos de medición, permiten a los investigadores probar la validez de constructo de las variables latentes de interés (Pituch & Stevens, 2019).

En el CFA, no le preocupan los factores de "descubrimiento" o "divulgación" como en el análisis factorial exploratorio, sino la cuantificación, prueba y confirmación de una estructura a priori propuesta (preconcebida) o hipotética de las relaciones entre un conjunto de medidas consideradas. A este respecto, el interés principal en CFA radica en examinar el patrón de relaciones entre los factores, así como aquellos entre ellos y las variables observadas. Por esta razón, para llevar a cabo un CFA, uno debe tener una idea inicial clara sobre la composición del conjunto de variables analizadas. Este método analítico es, por lo tanto, un enfoque de modelado que está diseñado para probar hipótesis sobre la estructura factorial subyacente, cuando se dan de antemano el número de factores involucrados y su interpretación en términos de las medidas manifiestas. Por lo tanto, uno puede ver que el CFA comienza donde termina una teoría en consideración, cuantificando todos los aspectos desconocidos de un modelo derivado de esa teoría y probando si el modelo es consistente con los datos disponibles (Raykov & Marcoulides, 2009).

*Matemáticas del análisis factorial confirmatorio (Mulaik, 2010).* Un modelo hipotético del análisis factorial confirmatorio tiene la misma forma matemática que en el análisis exploratorio de factores comunes. La hipótesis se afirma mediante la ecuación del modelo de análisis factorial habitual.

$$Y = \Lambda X + \Psi E$$

Donde:

"Y" es un vector aleatorio  $n \times 1$  de variables aleatorias observadas

" $\Lambda$ " es una matriz de cargas de patrones  $n \times r$

"X" es un vector aleatorio  $r \times 1$  de factores comunes latentes

" $\Psi$ " es una matriz diagonal  $n \times n$  de cargas de patrones de factor único

"E" es un vector aleatorio  $n \times 1$  de variables latentes de factor único

La matriz de covarianza reproducida para las variables observadas basadas en el modelo viene dada por la ecuación:

$$\Sigma = \Lambda\Phi\Lambda^t + \Theta = \Lambda\Phi\Lambda^t + \Psi^2$$

Donde:

" $\Sigma$ " es la matriz  $n \times n$  de varianzas-covarianzas de las puntuaciones observadas de una población.

" $\Phi$ " es la matriz  $r \times r$  de varianzas-covarianzas de los factores comunes.

" $\Theta$ " es la matriz (diagonal)  $n \times n$  de varianzas-covarianzas de los factores únicos.

***Especificación del Modelo.*** La teoría y/o la investigación previa son cruciales para especificar un modelo de CFA que se probará. La solución de una escala se prueba en base a la conceptualización de la escala como un factor global (es decir, unitario). Al principio del proceso de desarrollo de la medición, los investigadores pueden confiar completamente en la teoría para desarrollar un modelo CFA. Sin embargo, a medida que se usa una medida a lo largo del tiempo, el CFA se puede usar para replicar EFA (Análisis Factorial Exploratorio) u otros análisis que se hayan realizado sobre la medida. El análisis factorial confirmatorio puede no ser un análisis apropiado para usar si no hay una base subyacente sólida sobre la cual basar el modelo, y puede ser necesario un trabajo más preliminar, como la EFA o el desarrollo de la teoría (Harrington, 2009).

***Parámetros de un modelo CFA (Brown, 2015).*** Todos los modelos CFA contienen cargas de factores, variaciones únicas y variaciones de factores. Las cargas de factores son las pendientes de regresión para predecir los indicadores a partir de la variable latente. La varianza única es la varianza en el indicador que no se explica por las variables latentes. Por lo general, se supone que la varianza única es un error de medición y, por lo

tanto, a menudo se la conoce como tal (otros términos sinónimos incluyen varianza del error y falta de confiabilidad del indicador). En una solución no estandarizada, una variación del factor expresa la variabilidad de la muestra o la dispersión del factor, es decir, el grado en que la posición relativa de los participantes de la muestra en la dimensión latente es similar o diferente. Si se justifica sustancialmente, un CFA puede incluir covarianzas de error (denominadas unicidades correlacionadas, residuos o errores correlacionados), lo que sugiere que dos indicadores covarían por razones distintas a la influencia compartida de la variable latente. En las soluciones CFA, las covarianzas de error a menudo se especifican en función de los efectos del método (por ejemplo, los indicadores se miden por un método común), aunque son posibles otras fuentes de estas relaciones. Cuando la solución CFA consta de dos o más factores, generalmente se especifica una covarianza de factor (una correlación de factor que es la contraparte completamente estandarizada) para estimar la relación entre las dimensiones latentes (aunque es posible fijar las covarianzas de factor a cero, similar a una solución ortogonal de EFA). CFA a menudo se limita al análisis de estructuras de varianza-covarianza. En este caso, se estima que los parámetros antes mencionados (cargas de factores, varianzas de error y covarianzas, varianzas de factores y covarianzas) reproducen la matriz de varianza-covarianza de entrada. El análisis de las estructuras de covarianza se basa en el supuesto implícito de que los indicadores se miden como desviaciones de sus medias (es decir, todos los indicadores significan igual a cero). Sin embargo, el modelo CFA se puede ampliar para incluir el análisis de estructuras medias, en cuyo caso los parámetros CFA también se esfuerzan por reproducir las medias de muestra observadas de los indicadores (que se incluyen junto con las varianzas y covarianzas muestrales como datos de entrada). En consecuencia, tales modelos de CFA también incluyen estimaciones de parámetros de las intercepciones del indicador (valor predicho del indicador cuando el factor es cero) y las medias variables latentes, que a menudo se usan en CFA de grupos múltiples para probar si los distintos grupos difieren en su posición relativa en dimensiones latentes.

Las variables latentes en CFA pueden ser exógenas o endógenas. Una variable exógena es una variable que no es causada por otras variables en la solución. Por el contrario, una variable endógena es causada por una o más variables en el modelo (es decir, otras variables en la solución ejercen efectos directos sobre la variable. Por lo tanto,

las variables exógenas se pueden ver como sinónimos de variables X, independientes o predictoras (causales). Del mismo modo, las variables endógenas son equivalentes a las variables Y, dependientes o de criterio (resultado). Sin embargo, en el caso de los modelos estructurales, una variable endógena puede ser la causa de otra variable endógena

***Identificación del modelo CFA.*** Para estimar los parámetros en CFA, se debe identificar el modelo de medición. Se identifica un modelo si, sobre la base de información conocida (es decir, las variaciones y covarianzas en la matriz de entrada de muestra), es posible obtener un conjunto único de estimaciones de parámetros para cada parámetro en el modelo cuyos valores son desconocidos (por ejemplo, cargas de factores, correlaciones de factores). La identificación del modelo pertenece en parte a la diferencia entre el número de parámetros del modelo estimados libremente y el número de piezas de información en la matriz de varianza-covarianza de entrada. Antes de abordar este problema, se discute un aspecto de identificación específico para el análisis de variables latentes: escalar la variable latente (Brown, 2015).

***Identificación estadística.*** Además de escalar la variable latente, los parámetros de un modelo CFA solo se pueden estimar si el número de parámetros estimados libremente no excede el número de piezas de información en la matriz de varianza-covarianza de entrada. Un modelo se subidentifica cuando el número de parámetros desconocidos (estimados libremente) excede el número de información conocida (es decir, elementos de la matriz de varianza-covarianza de entrada). Un modelo no identificado no se puede resolver porque hay un número infinito de estimaciones de parámetros que dan como resultado un ajuste perfecto del modelo. Un modelo se sobreidentifica cuando la cantidad de conocimientos (es decir, la cantidad de variaciones y covarianzas en la matriz de entrada) excede la cantidad de parámetros del modelo estimados libremente (Brown, 2015).

***Estimación de los parámetros del modelo CFA.*** El objetivo de CFA es obtener estimaciones para cada parámetro del modelo de medición (es decir, cargas de factores, variaciones y covarianzas de factores, variaciones de error de indicador y posiblemente covarianzas de error) para producir una matriz de varianza-covarianza predicha

(simbolizada como  $\Sigma$ ) que se asemeja a la muestra de matriz de varianza-covarianza (simbolizada como S) lo más cerca posible. Este proceso implica una función de ajuste, una operación matemática para minimizar la diferencia entre  $\Sigma$  y S. Con mucho, la función de ajuste más utilizada en la investigación de CFA aplicada (y SEM en general) es ML (Brown, 2015). La función de ajuste que se minimiza en ML es:

$$F_{ML} = Ln|S| - Ln|\Sigma| + traza[(S)(\Sigma^{-1})] - p$$

Si el modelo incluye una o más variables indicadoras categóricas o si existe una no normalidad extrema, no es apropiado utilizar ML y existen varios métodos de estimación alternativos disponibles: (1) WLS, que se denomina distribución libre asintótica (ADF) en Amos 7.0; (2) mínimos cuadrados robustos ponderados (WLSMV); y (3) ULS. Sin embargo, cada uno de estos métodos de estimación tiene limitaciones. Para indicadores continuos no normales, se puede usar ML con SE robusto y  $\chi^2$  (MLM). En este momento, el programa Mplus tiene las mejores opciones para manejar datos categóricos debido a la disponibilidad del estimador WLSMV (Harrington, 2009). De los métodos de estimación que están ampliamente disponibles, incluso en Amos, ADF "estima el grado de sesgo y curtosis en los datos en bruto" y, por lo tanto, no hace suposiciones sobre la distribución de los datos (Kline, 2016). Aunque esto aborda el problema de la no normalidad en los datos, un inconveniente de este enfoque es que generalmente requiere tamaños de muestra muy grandes de 200 a 500 para modelos simples y miles de casos para modelos complejos. Además de los requisitos de tamaño de muestra, ADF o WLS no funcionan bien con datos categóricos, especialmente cuando las muestras no son lo suficientemente grandes (Brown, 2015).

### ***2.2.2 Índices de Ajuste del Análisis Factorial Confirmatorio***

***Estadísticas de prueba del modelo.*** Estas son las estadísticas de ajuste SEM originales. Generalmente son estadísticas de chi-cuadrado que prueban la hipótesis de ajuste exacto de que no hay diferencia entre las covarianzas predichas por el modelo, dadas las estimaciones de los parámetros, y la matriz de covarianza de la población. Rechazar esta hipótesis dice que (1) los datos contienen información de covarianza que habla en

contra del modelo, y (2) el investigador debe explicar las discrepancias modelo-datos que exceden las esperadas por error de muestreo (Kline, 2016). La prueba de chi-cuadrado que se acaba de describir es una prueba de aceptación-apoyo donde la hipótesis nula representa la creencia del investigador de que el modelo es correcto; por lo tanto, es el fracaso para rechazar la hipótesis nula, o la ausencia de significancia estadística (p. ej.,  $p \geq .05$ ), lo que respalda el modelo. Esta lógica es "al revés" de la prueba más típica de rechazo-apoyo donde el rechazo de la hipótesis nula (p. Ej.,  $P < .05$ ) apoya la teoría del investigador. De los dos, las pruebas de aceptación-apoyo son lógicamente más débiles porque el hecho de no refutar una afirmación (la hipótesis del ajuste exacto) no prueba que la afirmación sea verdadera. La baja potencia en las pruebas de aceptación-soporte significa que hay pocas posibilidades de detectar un modelo falso. Esto significa que analizar el modelo en una muestra que es demasiado pequeña (es decir, de baja potencia) hace que sea más probable que se retenga el modelo. Sin embargo, en las pruebas de rechazo-apoyo, la penalización por baja potencia debido a un tamaño de muestra insuficiente es que las hipótesis del investigador tienen menos probabilidades de ser respaldadas. De esta manera, una estadística de prueba del modelo es como un detector de humo: si suena la alarma, puede haber o no un incendio (discrepancia grave entre modelo y datos), pero es prudente tratar la alarma con seriedad (realice una evaluación más detallada de ajuste).

***Índices aproximados de ajuste.*** Los índices de ajuste aproximados no son pruebas de significación, por lo que no hay una decisión binaria sobre si rechazar o retener una hipótesis nula al igual que no hay una demarcación de los límites del error de muestreo. En cambio, estos índices están destinados a ser medidas continuas de correspondencia modelo-datos. Algunos se escalan como estadísticas de "maldad de ajuste" donde los valores más altos indican peor ajuste, pero otros son medidas de "bondad de ajuste" donde los valores más altos indican un mejor ajuste. Los valores de algunos índices de bondad de ajuste están más o menos estandarizados, de modo que su rango es 0-1.0, donde un valor de 1.0 indica el mejor resultado. A continuación, se describen cuatro categorías de índices de ajuste aproximados (Kline, 2016). Estas categorías no son mutuamente excluyentes porque algunos índices se pueden clasificar en más de uno:



1. *Los índices de ajuste absoluto* miden qué tan bien un modelo a priori explica los datos. Ese modelo es el modelo del investigador porque no hay otro punto de referencia para un índice de ajuste absoluto. Explicar los datos no dice por sí solo que el modelo sea adecuado. Esto se debe a que se puede hacer un modelo mal especificado para explicar los datos agregando parámetros libres al punto donde no quedan grados de libertad ( $df_M = 0$ ); es decir, la mayoría de los modelos recién identificados explicarían perfectamente las covarianzas observadas.
2. *Los índices de ajuste incrementales* (relativos, comparativos) miden la mejora relativa en el ajuste del modelo del investigador sobre el de un modelo de referencia. El modelo de referencia suele ser el modelo de independencia (nulo), que supone covarianzas de cero entre las variables endógenas. El modelo nulo en Mplus usa las covarianzas de muestra entre las variables exógenas, pero otros programas de computadora, como EQS y LISREL, también fijan las covarianzas entre pares de variables exógenas medidas a cero. Verifique la definición del modelo nulo en la documentación de su herramienta informática SEM. La suposición de cero covarianzas es a menudo inverosímil. Esta es la razón por la que los índices de ajuste incremental basados en el modelo nulo "efectivamente dicen, ¿Cómo está mi modelo, en comparación con el peor modelo que existe?. Esto significa que el modelo nulo es un argumento "hombre de paja" que probablemente sea falso.
3. *Los índices ajustados por parsimonia* incluyen en sus fórmulas una corrección o "penalización" por la complejidad del modelo. La misma pena puede ser vista como una "recompensa" por parsimonia. El modelo de parsimonia se define como la escasez de parámetros libremente estimados en relación con el número de observaciones. Parsimonia está relacionado con  $df_M$ , pero los dos no son sinónimos. Esto se debe a que  $df_M$  no es una medida proporcional de la relación entre observaciones y parámetros. Por ejemplo, el valor de  $df_M$  puede ser relativamente alto cuando hay muchas variables observadas, pero aún puede ser proporcionalmente pequeño en comparación con el número de observaciones cuando hay muchos parámetros libres en un modelo muy complejo. Se define la relación de parsimonia (PR) como la relación de  $df_M$  del modelo del investigador sobre los grados de libertad del modelo nulo. El modelo ideal (el más parsimonioso) tendría un PR de 1.0, que dice que el

modelo tiene tantos grados de libertad como hay disponibles en los datos (según el modelo nulo). Dichos modelos son potencialmente más no confirmables con los datos que los modelos en los que el valor del RP es inferior a 1.0.

4. Los índices de ajuste predictivo estiman el ajuste del modelo en muestras de réplica hipotéticas del mismo tamaño y extraídas aleatoriamente de la misma población que la muestra original. Por lo tanto, estos índices pueden verse como basados en la población en lugar de basados en la muestra. También pueden corregir los grados de libertad del modelo o el tamaño de la muestra. Existe un contexto específico para los índices de ajuste predictivo que se describe más adelante en este capítulo, pero la mayoría de las aplicaciones de SEM no se incluyen en él.

Las fórmulas de algunos índices de ajuste aproximados incluyen estadísticas de prueba del modelo. Existe una relación similar en los análisis de datos más estándar entre las estadísticas de prueba y las medidas del tamaño del efecto: muchos tamaños de efectos se pueden expresar como funciones de las estadísticas de prueba y viceversa (Kline, 2016). La relación entre las estadísticas de prueba del modelo y los índices de ajuste aproximados significa que ambos se basan en los mismos supuestos de distribución. Si estos supuestos son insostenibles, entonces los valores tanto del índice de ajuste aproximado como del estadístico de prueba correspondiente (y su valor  $p$ ) pueden ser inexactos.

### ***2.2.3 Tamaño de muestra***

Aunque los investigadores coinciden en que cuanto mayor sea el tamaño de la muestra, mejor para el CFA, no hay un acuerdo universal sobre qué tan grande es lo suficientemente grande. Además, algunos procedimientos de estimación (como el ADF utilizado para datos no normales) requieren tamaños de muestra aún mayores que los necesarios con datos distribuidos normalmente y estimación de ML. Desafortunadamente, no hay una manera fácil de determinar el tamaño de muestra necesario para CFA. Se discuten brevemente los cuatro enfoques existentes: reglas generales, el método Satorra-Saris, la aproximación MacCallum y la aproximación de Monte Carlo (Harrington, 2009).

***Reglas generales.*** Existen algunas pautas muy generales para los tamaños de muestra: menos de 100 se considera "pequeño" y puede que solo sea apropiado para

modelos muy simples; 100 a 200 es "medio" y puede ser un tamaño de muestra mínimo aceptable si el modelo no es demasiado complejo; y más de 200 es "grande", lo que probablemente sea aceptable para la mayoría de los modelos (Kline, 2016). Los tamaños de muestra pequeños pueden ocasionar problemas técnicos al ejecutar el análisis (como la no convergencia o soluciones inadecuadas) y baja potencia (Kline, 2016).

Para los modelos CFA con tamaños de muestra pequeños o medianos, Kline (2016) sugiere utilizar indicadores con buenas propiedades psicométricas y cargas de factores estandarizadas mayores de 0,60, imponiendo restricciones de igualdad (es decir, forzando parámetros a ser iguales en lugar de libremente estimado) sobre las cargas no estandarizadas de indicadores sobre el mismo factor, y el uso de parcelas (es decir, grupos de indicadores) para elementos categóricos. Sin embargo, se recomienda no usar parcelas de artículos o cualquiera de las múltiples reglas generales para determinar cuál debería ser el tamaño de la muestra (Brown, 2015).

Lee y Song (2004) llevaron a cabo un estudio de simulación comparando ML y estimación bayesiana con pequeños tamaños de muestra. Llegaron a la conclusión de que ML no se recomienda con tamaños de muestra pequeños, incluso cuando los datos se distribuyen normalmente; sin embargo, se recomienda el enfoque bayesiano con tamaños de muestra pequeños siempre que el tamaño de la muestra sea dos o tres veces mayor que el número de parámetros desconocidos que se estimarán. El rendimiento del enfoque bayesiano mejora con muestras más grandes y produce estimaciones precisas de los parámetros y una prueba confiable de bondad de ajuste (Lee & Song, 2004) cuando la relación del tamaño de la muestra a los parámetros es 4: 1 o 5: 1. Los hallazgos de Lee y Song (2004) sugieren que, en algunos casos, estas reglas generales pueden proporcionar una guía razonable para las estimaciones del tamaño de la muestra, al menos para los datos distribuidos normalmente. Sin embargo, se advierten que ninguna regla general es aplicable en todas las situaciones porque el tamaño de muestra necesario para un estudio depende de muchos factores, incluido el tamaño del modelo, la distribución de las variables, la cantidad de datos faltantes, confiabilidad de las variables y fortaleza de las relaciones entre las variables (Muthén & Muthén, 2002). Finalmente, Gignac (2006) sugiere que los requisitos de tamaño de la muestra deben tratarse como recomendaciones

para ser probadas y que en algunos casos el SEM puede ser apropiado en tamaños de muestra tan pequeños como 100 (Gignac, 2006).

**Método Satorra-Saris.** Satorra y Saris (1985) citado por Brown (2015) introdujeron el primer enfoque más conocido para realizar análisis de potencia para SEM de indicadores múltiples. Este enfoque se centra en la potencia de la prueba de diferencia de  $\chi^2$  para detectar errores de especificación asociados con un solo parámetro. El investigador especifica un modelo asociado con una hipótesis nula ( $H_0$ ) que se evalúa en comparación con un modelo que representa la hipótesis alternativa ( $H_1$ ). En términos más simples,  $H_1$  refleja el modelo "verdadero" (es decir, contiene los valores de población verdaderos para todos los parámetros), y  $H_0$  es idéntico a  $H_1$  a excepción de los parámetros a probar. Por lo tanto, se genera una matriz de covarianza de la población a partir de las estimaciones de parámetros de  $H_1$ . Esta matriz de covarianza se utiliza como entrada en la prueba de  $H_0$ , que contiene el parámetro mal especificado y el tamaño de muestra de interés. En la práctica, el parámetro mal especificado se suele fijar a cero y, por lo tanto, el análisis se centra en la potencia del modelo  $\chi^2$  para detectar que un parámetro es diferente de cero. La prueba de  $H_0$  produce un valor de modelo  $\chi^2$  distinto de cero resultantes del parámetro mal especificado. Este valor representa un parámetro de no centralidad (NCP, a menudo simbolizado como  $l$ , pero no debe confundirse con un factor de carga) de la distribución no central de  $\chi^2$ . La distribución  $\chi^2$  no central refleja la distribución  $\chi^2$  cuando la hipótesis nula es falsa. Con el valor NCP resultante, la potencia de la prueba se puede determinar a partir de valores tabulados o rutinas simples en paquetes de software comerciales como SPSS y SAS. El método Satorra-Saris es superior a las reglas generales porque es un enfoque cuantitativo basado en modelos para determinar el tamaño de muestra requerido. Sin embargo, este enfoque tiene varias desventajas ya que el método requiere que el investigador especifique los valores alternativos que se probarán. Satorra demostró más tarde que los índices de modificación pueden usarse como valores aproximados de NCP para cada restricción en el modelo.

**Aproximación MacCallum.** Maccallum considera que el análisis de potencia debe basarse en el error cuadrático medio del índice de ajuste de estimación, en lugar del GFI debido a la influencia indeseable de los grados de libertad en los análisis de poder basados

en GFI. Además, se examina el papel del error del modelo en los requisitos de tamaño de muestra para el análisis factorial y concluyeron que las reglas tradicionales del tamaño de la muestra son de valor limitado y EFA puede funcionar bien cuando las comunidades son altas, casi independientemente del tamaño de la muestra, el error del modelo o el nivel de sobre-determinación de los factores.

*Aproximación Monte Carlo.* Muthén y Muthén (2002) demostraron el uso de un estudio de Monte Carlo con Mplus para identificar el tamaño de muestra apropiado y determinar la potencia para CFA. Los estudios de Monte Carlo se realizan generando datos de una población con valores de parámetros hipotéticos. Se extraen varias muestras, se estima el modelo para cada muestra y los valores de parámetros y errores estándar se promedian sobre las muestras (Muthén & Muthén, 2002). Antes de realizar el estudio de Monte Carlo, el investigador debe decidir sobre el modelo que se probará y seleccionar los valores de población para cada parámetro en el modelo. Las estimaciones pueden provenir de la teoría o la investigación previa, y las mejores estimaciones provienen de estudios anteriores. Luego, el investigador debe decidir cuántas muestras se extraerán (es decir, cuántas réplicas habrá) y seleccionar múltiples valores de semillas (es decir, los puntos de partida para la selección aleatoria de las muestras). Debe haber suficientes repeticiones para lograr resultados estables, y Muthén y Muthén (2002) usaron 10,000 repeticiones en su estudio. Antes de entrar en pánico y decidir que esto es abrumador e imposible, recuerde que el software de la computadora hará la gran mayoría del trabajo por usted. En realidad, no ejecutará 10,000 repeticiones, sino que programará el software para que lo haga por usted. Muthén y Muthén (2002) proporcionan los archivos Mplus para realizar los análisis que informan.

Una ventaja importante del enfoque de Monte Carlo es que permite al investigador considerar aspectos importantes de los datos, como la falta y la normalidad. Se pueden especificar patrones y cantidades de datos faltantes, asimetría y curtosis. En su estudio de Monte Carlo, Muthén y Muthén (2002) utilizaron la estimación de ML; Se utilizó un estimador que es robusto a la no normalidad con los datos no normales para estimar los SE. Para su modelo CFA, Muthén y Muthén (2002) encontraron que se necesitaba un tamaño de muestra de 150 cuando los datos se distribuían normalmente y no faltaban

datos. El tamaño de muestra necesario aumentó a 175 cuando había datos faltantes al azar, 265 para datos completos no normales y 315 cuando los datos no eran normales y tenían alguna falta. Estos resultados indican claramente el impacto de la distribución y la falta de datos en los requisitos de tamaño de muestra necesarios para una potencia adecuada.

## **2.3 Marcos conceptuales o glosario**

### ***2.3.1 PQH-9.***

El cuestionario de salud del paciente (PHQ) es una herramienta de diagnóstico para trastornos de salud mental utilizada por profesionales de la salud que es rápida y fácil de completar para los pacientes (Borghero et al., 2018).

### ***2.3.2 SEM***

El término modelado de ecuaciones estructurales (SEM) no designa una técnica estadística única, sino que se refiere a una familia de procedimientos relacionados. Otros términos como análisis de estructura de covarianza, modelado de estructura de covarianza o análisis de estructuras de covarianza también se utilizan en la literatura para clasificar estas técnicas bajo una sola etiqueta (Kline, 2016).

### ***2.3.3 Causalidad probabilística***

Los modelos analizados en SEM generalmente asumen causalidad probabilística, no causalidad determinista. Esto último significa que, dado un cambio en una variable causal, se observa la misma consecuencia en todos los casos en la variable de resultado. En contraste, la causalidad probabilística permite que ocurran cambios en los resultados con alguna probabilidad  $<1.0$ . La estimación de estas probabilidades (efectos) con datos de muestra generalmente se basa en supuestos de distribución específicos, como la normalidad. La causalidad como una relación funcional entre dos variables cuantitativas se conserva en este punto de vista, pero se supone que los efectos causales cambian una distribución de probabilidad (Kline, 2016).

#### ***2.3.4 Distribución normal multivariante***

Una generalización de la distribución normal univariada a más de una variable es la distribución normal multivariada. Gran parte de la teoría estadística básica que subyace a la inferencia estadística en el análisis factorial y el modelado de ecuaciones estructurales se basa en la distribución normal multivariada. Algunas de las razones de esto es que a menudo es una excelente aproximación a muchas distribuciones del mundo real, mientras que al mismo tiempo es matemáticamente manejable. También se presta bien a las relaciones funcionales lineales entre variables porque las transformaciones lineales de variables multivariadas, normalmente distribuidas, son variables que también tienen distribuciones normales multivariadas. Y las distribuciones de muestreo de muchas estadísticas multivariadas, incluidas las de análisis factorial, se comportan aproximadamente como distribuciones normales multivariadas debido al efecto descrito por el teorema del límite central (Mulaik, 2010).

## **CAPÍTULO 3: METODOLOGÍA**

### **3.1 Tipo y diseño de investigación**

Investigación aplicada, de diseño experimental.

### **3.2 Unidad de análisis**

Un registro de la Encuesta ENDES-INEI

### **3.3 Población de estudio**

La población se encuentra conformado por 34476 registros de la encuesta ENDES-INEI 2018

### **3.4 Tamaño de muestra**

El tamaño de muestra estará conformado por muestra elegidas mediante el muestreo Bootstrap con 100 réplicas en cantidad de 50, 100, 500, 1000 y 5000 tamaño de muestra.

### **3.5 Selección de muestra**

Muestreo aleatorio simple mediante Bootstrap.



### **3.6 Técnicas de recolección de Datos**

Se realizo mediante la técnica de minería de datos por intermedio de la metodología CRISP-DM.

### **3.7 Análisis e interpretación de la información**

Se aplico el análisis factorial confirmatorio y se obtendrán las medidas de bondad de ajuste para cada uno de los tamaños de muestra.

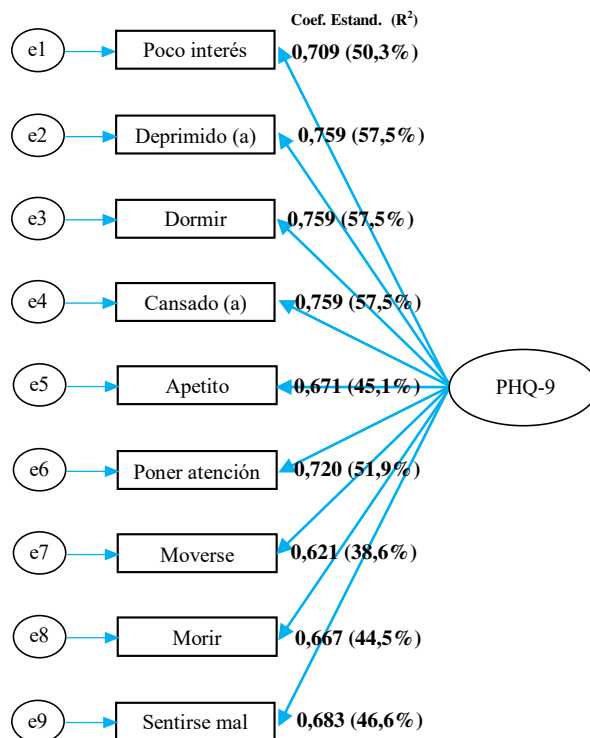
## CAPÍTULO 4: RESULTADOS Y DISCUSIÓN

### 4.1 Resultados

La estimación del modelo paramétrico se realizó mediante el método de máxima verosimilitud.

*Figura 1*

**Parámetros del análisis factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.**



Las medidas de bondad de ajuste para el modelo factorial confirmatorio paramétrico del cuestionario PHQ-9 aplicado a la población peruana, se presentan en la tabla 1.

**Tabla 1**

*Medidas de bondad de ajuste del modelo factorial confirmatorio paramétrico sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018*

| <b>Índice de bondad de Ajuste</b> | <b>Valor</b>      |            |
|-----------------------------------|-------------------|------------|
|                                   | Chi-Cuadrado      | 9020,030   |
|                                   | Chi-Cuadrado / GL | 334,075    |
| Absoluto                          | GFI               | 0,943      |
|                                   | RMR               | 0,015      |
|                                   | SRMR              | 0,047      |
|                                   | RMSEA             | 0,098      |
|                                   |                   |            |
| Parsimónico                       | AGFI              | 0,906      |
|                                   | PNFI              | 0,694      |
| Comparativo                       | AIC               | 478202,180 |
|                                   | BIC               | 478354,244 |
| Incremental                       | CFI               | 0,925      |
|                                   | TLI               | 0,901      |
|                                   | RNI               | 0,925      |
|                                   | NFI               | 0,925      |
|                                   | NNFI              | 0,901      |

Las medidas de bondad de ajuste muestran que tan bueno es el modelo de medida estimado, observándose en los resultados paramétricos que muchas de estas medidas presentan un ajuste aceptable, ya que todos ellos fueron influenciados por el tamaño de los casos analizados y el método de estimación de los parámetros, que en este caso fue la estimación de máxima verosimilitud.

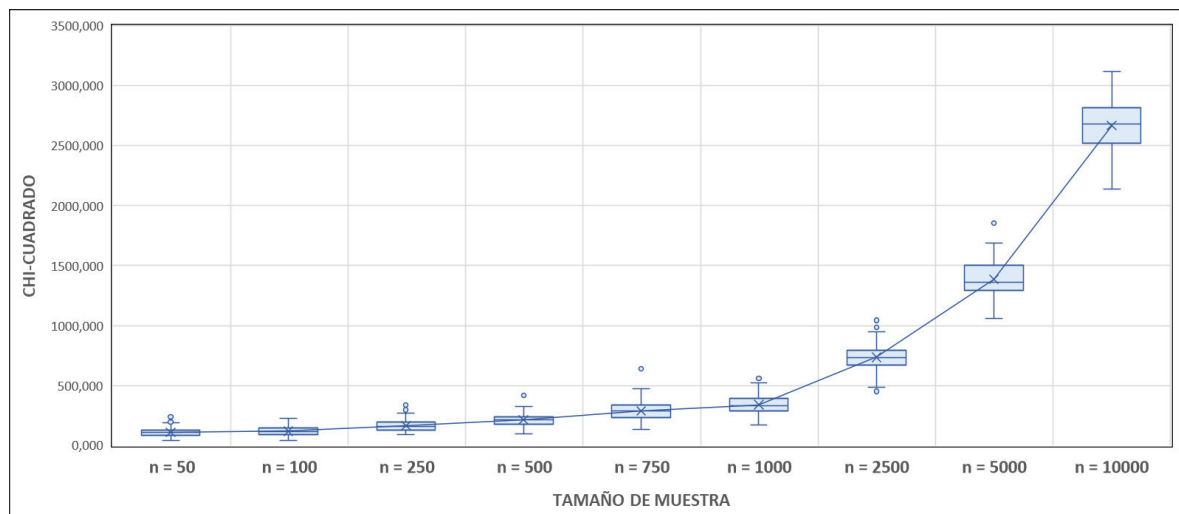
**Tabla 2**

***Chi-Cuadrado del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018***

| Estadístico         | Tamaño de Muestra |         |         |         |         |         |          |          |          |
|---------------------|-------------------|---------|---------|---------|---------|---------|----------|----------|----------|
|                     | 50                | 100     | 250     | 500     | 750     | 1000    | 2500     | 5000     | 10000    |
| Promedio            | 109,441           | 118,219 | 164,294 | 214,011 | 287,769 | 339,317 | 736,261  | 1385,153 | 2668,674 |
| Desviación estándar | 40,265            | 37,241  | 50,108  | 55,409  | 77,780  | 75,808  | 114,109  | 156,248  | 213,286  |
| Mínimo              | 42,792            | 41,695  | 89,105  | 93,812  | 132,612 | 170,882 | 450,894  | 1059,099 | 2140,918 |
| Cuartil 1           | 81,486            | 90,195  | 127,800 | 175,977 | 231,244 | 287,499 | 667,428  | 1295,211 | 2527,700 |
| Mediana             | 107,505           | 115,002 | 156,481 | 214,920 | 286,169 | 332,480 | 734,702  | 1361,012 | 2681,037 |
| Cuartil 3           | 125,558           | 143,722 | 193,014 | 240,935 | 333,609 | 391,969 | 791,554  | 1497,192 | 2814,686 |
| Máximo              | 276,397           | 225,128 | 337,581 | 416,918 | 641,335 | 557,486 | 1082,400 | 1856,077 | 3119,056 |

**Figura 2**

**Chi-Cuadrado del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.**



La medida de Chi-cuadrado presenta una medida creciente en función al crecimiento del tamaño de muestra.

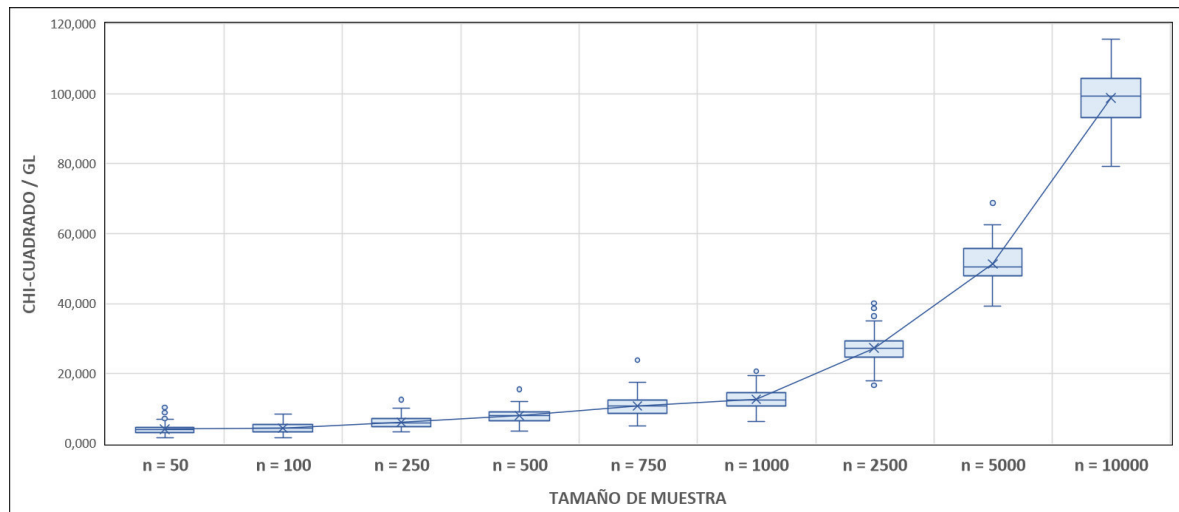
**Tabla 3**

*Chi-Cuadrado/gl del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018*

| Estadístico         | Tamaño de Muestra |       |        |        |        |        |        |        |         |
|---------------------|-------------------|-------|--------|--------|--------|--------|--------|--------|---------|
|                     | 50                | 100   | 250    | 500    | 750    | 1000   | 2500   | 5000   | 10000   |
| Promedio            | 4,053             | 4,378 | 6,085  | 7,926  | 10,658 | 12,567 | 27,269 | 51,302 | 98,840  |
| Desviación estándar | 1,491             | 1,379 | 1,856  | 2,052  | 2,881  | 2,808  | 4,226  | 5,787  | 7,899   |
| Mínimo              | 1,585             | 1,544 | 3,300  | 3,475  | 4,912  | 6,329  | 16,700 | 39,226 | 79,293  |
| Cuartil 1           | 3,018             | 3,341 | 4,733  | 6,518  | 8,565  | 10,648 | 24,720 | 47,971 | 93,619  |
| Mediana             | 3,982             | 4,259 | 5,796  | 7,960  | 10,599 | 12,314 | 27,211 | 50,408 | 99,298  |
| Cuartil 3           | 4,650             | 5,323 | 7,149  | 8,924  | 12,356 | 14,517 | 29,317 | 55,452 | 104,248 |
| Máximo              | 10,237            | 8,338 | 12,503 | 15,441 | 23,753 | 20,648 | 40,089 | 68,744 | 115,521 |

**Figura 3**

**Chi-Cuadrado/gl del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.**



La medida de Chi-cuadrado/gl presenta una medida creciente en función al crecimiento del tamaño de muestra.

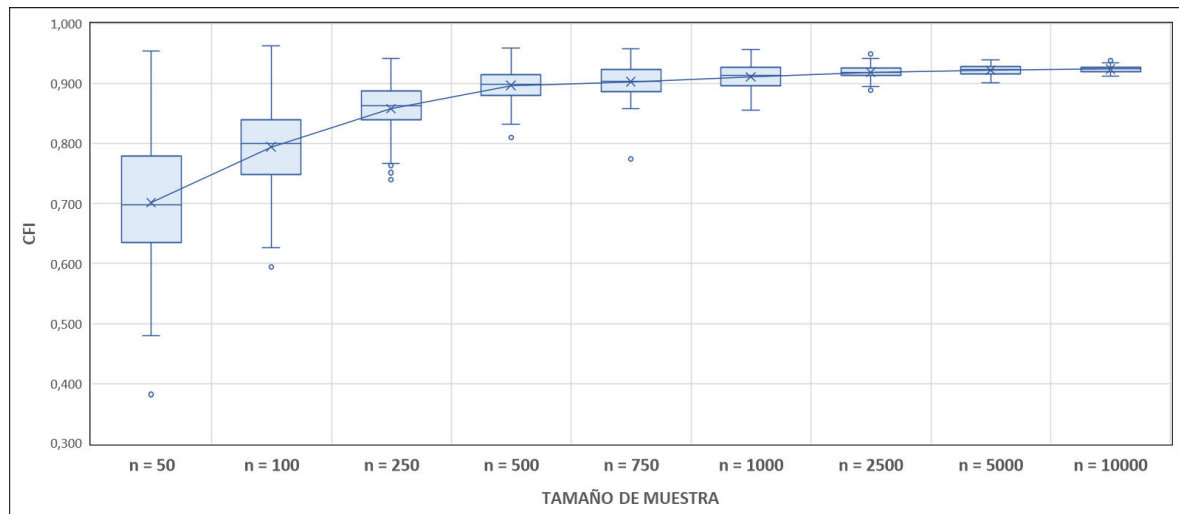
**Tabla 4**

*CFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018*

| Estadístico         | Tamaño de Muestra |       |       |       |       |       |       |       |       |
|---------------------|-------------------|-------|-------|-------|-------|-------|-------|-------|-------|
|                     | 50                | 100   | 250   | 500   | 750   | 1000  | 2500  | 5000  | 10000 |
| Promedio            | 0,702             | 0,795 | 0,859 | 0,897 | 0,903 | 0,911 | 0,919 | 0,923 | 0,925 |
| Desviación estandar | 0,109             | 0,073 | 0,043 | 0,029 | 0,027 | 0,021 | 0,012 | 0,009 | 0,006 |
| Mínimo              | 0,382             | 0,594 | 0,740 | 0,810 | 0,775 | 0,855 | 0,889 | 0,902 | 0,913 |
| Cuartil 1           | 0,635             | 0,749 | 0,840 | 0,884 | 0,887 | 0,896 | 0,914 | 0,917 | 0,921 |
| Mediana             | 0,698             | 0,800 | 0,864 | 0,900 | 0,904 | 0,914 | 0,919 | 0,924 | 0,925 |
| Cuartil 3           | 0,780             | 0,839 | 0,888 | 0,915 | 0,924 | 0,928 | 0,926 | 0,928 | 0,928 |
| Máximo              | 0,955             | 0,963 | 0,943 | 0,960 | 0,959 | 0,957 | 0,952 | 0,940 | 0,939 |

**Figura 4**

**CFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.**



La medida de CFI presenta una medida creciente en función al crecimiento del tamaño de muestra, reduciéndose su intervalo de variación a medida que crece el tamaño de muestra.

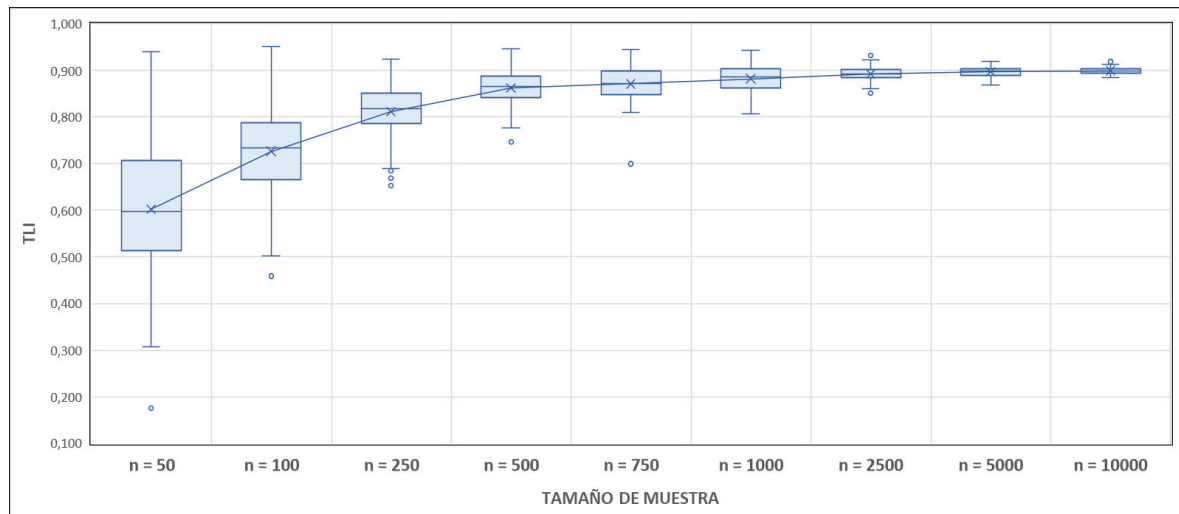
**Tabla 5**

*TLI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018*

| Estadístico         | Tamaño de Muestra |       |       |       |       |       |       |       |       |
|---------------------|-------------------|-------|-------|-------|-------|-------|-------|-------|-------|
|                     | 50                | 100   | 250   | 500   | 750   | 1000  | 2500  | 5000  | 10000 |
| Promedio            | 0,603             | 0,727 | 0,812 | 0,863 | 0,871 | 0,882 | 0,892 | 0,897 | 0,900 |
| Desviación estandar | 0,145             | 0,098 | 0,057 | 0,039 | 0,035 | 0,027 | 0,015 | 0,011 | 0,007 |
| Mínimo              | 0,176             | 0,459 | 0,653 | 0,747 | 0,700 | 0,807 | 0,852 | 0,869 | 0,884 |
| Cuartil 1           | 0,513             | 0,665 | 0,787 | 0,845 | 0,849 | 0,862 | 0,885 | 0,889 | 0,894 |
| Mediana             | 0,597             | 0,734 | 0,819 | 0,866 | 0,872 | 0,886 | 0,892 | 0,898 | 0,899 |
| Cuartil 3           | 0,706             | 0,786 | 0,851 | 0,886 | 0,899 | 0,904 | 0,902 | 0,904 | 0,904 |
| Máximo              | 0,940             | 0,951 | 0,924 | 0,946 | 0,945 | 0,943 | 0,936 | 0,920 | 0,919 |

**Figura 5**

**TLI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.**



La medida de TLI presenta una medida creciente en función al crecimiento del tamaño de muestra, reduciéndose su intervalo de variación a medida que crece el tamaño de muestra.

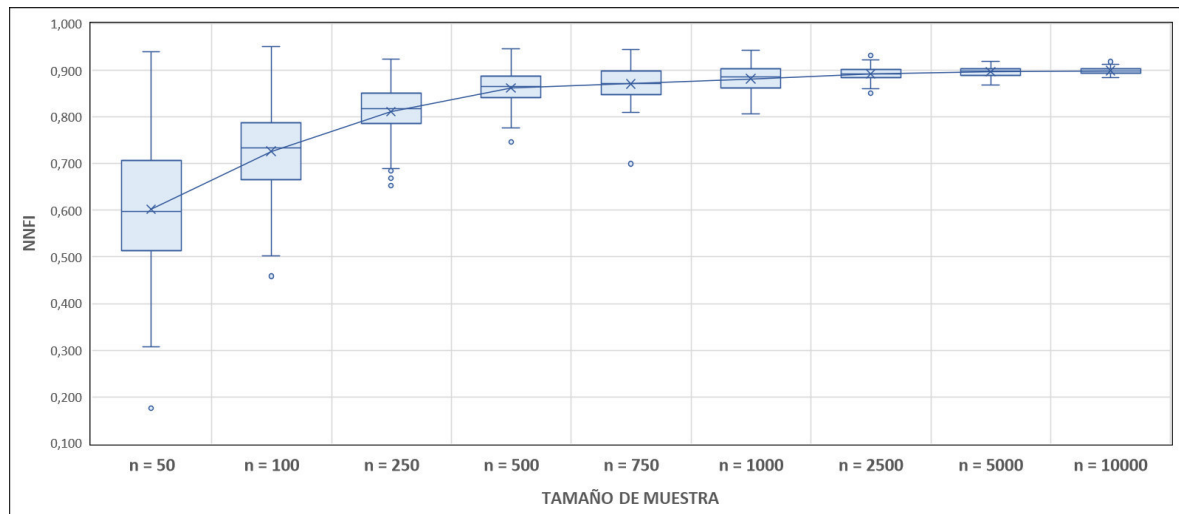
**Tabla 6**

*NNFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018*

| Estadístico         | Tamaño de Muestra |       |       |       |       |       |       |       |       |
|---------------------|-------------------|-------|-------|-------|-------|-------|-------|-------|-------|
|                     | 50                | 100   | 250   | 500   | 750   | 1000  | 2500  | 5000  | 10000 |
| Promedio            | 0,603             | 0,727 | 0,812 | 0,863 | 0,871 | 0,882 | 0,892 | 0,897 | 0,900 |
| Desviación estándar | 0,145             | 0,098 | 0,057 | 0,039 | 0,035 | 0,027 | 0,015 | 0,011 | 0,007 |
| Mínimo              | 0,176             | 0,459 | 0,653 | 0,747 | 0,700 | 0,807 | 0,852 | 0,869 | 0,884 |
| Cuartil 1           | 0,513             | 0,665 | 0,787 | 0,845 | 0,849 | 0,862 | 0,885 | 0,889 | 0,894 |
| Mediana             | 0,597             | 0,734 | 0,819 | 0,866 | 0,872 | 0,886 | 0,892 | 0,898 | 0,899 |
| Cuartil 3           | 0,706             | 0,786 | 0,851 | 0,886 | 0,899 | 0,904 | 0,902 | 0,904 | 0,904 |
| Máximo              | 0,940             | 0,951 | 0,924 | 0,946 | 0,945 | 0,943 | 0,936 | 0,920 | 0,919 |

**Figura 6**

**NNFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.**



La medida de NNFI presenta una medida creciente en función al crecimiento del tamaño de muestra, reduciéndose su intervalo de variación a medida que crece el tamaño de muestra.



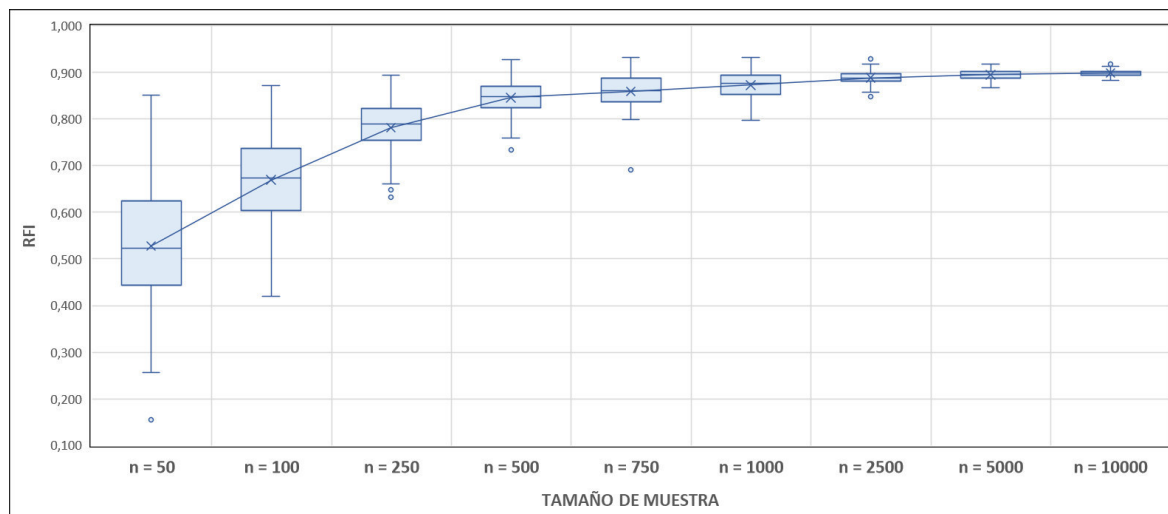
**Tabla 7**

*RFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018*

| Estadístico         | Tamaño de Muestra |       |       |       |       |       |       |       |       |
|---------------------|-------------------|-------|-------|-------|-------|-------|-------|-------|-------|
|                     | 50                | 100   | 250   | 500   | 750   | 1000  | 2500  | 5000  | 10000 |
| Promedio            | 0,528             | 0,669 | 0,782 | 0,846 | 0,860 | 0,873 | 0,889 | 0,895 | 0,899 |
| Desviación estándar | 0,129             | 0,092 | 0,056 | 0,038 | 0,035 | 0,027 | 0,015 | 0,011 | 0,007 |
| Mínimo              | 0,155             | 0,421 | 0,632 | 0,734 | 0,691 | 0,797 | 0,848 | 0,867 | 0,883 |
| Cuartil 1           | 0,444             | 0,606 | 0,756 | 0,828 | 0,837 | 0,853 | 0,882 | 0,887 | 0,894 |
| Mediana             | 0,523             | 0,674 | 0,789 | 0,848 | 0,861 | 0,877 | 0,888 | 0,896 | 0,898 |
| Cuartil 3           | 0,623             | 0,735 | 0,820 | 0,869 | 0,887 | 0,894 | 0,898 | 0,902 | 0,903 |
| Máximo              | 0,852             | 0,872 | 0,894 | 0,927 | 0,932 | 0,933 | 0,933 | 0,919 | 0,918 |

**Figura 7**

**RFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.**



La medida de RFI presenta una medida creciente en función al crecimiento del tamaño de muestra, reduciéndose su intervalo de variación a medida que crece el tamaño de muestra.

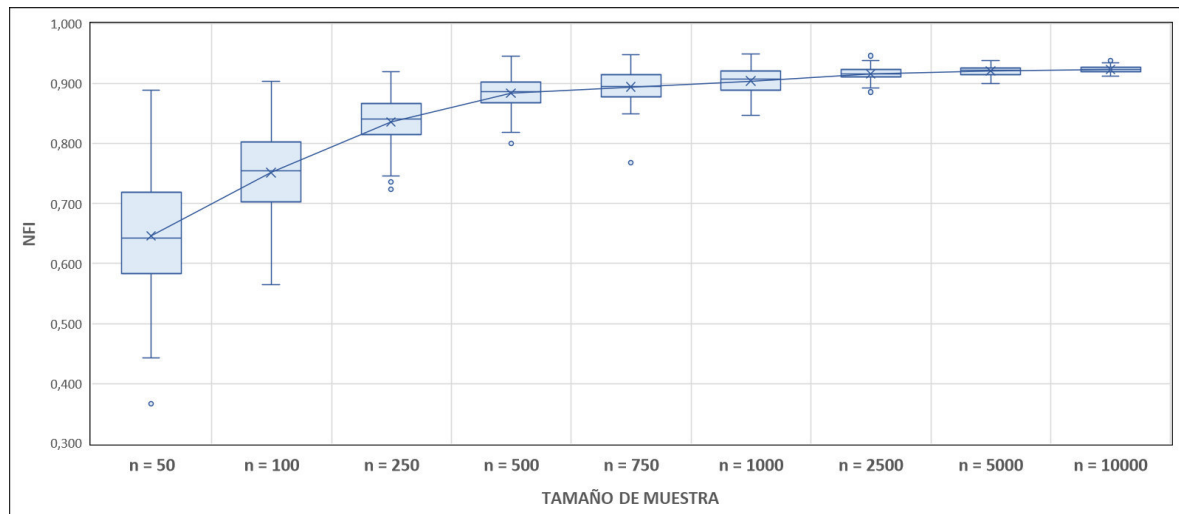
**Tabla 8**

*NFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018*

| Estadístico         | Tamaño de Muestra |       |       |       |       |       |       |       |       |
|---------------------|-------------------|-------|-------|-------|-------|-------|-------|-------|-------|
|                     | 50                | 100   | 250   | 500   | 750   | 1000  | 2500  | 5000  | 10000 |
| Promedio            | 0,646             | 0,752 | 0,836 | 0,884 | 0,895 | 0,905 | 0,916 | 0,921 | 0,924 |
| Desviación estándar | 0,097             | 0,069 | 0,042 | 0,029 | 0,026 | 0,020 | 0,011 | 0,009 | 0,006 |
| Mínimo              | 0,366             | 0,565 | 0,724 | 0,801 | 0,768 | 0,847 | 0,886 | 0,900 | 0,912 |
| Cuartil 1           | 0,583             | 0,704 | 0,817 | 0,871 | 0,878 | 0,890 | 0,911 | 0,916 | 0,920 |
| Mediana             | 0,643             | 0,755 | 0,842 | 0,886 | 0,896 | 0,908 | 0,916 | 0,922 | 0,924 |
| Cuartil 3           | 0,718             | 0,802 | 0,865 | 0,901 | 0,915 | 0,921 | 0,923 | 0,927 | 0,927 |
| Máximo              | 0,889             | 0,904 | 0,920 | 0,946 | 0,949 | 0,950 | 0,949 | 0,939 | 0,939 |

**Figura 8**

**NFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.**



La medida de NFI presenta una medida creciente en función al crecimiento del tamaño de muestra, reduciéndose su intervalo de variación a medida que crece el tamaño de muestra.

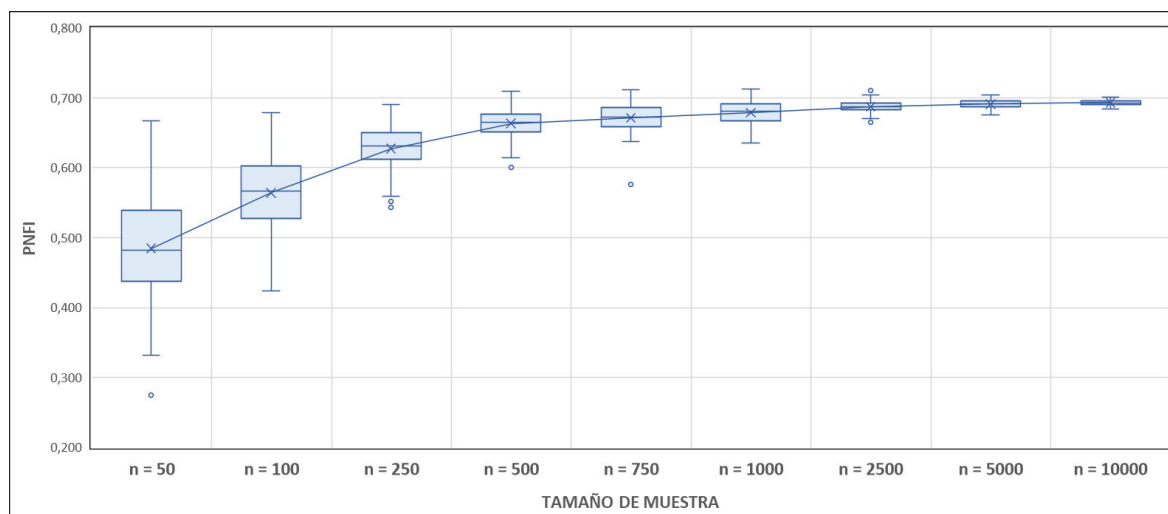
**Tabla 9**

*PNFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018*

| Estadístico         | Tamaño de Muestra |       |       |       |       |       |       |       |       |
|---------------------|-------------------|-------|-------|-------|-------|-------|-------|-------|-------|
|                     | 50                | 100   | 250   | 500   | 750   | 1000  | 2500  | 5000  | 10000 |
| Promedio            | 0,485             | 0,564 | 0,627 | 0,663 | 0,671 | 0,679 | 0,687 | 0,691 | 0,693 |
| Desviación estándar | 0,073             | 0,052 | 0,031 | 0,021 | 0,020 | 0,015 | 0,009 | 0,006 | 0,004 |
| Mínimo              | 0,275             | 0,424 | 0,543 | 0,601 | 0,576 | 0,636 | 0,665 | 0,675 | 0,684 |
| Cuartil 1           | 0,437             | 0,528 | 0,613 | 0,653 | 0,659 | 0,667 | 0,683 | 0,687 | 0,690 |
| Mediana             | 0,482             | 0,566 | 0,631 | 0,665 | 0,672 | 0,681 | 0,687 | 0,692 | 0,693 |
| Cuartil 3           | 0,538             | 0,601 | 0,649 | 0,676 | 0,686 | 0,691 | 0,693 | 0,695 | 0,695 |
| Máximo              | 0,667             | 0,678 | 0,690 | 0,709 | 0,712 | 0,712 | 0,712 | 0,704 | 0,704 |

**Figura 9**

**PNFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.**



La medida de PNFI presenta una medida creciente en función al crecimiento del tamaño de muestra, reduciéndose su intervalo de variación a medida que crece el tamaño de muestra.

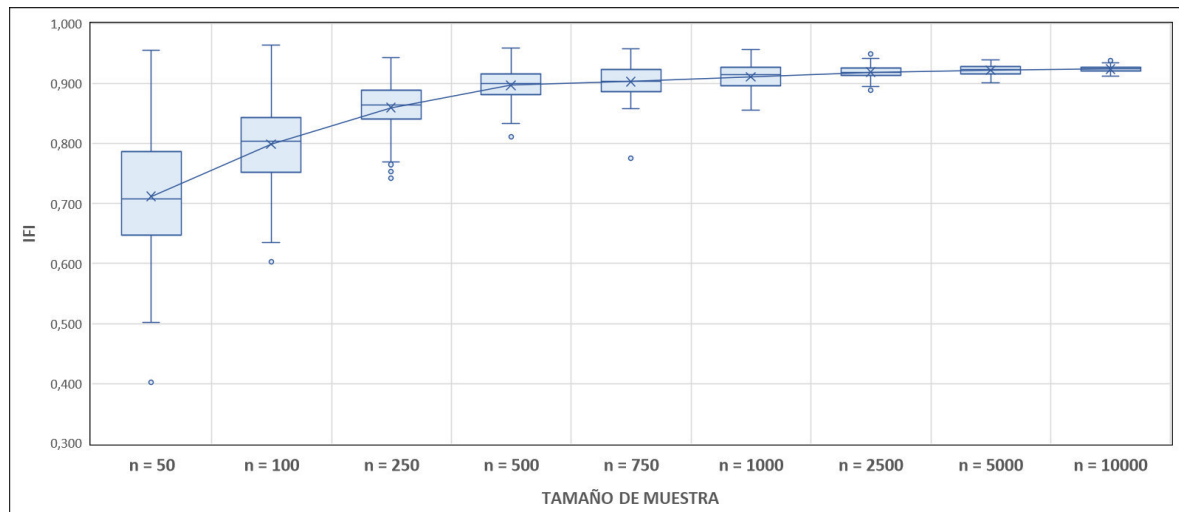
**Tabla 10**

*IFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018*

| Estadístico         | Tamaño de Muestra |       |       |       |       |       |       |       |       |
|---------------------|-------------------|-------|-------|-------|-------|-------|-------|-------|-------|
|                     | 50                | 100   | 250   | 500   | 750   | 1000  | 2500  | 5000  | 10000 |
| Promedio            | 0,712             | 0,799 | 0,860 | 0,898 | 0,904 | 0,912 | 0,919 | 0,923 | 0,925 |
| Desviación estándar | 0,105             | 0,072 | 0,043 | 0,029 | 0,027 | 0,020 | 0,012 | 0,009 | 0,006 |
| Mínimo              | 0,402             | 0,603 | 0,742 | 0,811 | 0,776 | 0,856 | 0,889 | 0,902 | 0,913 |
| Cuartil 1           | 0,648             | 0,753 | 0,842 | 0,884 | 0,887 | 0,897 | 0,914 | 0,917 | 0,921 |
| Mediana             | 0,708             | 0,804 | 0,865 | 0,900 | 0,905 | 0,915 | 0,919 | 0,924 | 0,925 |
| Cuartil 3           | 0,786             | 0,843 | 0,889 | 0,915 | 0,924 | 0,928 | 0,926 | 0,928 | 0,928 |
| Máximo              | 0,956             | 0,964 | 0,943 | 0,960 | 0,959 | 0,957 | 0,952 | 0,940 | 0,939 |

**Figura 10**

**IFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.**



La medida de IFI presenta una medida creciente en función al crecimiento del tamaño de muestra, reduciéndose su intervalo de variación a medida que crece el tamaño de muestra.

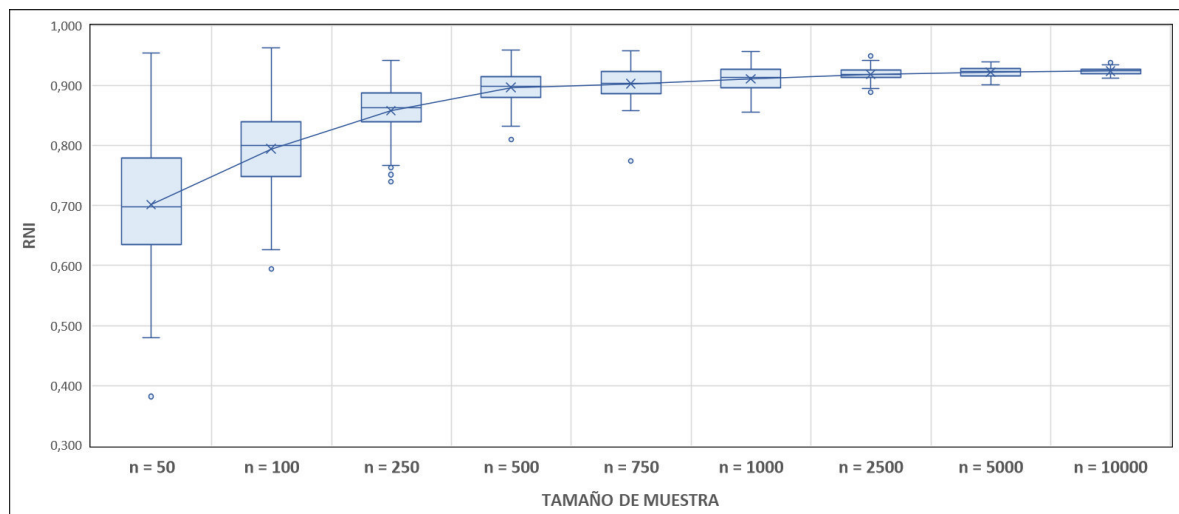
**Tabla 11**

*RNI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018*

| Estadístico         | Tamaño de Muestra |       |       |       |       |       |       |       |       |
|---------------------|-------------------|-------|-------|-------|-------|-------|-------|-------|-------|
|                     | 50                | 100   | 250   | 500   | 750   | 1000  | 2500  | 5000  | 10000 |
| Promedio            | 0,702             | 0,795 | 0,859 | 0,897 | 0,903 | 0,911 | 0,919 | 0,923 | 0,925 |
| Desviación estándar | 0,109             | 0,073 | 0,043 | 0,029 | 0,027 | 0,021 | 0,012 | 0,009 | 0,006 |
| Mínimo              | 0,382             | 0,594 | 0,740 | 0,810 | 0,775 | 0,855 | 0,889 | 0,902 | 0,913 |
| Cuartil 1           | 0,635             | 0,749 | 0,840 | 0,884 | 0,887 | 0,896 | 0,914 | 0,917 | 0,921 |
| Mediana             | 0,698             | 0,800 | 0,864 | 0,900 | 0,904 | 0,914 | 0,919 | 0,924 | 0,925 |
| Cuartil 3           | 0,780             | 0,839 | 0,888 | 0,915 | 0,924 | 0,928 | 0,926 | 0,928 | 0,928 |
| Máximo              | 0,955             | 0,963 | 0,943 | 0,960 | 0,959 | 0,957 | 0,952 | 0,940 | 0,939 |

**Figura 11**

**RNI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.**



La medida de RNI presenta una medida creciente en función al crecimiento del tamaño de muestra, reduciéndose su intervalo de variación a medida que crece el tamaño de muestra.

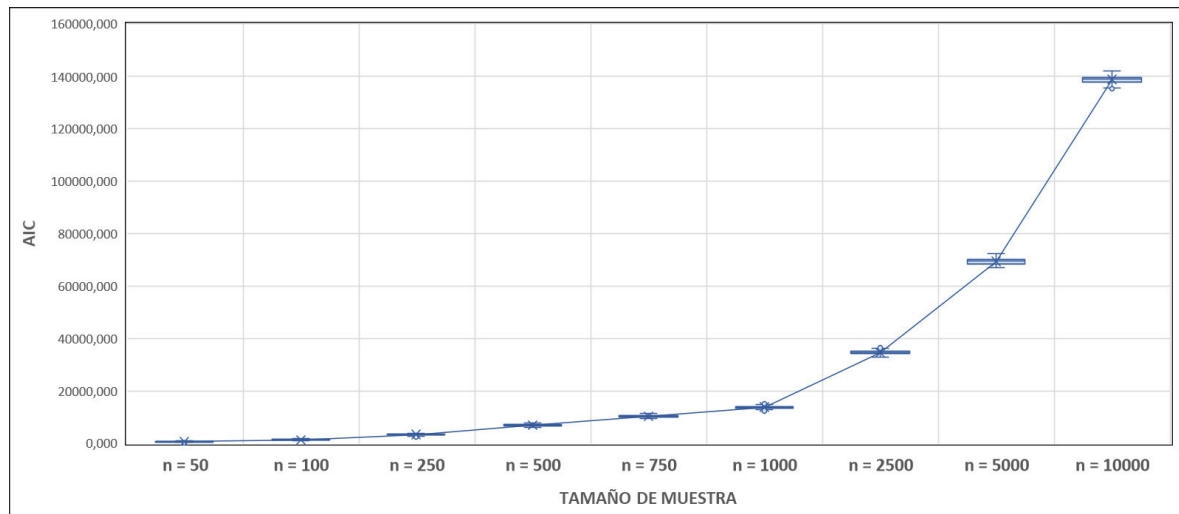
**Tabla 12**

*AIC del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018*

| Estadístico         | Tamaño de Muestra |          |          |          |           |           |           |           |            |
|---------------------|-------------------|----------|----------|----------|-----------|-----------|-----------|-----------|------------|
|                     | 50                | 100      | 250      | 500      | 750       | 1000      | 2500      | 5000      | 10000      |
| Promedio            | 653,514           | 1339,226 | 3367,650 | 6990,734 | 10376,478 | 13778,772 | 34612,060 | 69404,575 | 138779,581 |
| Desviación estándar | 130,626           | 159,591  | 290,982  | 350,798  | 383,221   | 470,807   | 798,502   | 1213,662  | 1312,788   |
| Mínimo              | 195,453           | 951,294  | 2479,322 | 6186,583 | 9410,259  | 12420,309 | 33007,897 | 67048,977 | 135324,879 |
| Cuartil 1           | 586,929           | 1226,655 | 3211,412 | 6747,572 | 10117,949 | 13495,692 | 34227,948 | 68433,099 | 137964,269 |
| Mediana             | 673,931           | 1332,592 | 3384,999 | 6967,546 | 10351,319 | 13821,452 | 34606,520 | 69416,680 | 138859,771 |
| Cuartil 3           | 745,317           | 1460,128 | 3561,265 | 7249,229 | 10614,960 | 14055,543 | 35064,949 | 70134,658 | 139616,537 |
| Máximo              | 909,643           | 1750,506 | 3876,600 | 7792,870 | 11359,648 | 15043,207 | 37218,030 | 72353,027 | 141919,388 |

**Figura 12**

**AIC del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.**



La medida de AIC presenta una medida creciente en función al crecimiento del tamaño de muestra.

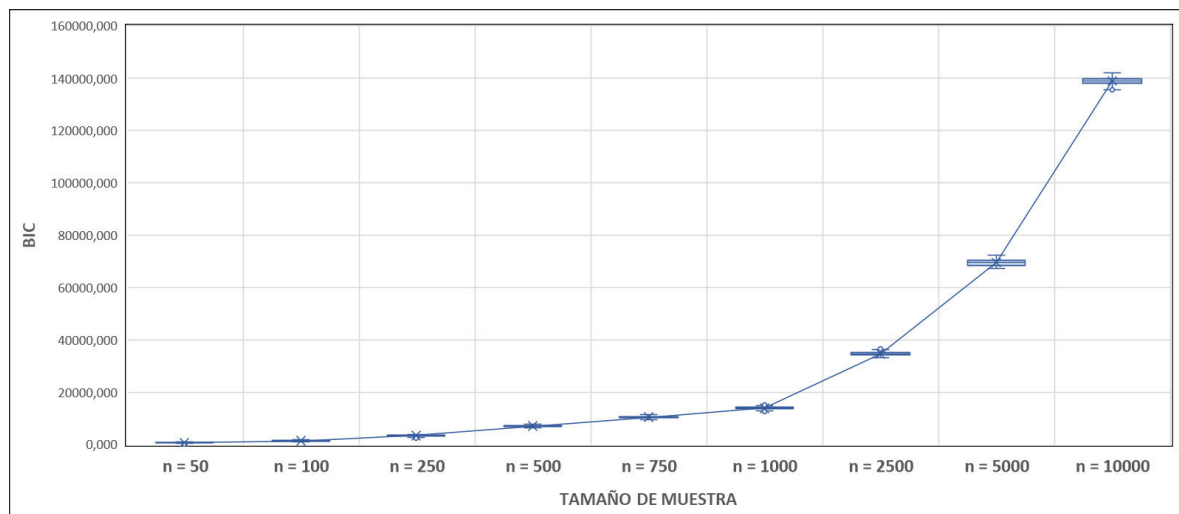
**Tabla 13**

*BIC del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018*

| Estadístico         | Tamaño de Muestra |          |          |          |           |           |           |           |            |
|---------------------|-------------------|----------|----------|----------|-----------|-----------|-----------|-----------|------------|
|                     | 50                | 100      | 250      | 500      | 750       | 1000      | 2500      | 5000      | 10000      |
| Promedio            | 687,930           | 1386,119 | 3431,036 | 7066,597 | 10459,639 | 13867,111 | 34716,893 | 69521,884 | 138909,367 |
| Desviación estándar | 130,626           | 159,591  | 290,982  | 350,798  | 383,221   | 470,807   | 798,502   | 1213,662  | 1312,788   |
| Mínimo              | 229,870           | 998,187  | 2542,708 | 6262,446 | 9493,420  | 12508,649 | 33112,730 | 67166,286 | 135454,665 |
| Cuartil 1           | 621,345           | 1273,548 | 3274,798 | 6823,435 | 10201,110 | 13584,032 | 34332,781 | 68550,408 | 138094,055 |
| Mediana             | 708,348           | 1379,485 | 3448,385 | 7043,409 | 10434,480 | 13909,792 | 34711,353 | 69533,990 | 138989,557 |
| Cuartil 3           | 779,733           | 1507,021 | 3624,651 | 7325,092 | 10698,121 | 14143,883 | 35169,782 | 70251,967 | 139746,323 |
| Máximo              | 944,059           | 1797,399 | 3939,987 | 7868,733 | 11442,810 | 15131,546 | 37322,863 | 72470,337 | 142049,174 |

**Figura 13**

**BIC del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.**



La medida de BIC presenta una medida creciente en función al crecimiento del tamaño de muestra.

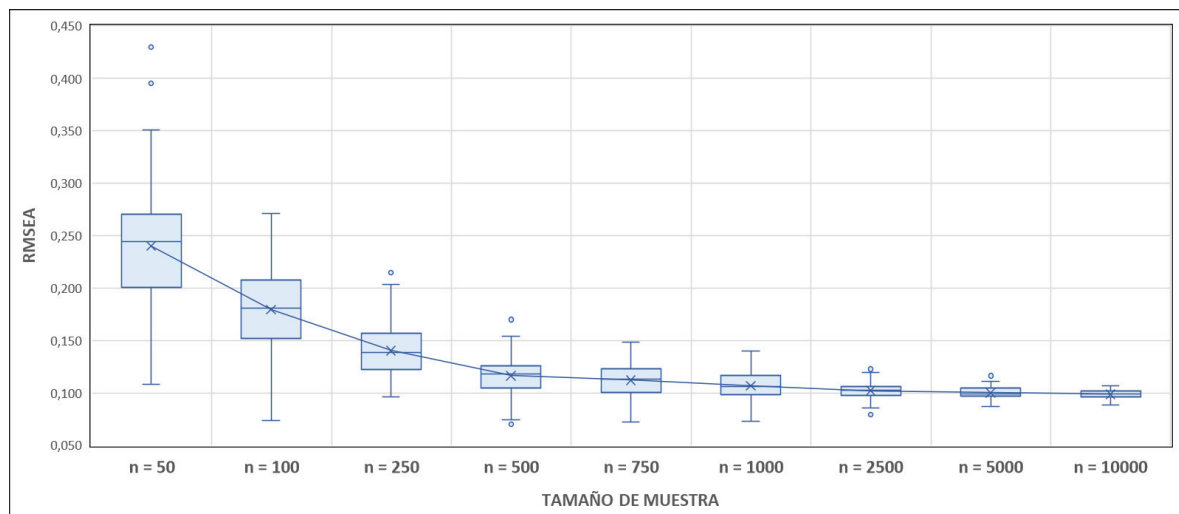
**Tabla 14**

*RMSEA del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018*

| Estadístico         | Tamaño de Muestra |       |       |       |       |       |       |       |       |
|---------------------|-------------------|-------|-------|-------|-------|-------|-------|-------|-------|
|                     | 50                | 100   | 250   | 500   | 750   | 1000  | 2500  | 5000  | 10000 |
| Promedio            | 0,240             | 0,180 | 0,140 | 0,116 | 0,112 | 0,107 | 0,102 | 0,100 | 0,099 |
| Desviación estándar | 0,059             | 0,039 | 0,025 | 0,018 | 0,017 | 0,013 | 0,008 | 0,006 | 0,004 |
| Mínimo              | 0,108             | 0,074 | 0,096 | 0,070 | 0,072 | 0,073 | 0,079 | 0,087 | 0,088 |
| Cuartil 1           | 0,201             | 0,153 | 0,122 | 0,105 | 0,100 | 0,098 | 0,097 | 0,097 | 0,096 |
| Mediana             | 0,244             | 0,181 | 0,139 | 0,118 | 0,113 | 0,106 | 0,102 | 0,099 | 0,099 |
| Cuartil 3           | 0,270             | 0,208 | 0,157 | 0,126 | 0,123 | 0,116 | 0,106 | 0,104 | 0,102 |
| Máximo              | 0,430             | 0,271 | 0,215 | 0,170 | 0,174 | 0,140 | 0,125 | 0,116 | 0,107 |

**Figura 14**

**RMSEA del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.**



La medida de RMSEA presenta una medida decreciente en función al crecimiento del tamaño de muestra, reduciéndose su intervalo de variación a medida que crece el tamaño de muestra.



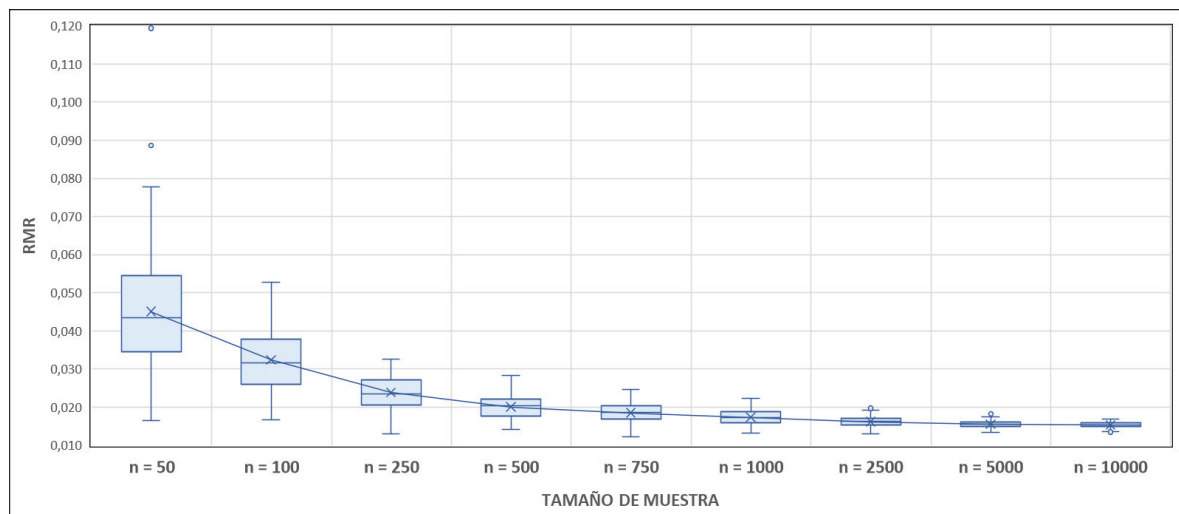
**Tabla 15**

*RMR del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018*

| Estadístico         | Tamaño de Muestra |       |       |       |       |       |       |       |       |
|---------------------|-------------------|-------|-------|-------|-------|-------|-------|-------|-------|
|                     | 50                | 100   | 250   | 500   | 750   | 1000  | 2500  | 5000  | 10000 |
| Promedio            | 0,045             | 0,032 | 0,024 | 0,020 | 0,019 | 0,017 | 0,016 | 0,016 | 0,015 |
| Desviación estándar | 0,016             | 0,008 | 0,004 | 0,003 | 0,003 | 0,002 | 0,001 | 0,001 | 0,001 |
| Mínimo              | 0,017             | 0,017 | 0,013 | 0,014 | 0,012 | 0,013 | 0,013 | 0,013 | 0,014 |
| Cuartil 1           | 0,035             | 0,026 | 0,021 | 0,018 | 0,017 | 0,016 | 0,015 | 0,015 | 0,015 |
| Mediana             | 0,043             | 0,032 | 0,024 | 0,020 | 0,019 | 0,017 | 0,016 | 0,016 | 0,015 |
| Cuartil 3           | 0,054             | 0,038 | 0,027 | 0,022 | 0,020 | 0,019 | 0,017 | 0,016 | 0,016 |
| Máximo              | 0,119             | 0,053 | 0,033 | 0,028 | 0,025 | 0,022 | 0,020 | 0,019 | 0,017 |

**Figura 15**

**RMR del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.**



La medida de RMR presenta una medida decreciente en función al crecimiento del tamaño de muestra, reduciéndose su intervalo de variación a medida que crece el tamaño de muestra.

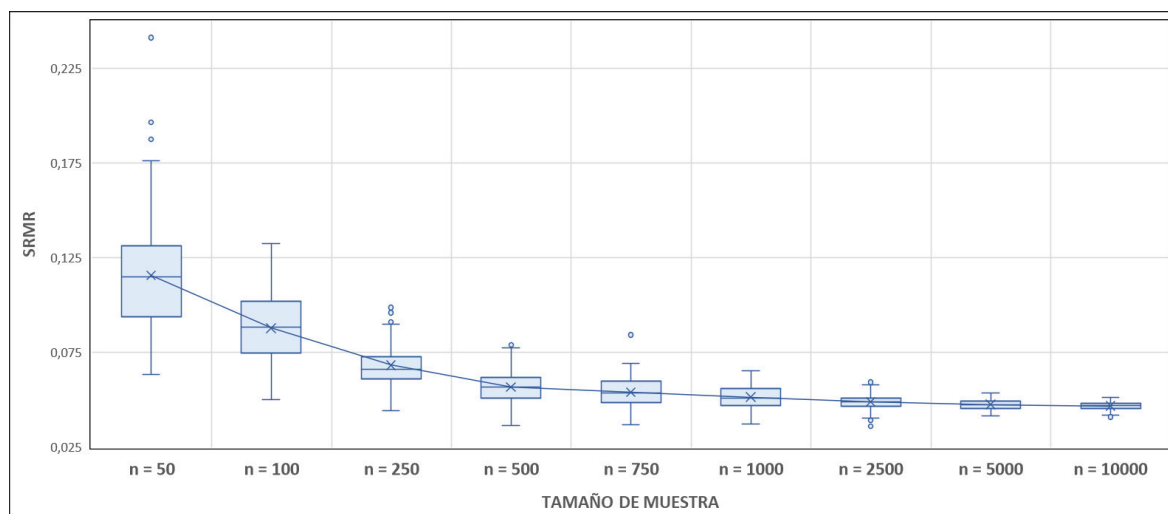
**Tabla 16**

*SRMR del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018*

| Estadístico         | Tamaño de Muestra |       |       |       |       |       |       |       |       |
|---------------------|-------------------|-------|-------|-------|-------|-------|-------|-------|-------|
|                     | 50                | 100   | 250   | 500   | 750   | 1000  | 2500  | 5000  | 10000 |
| Promedio            | 0,116             | 0,088 | 0,068 | 0,057 | 0,054 | 0,052 | 0,049 | 0,048 | 0,047 |
| Desviación estándar | 0,030             | 0,018 | 0,011 | 0,009 | 0,008 | 0,006 | 0,004 | 0,003 | 0,002 |
| Mínimo              | 0,064             | 0,050 | 0,044 | 0,036 | 0,037 | 0,037 | 0,036 | 0,042 | 0,041 |
| Cuartil 1           | 0,094             | 0,075 | 0,061 | 0,051 | 0,049 | 0,047 | 0,047 | 0,046 | 0,046 |
| Mediana             | 0,115             | 0,088 | 0,066 | 0,057 | 0,054 | 0,051 | 0,049 | 0,048 | 0,047 |
| Cuartil 3           | 0,131             | 0,102 | 0,072 | 0,062 | 0,060 | 0,056 | 0,051 | 0,049 | 0,048 |
| Máximo              | 0,241             | 0,133 | 0,101 | 0,079 | 0,084 | 0,066 | 0,060 | 0,054 | 0,052 |

**Figura 16**

**SRMR del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.**



La medida de SRMR presenta una medida decreciente en función al crecimiento del tamaño de muestra, reduciéndose su intervalo de variación a medida que crece el tamaño de muestra.

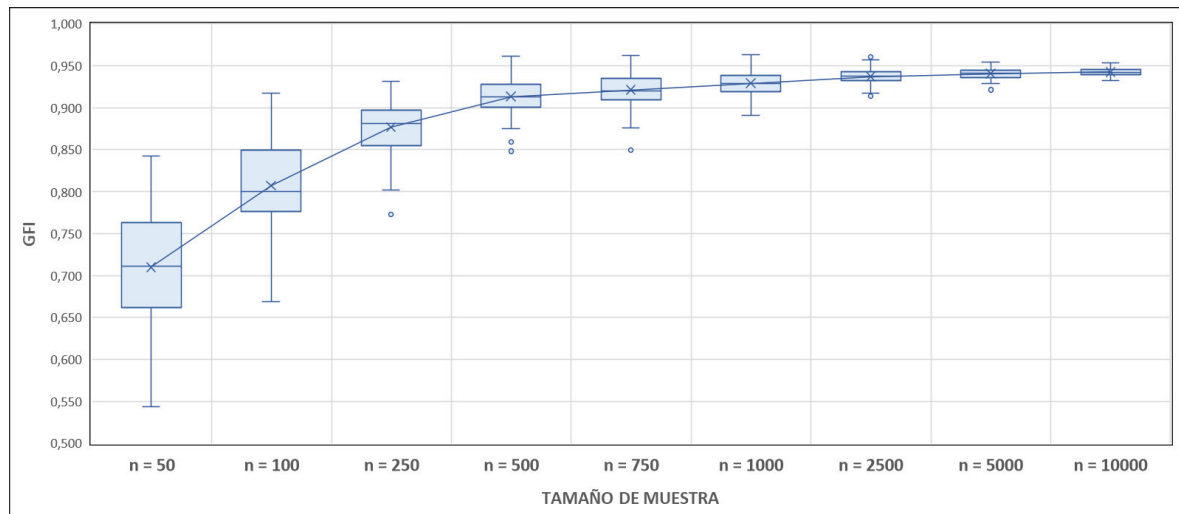
**Tabla 17**

*GFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018*

| Estadístico         | Tamaño de Muestra |       |       |       |       |       |       |       |       |
|---------------------|-------------------|-------|-------|-------|-------|-------|-------|-------|-------|
|                     | 50                | 100   | 250   | 500   | 750   | 1000  | 2500  | 5000  | 10000 |
| Promedio            | 0,710             | 0,807 | 0,877 | 0,913 | 0,921 | 0,929 | 0,937 | 0,941 | 0,943 |
| Desviación estándar | 0,073             | 0,049 | 0,030 | 0,020 | 0,019 | 0,015 | 0,009 | 0,006 | 0,004 |
| Mínimo              | 0,544             | 0,669 | 0,773 | 0,848 | 0,850 | 0,891 | 0,914 | 0,921 | 0,933 |
| Cuartil 1           | 0,664             | 0,777 | 0,855 | 0,901 | 0,910 | 0,919 | 0,933 | 0,936 | 0,940 |
| Mediana             | 0,711             | 0,800 | 0,881 | 0,913 | 0,920 | 0,929 | 0,937 | 0,941 | 0,942 |
| Cuartil 3           | 0,763             | 0,848 | 0,897 | 0,927 | 0,935 | 0,939 | 0,943 | 0,945 | 0,946 |
| Máximo              | 0,842             | 0,917 | 0,932 | 0,961 | 0,963 | 0,963 | 0,961 | 0,954 | 0,953 |

**Figura 17**

**GFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.**



La medida de GFI presenta una medida creciente en función al crecimiento del tamaño de muestra, reduciéndose su intervalo de variación a medida que crece el tamaño de muestra.

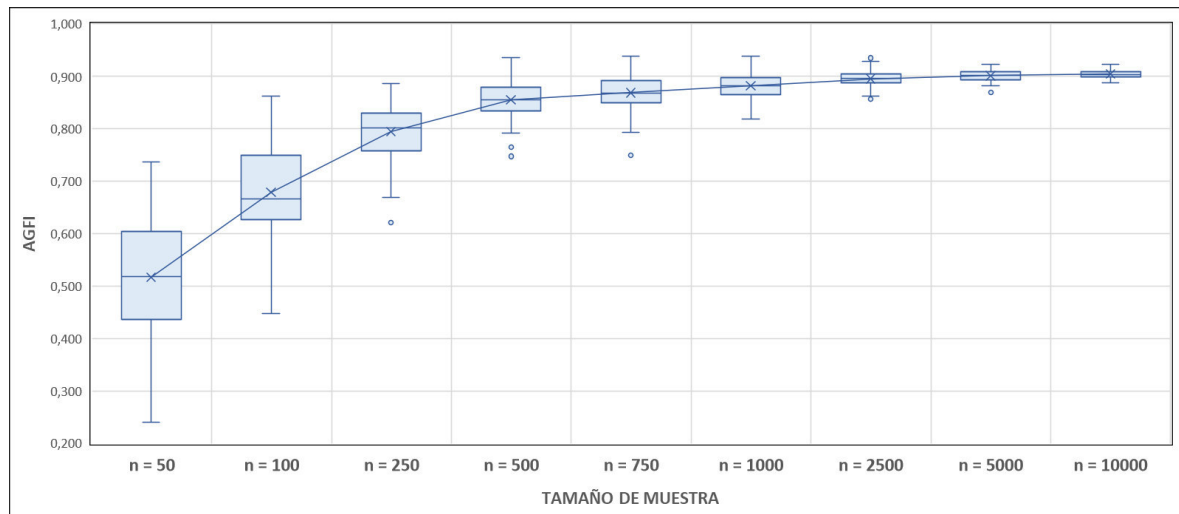
**Tabla 18**

*AGFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018*

| Estadístico         | Tamaño de Muestra |       |       |       |       |       |       |       |       |
|---------------------|-------------------|-------|-------|-------|-------|-------|-------|-------|-------|
|                     | 50                | 100   | 250   | 500   | 750   | 1000  | 2500  | 5000  | 10000 |
| Promedio            | 0,517             | 0,679 | 0,795 | 0,855 | 0,869 | 0,882 | 0,895 | 0,901 | 0,904 |
| Desviación estándar | 0,121             | 0,082 | 0,050 | 0,034 | 0,032 | 0,024 | 0,015 | 0,010 | 0,007 |
| Mínimo              | 0,240             | 0,448 | 0,621 | 0,747 | 0,749 | 0,819 | 0,857 | 0,869 | 0,888 |
| Cuartil 1           | 0,440             | 0,628 | 0,759 | 0,835 | 0,850 | 0,865 | 0,888 | 0,893 | 0,899 |
| Mediana             | 0,518             | 0,667 | 0,802 | 0,855 | 0,867 | 0,882 | 0,896 | 0,902 | 0,904 |
| Cuartil 3           | 0,605             | 0,746 | 0,829 | 0,879 | 0,892 | 0,898 | 0,905 | 0,908 | 0,909 |
| Máximo              | 0,737             | 0,862 | 0,886 | 0,935 | 0,938 | 0,939 | 0,935 | 0,924 | 0,922 |

**Figura 18**

**AGFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.**



La medida de AGFI presenta una medida creciente en función al crecimiento del tamaño de muestra, reduciéndose su intervalo de variación a medida que crece el tamaño de muestra.

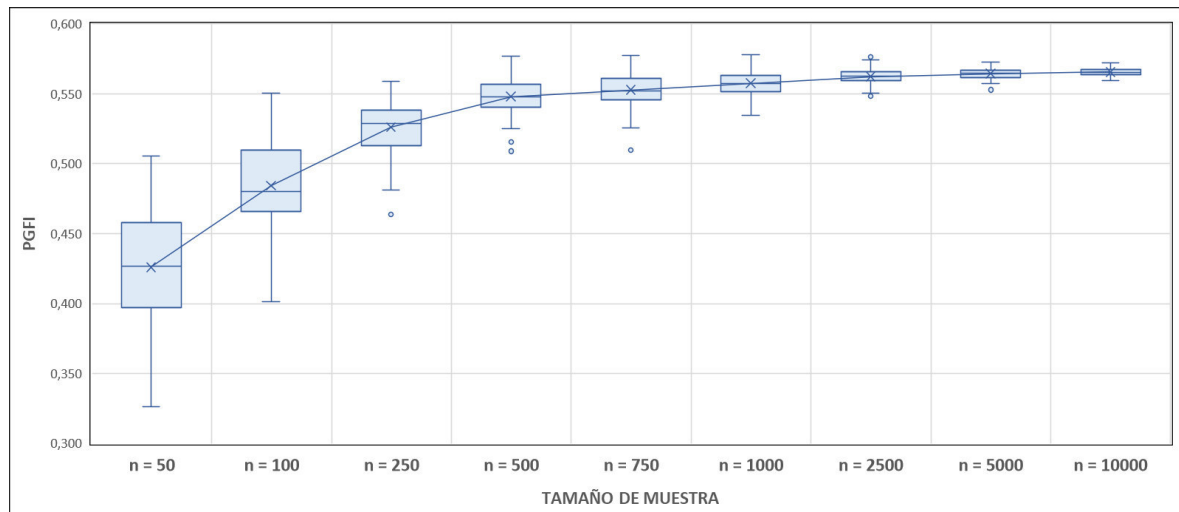
**Tabla 19**

*PGFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018*

| Estadístico         | Tamaño de Muestra |       |       |       |       |       |       |       |       |
|---------------------|-------------------|-------|-------|-------|-------|-------|-------|-------|-------|
|                     | 50                | 100   | 250   | 500   | 750   | 1000  | 2500  | 5000  | 10000 |
| Promedio            | 0,426             | 0,484 | 0,526 | 0,548 | 0,553 | 0,558 | 0,562 | 0,564 | 0,566 |
| Desviación estándar | 0,044             | 0,030 | 0,018 | 0,012 | 0,011 | 0,009 | 0,005 | 0,004 | 0,003 |
| Mínimo              | 0,326             | 0,401 | 0,464 | 0,509 | 0,510 | 0,535 | 0,549 | 0,553 | 0,560 |
| Cuartil 1           | 0,398             | 0,466 | 0,513 | 0,541 | 0,546 | 0,552 | 0,560 | 0,562 | 0,564 |
| Mediana             | 0,427             | 0,480 | 0,529 | 0,548 | 0,552 | 0,558 | 0,562 | 0,565 | 0,565 |
| Cuartil 3           | 0,458             | 0,509 | 0,538 | 0,556 | 0,561 | 0,563 | 0,566 | 0,567 | 0,567 |
| Máximo              | 0,505             | 0,550 | 0,559 | 0,577 | 0,578 | 0,578 | 0,577 | 0,572 | 0,572 |

**Figura 19**

**PGFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.**



La medida de PGFI presenta una medida creciente en función al crecimiento del tamaño de muestra, reduciéndose su intervalo de variación a medida que crece el tamaño de muestra.

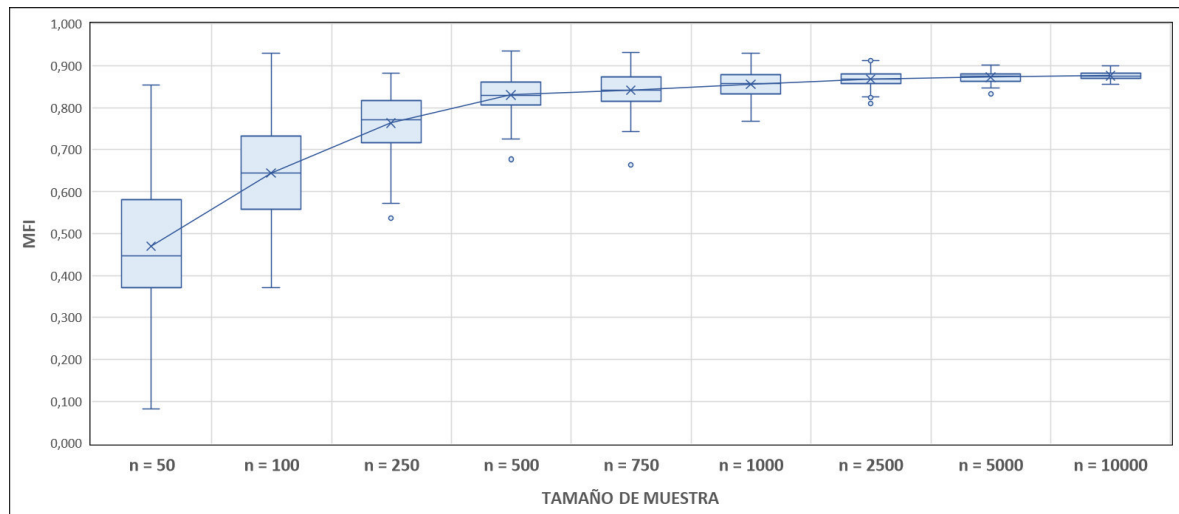
**Tabla 20**

*MFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018*

| Estadístico         | Tamaño de Muestra |       |       |       |       |       |       |       |       |
|---------------------|-------------------|-------|-------|-------|-------|-------|-------|-------|-------|
|                     | 50                | 100   | 250   | 500   | 750   | 1000  | 2500  | 5000  | 10000 |
| Promedio            | 0,470             | 0,645 | 0,764 | 0,831 | 0,842 | 0,856 | 0,868 | 0,873 | 0,876 |
| Desviación estándar | 0,163             | 0,118 | 0,073 | 0,045 | 0,043 | 0,032 | 0,020 | 0,014 | 0,009 |
| Mínimo              | 0,083             | 0,371 | 0,537 | 0,677 | 0,664 | 0,767 | 0,810 | 0,833 | 0,857 |
| Cuartil 1           | 0,373             | 0,558 | 0,717 | 0,807 | 0,815 | 0,833 | 0,858 | 0,863 | 0,870 |
| Mediana             | 0,447             | 0,644 | 0,772 | 0,829 | 0,841 | 0,858 | 0,868 | 0,875 | 0,876 |
| Cuartil 3           | 0,580             | 0,729 | 0,817 | 0,862 | 0,873 | 0,878 | 0,880 | 0,881 | 0,882 |
| Máximo              | 0,854             | 0,929 | 0,883 | 0,935 | 0,932 | 0,931 | 0,919 | 0,902 | 0,900 |

**Figura 20**

**MFI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.**



La medida de MFI presenta una medida creciente en función al crecimiento del tamaño de muestra, reduciéndose su intervalo de variación a medida que crece el tamaño de muestra.

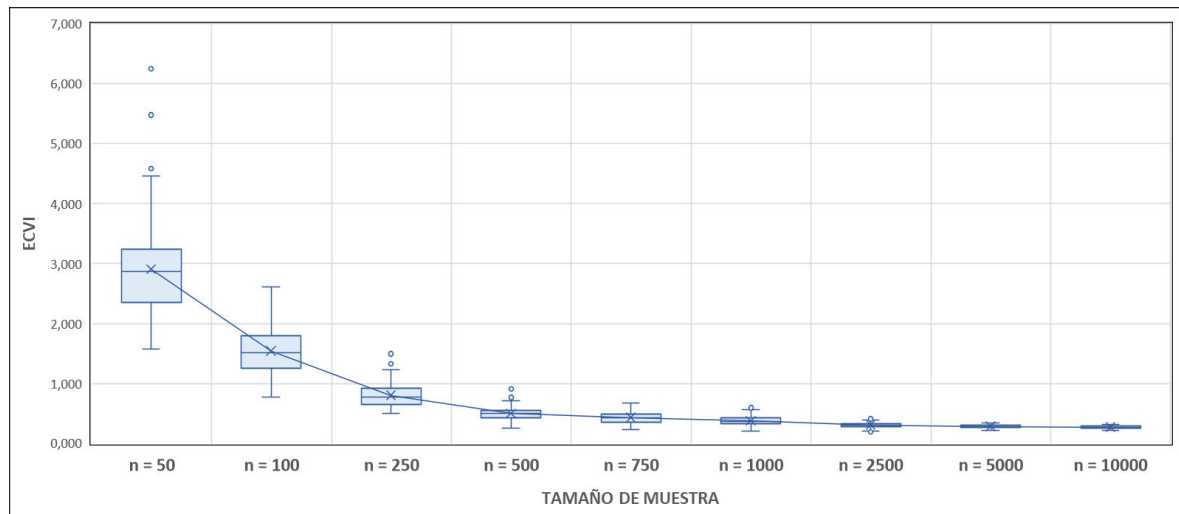
**Tabla 21**

*ECVI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018*

| Estadístico         | Tamaño de Muestra |       |       |       |       |       |       |       |       |
|---------------------|-------------------|-------|-------|-------|-------|-------|-------|-------|-------|
|                     | 50                | 100   | 250   | 500   | 750   | 1000  | 2500  | 5000  | 10000 |
| Promedio            | 2,909             | 1,542 | 0,801 | 0,500 | 0,432 | 0,375 | 0,309 | 0,284 | 0,270 |
| Desviación estándar | 0,805             | 0,372 | 0,200 | 0,111 | 0,104 | 0,076 | 0,046 | 0,031 | 0,021 |
| Mínimo              | 1,576             | 0,777 | 0,500 | 0,260 | 0,225 | 0,207 | 0,195 | 0,219 | 0,218 |
| Cuartil 1           | 2,350             | 1,262 | 0,655 | 0,424 | 0,356 | 0,323 | 0,281 | 0,266 | 0,256 |
| Mediana             | 2,870             | 1,510 | 0,770 | 0,502 | 0,430 | 0,368 | 0,308 | 0,279 | 0,272 |
| Cuartil 3           | 3,231             | 1,797 | 0,916 | 0,554 | 0,493 | 0,428 | 0,331 | 0,307 | 0,285 |
| Máximo              | 6,248             | 2,611 | 1,494 | 0,906 | 0,903 | 0,593 | 0,447 | 0,378 | 0,316 |

**Figura 21**

**ECVI del modelo factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, ENDES 2018.**



La medida de ECVI presenta una medida decreciente en función al crecimiento del tamaño de muestra, reduciéndose su intervalo de variación a medida que crece el tamaño de muestra.

## **4.2 Prueba de hipótesis**

El tamaño de muestra presenta un efecto directo en los índices de ajuste del análisis factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana, tal como se muestran en las figuras y tablas 2 al 21, donde se evidencia que el tamaño de muestra presenta un efecto directo con concavidad hacia arriba en el  $X^2$ ,  $X^2/df$ , AIC y BIC, efecto directo con concavidad hacia abajo en el CFI, TLI, NNFI, RFI, NFI, PNFI, IFI, RNI, GFI, AGFI, PGFI Y MFI, y en algunas medidas de bondad de ajuste un efecto indirecto en el RMSEA, RMR, SRMR y ECVI. Por lo que se contrasta un efecto directo e inverso del tamaño de muestra sobre las medidas de bondad de ajuste.

## **4.3 Discusión**

Las medidas de bondad de ajuste en su totalidad presentan efecto respecto al tamaño de muestra, tal como lo indicaron Cangur & Ercan (2015) y Ramírez et al., (1969).

Así mismo, en cuanto al GFI se observa un crecimiento cóncavo hacia abajo a medida que el tamaño de muestra crece lo cual coincide con los resultados obtenidos por Miles & Shevlin, (1998), donde concluyeron que el rendimiento del GFI mejora al aumentar el tamaño de la muestra.

Los valores del RMSEA son favorecidos con los tamaños de muestras grandes, tal como lo indicaron MacCallum et al., (1996), en la que recomendaron el uso del RMSEA en casos con muestras muy grandes. Del mismo modo, los resultados del RMR, SRMR y ECVI son favorecidos a medida que se cuenta con tamaños de muestras grandes.

Por otro lado, los resultados muestran que las medidas de bondad de ajuste ya sea en tamaños de muestras pequeños o grandes no significan plenamente un mal ajuste o buen ajuste del modelo estimado, tal como lo indica Iacobucci, (2010) en no tomar las reglas generales demasiado en serio, más por el contrario buscarle interpretabilidad y ser parsimonioso en el modelamiento de las variables de estudio.



## **CONCLUSIONES**

- El tamaño de muestra presenta efecto directo o indirecto según corresponda en los índices de ajuste del análisis factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana.
- El tamaño de muestra presenta efecto directo en los índices de ajuste absoluto del análisis factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana.
- El tamaño de muestra presenta efecto directo en los índices de ajuste parsimónico del análisis factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana.
- El tamaño de muestra presenta efecto directo en los índices de ajuste comparativo del análisis factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana.
- El tamaño de muestra presenta efecto directo en los índices de ajuste predictivo del análisis factorial confirmatorio sobre el cuestionario PHQ-9 aplicado a la población peruana.

## **RECOMENDACIONES**

- A los estudiantes, replicar el presente estudio para los diferentes tipos de métodos de estimación.
- A los investigadores, la realización del análisis factorial confirmatorio debería ser una práctica habitual como fundamento científico de validez del instrumento de medición que utilizan para sus investigaciones, a fin de dar el soporte científico a los datos y por ende a los resultados de emanan de las mismas.

## REFERENCIAS BIBLIOGRÁFICAS

- Borghero, F., Martínez, V., Zitko, P., & Vöhringer, P. A. (2018). *Tamizaje de episodio depresion en adolescentes Validacion PHQ -9*. 479–486.
- Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research* (T. G. Press (ed.); Second edi).
- Cangur, S., & Ercan, I. (2015). Comparison of model fit indices used in structural equation modeling under multivariate normality. *Journal of Modern Applied Statistical Methods*, *14*(1), 152–167. <https://doi.org/10.22237/jmasm/1430453580>
- Catena, A., Ramos, M. M., & Trujillo, H. M. (2003). Análisis multivariado: Un manual para investigadores. In *Editorial Biblioteca Nueva*.
- Denis, D. J. (2016). Applied Univariate, Bivariate, and Multivariate Statistics. In *Journal of Statistical Software* (Vol. 72, Issue Book Review 2). <https://doi.org/10.18637/jss.v072.b02>
- Fan, X., Thompson, B., & Wang, L. (1999). Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes. *Structural Equation Modeling*, *6*(1), 56–83. <https://doi.org/10.1080/10705519909540119>
- Gignac, G. E. (2006). Self-reported emotional intelligence and life satisfaction: Testing incremental predictive validity hypotheses via structural equation modeling (SEM) in a small sample. *Personality and Individual Differences*, *40*(8), 1569–1577.

<https://doi.org/https://doi.org/10.1016/j.paid.2006.01.001>

Gomer, B., Jiang, G., & Yuan, K. H. (2019). New Effect Size Measures for Structural Equation Modeling. *Structural Equation Modeling*, 26(3), 371–389.

<https://doi.org/10.1080/10705511.2018.1545231>

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1999). *Analysis multivariante (Hair).pdf*.

Harrington, D. (2009). Confirmatory Factor Analysis. In Oxford University Press (Ed.), *Journal of Chemical Information and Modeling* (Vol. 53, Issue 9).

Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural Equation Modelling : Guidelines for Determining Model Fit Structural equation modelling : guidelines for determining model fit. *Dublin Institute of Technology ARROW @ DIT*, 6(1), 53–60. [www.ejbrm.com](http://www.ejbrm.com)

Huba, G. J., & Harlow, L. L. (1987). Robust Structural Equation Models: Implications for Developmental Psychology. *Child Development*, 58(1), 147–166.

<https://doi.org/10.2307/1130297>

Iacobucci, D. (2010). Structural equations modeling: Fit Indices, sample size, and advanced topics. *Journal of Consumer Psychology*, 20(1), 90–98.

<https://doi.org/10.1016/J.JCPS.2009.09.003>

Igundunasse, A. (2016). The impact of small samples sizes in factor analytical studies. *Ife Psychologia*, 24(1), 54. <https://search.proquest.com/docview/1874040308>

Keith, T. Z. (2019). Multiple Regression and Beyond. In *Multiple Regression and Beyond*. <https://doi.org/10.4324/9781315749099>

Kline, R. B. (2016). *Principles and Practice of Structural Equation Modeling* (THE GUI LFORD PRESS (ed.); Fourth Edi).

- Koyuncu, İ., & Kılıç, A. F. (2019). The use of exploratory and confirmatory factor analyses: A document analysis. *Eğitim ve Bilim*, 44(198), 361–388.  
<https://doi.org/10.15390/EB.2019.7665>
- Lee, S.-Y., & Song, X.-Y. (2004). Evaluation of the Bayesian and Maximum Likelihood Approaches in Analyzing Structural Equation Models with Small Sample Sizes. In *Multivariate Behavioral Research* (Vol. 39, Issue 4, pp. 653–686). Lawrence Erlbaum. [https://doi.org/10.1207/s15327906mbr3904\\_4](https://doi.org/10.1207/s15327906mbr3904_4)
- 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>
- Miles, J. N., & Shevlin, M. (1998). Effects of sample size, model specification and factor loadings on the GFI in confirmatory factor analysis. *Personality and Individual Differences*, 25(1), 85–90.
- Mora-Esquivel, R., & Halabí, T. V. (2017). constructos latentes : Un método práctico Tomas Vargas Halabí Tamaño de la muestra en modelos de ecuaciones estructurales con constructos latentes : Un método práctico Sample sizes using structural equation modeling with latent variables : A practical met. *Revista Actualidades Investigativas En Educación*, 17(1), 1–34.  
<https://doi.org/http://dx.doi.org/10.15517/aie.v17i1.27294>
- Mulaik, S. A. (2010). Foundations of Factor Analysis. In *Chapman and Hall/CRC; 2 edition* (September 25, 2009). <https://doi.org/10.1201/b15851>
- Muthén, L. K., & Muthén, B. O. (2002). How to Use a Monte Carlo Study to Decide on Sample Size and Determine Power. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(4), 599–620.

[https://doi.org/10.1207/S15328007SEM0904\\_8](https://doi.org/10.1207/S15328007SEM0904_8)

Pituch, K., & Stevens, J. (2019). Applied Multivariate Statistics for the Social Sciences.

In *Sensory Evaluation of Sound*. <https://doi.org/10.1201/9780429429422-7>

Ramírez, I. C., Barrera, C., & Juan C. Correa, J. C. C. (1969). Efecto del tamaño de muestra y el número de réplicas bootstrap. *Ingeniería Y Competitividad*, 15(1), 93–101. <https://doi.org/10.25100/iyc.v15i1.2623>

Raykov, T., & Marcoulides, G. A. (2009). An Introduction to Applied Multivariate Analysis. In *International Statistical Review* (Vol. 77, Issue 1).

[https://doi.org/10.1111/j.1751-5823.2009.00074\\_18.x](https://doi.org/10.1111/j.1751-5823.2009.00074_18.x)

Rendón-Macías, M. E., & Villasís-Keever, M. Á. (2017). The research protocol V: The calculation of sample size. *Revista Alergia Mexico*, 64(2), 220–227.

<https://doi.org/10.29262/ram.v64i2.267>

Rožman, M., Tominc, P., & Milfelner, B. (2020). A Comparative Study Using Two SEM Techniques on Different Samples Sizes for Determining Factors of Older Employee's Motivation and Satisfaction. *Sustainability*, 12(6), 2189.

<https://doi.org/10.3390/su12062189>

Tanaka, J. S. (1987). “How Big Is Big Enough?”: Sample Size and Goodness of Fit in Structural Equation Models with Latent Variables. *Child Development*, 58(1), 134–146. <https://doi.org/10.2307/1130296>

Thoemmes, F., Rosseel, Y., & Textor, J. (2018). Local fit evaluation of structural equation models using graphical criteria. *Psychological Methods*, 23(1), 27–41.

<https://doi.org/10.1037/met0000147>

## ANEXOS

### CUESTIONARIO SOBRE LA SALUD DEL PACIENTE-9 (PHQ-9)

En las últimas 2 semanas, ¿con qué frecuencia ha sentido molestia por cualquiera de los siguientes problemas?

(Marque su respuesta con una "✓")

|  | Para nada | Varios días | Más de la mitad de los días | Casi todos los días |
|--|-----------|-------------|-----------------------------|---------------------|
| 1. Poco interés o placer en hacer las cosas  | 0         | 1           | 2                           | 3                   |
| 2. Sentirse desanimado/a, deprimido/a o sin esperanzas   | 0         | 1           | 2                           | 3                   |
| 3. Problemas para dormir o mantenerse dormido/a, o dormir demasiado  | 0         | 1           | 2                           | 3                   |
| 4. Sentirse cansado/a o con poca energía   | 0         | 1           | 2                           | 3                   |
| 5. Poco apetito o comer en exceso  | 0         | 1           | 2                           | 3                   |
| 6. Sentirse mal acerca de sí mismo/a – o sentir que es un/una fracasado/a o que se ha fallado a sí mismo/a o a su familia  | 0         | 1           | 2                           | 3                   |
| 7. Dificultad para concentrarse en las cosas, tales como leer el periódico o ver televisión  | 0         | 1           | 2                           | 3                   |
| 8. ¿Moverse o hablar tan despacio que otras personas lo pueden haber notarlo? O lo contrario – estar tan inquieto/a o intranquilo/a que se ha estado moviendo mucho más de lo normal | 0         | 1           | 2                           | 3                   |
| 9. Pensamientos de que sería mejor estar muerto/a o que quisiera lastimarse de alguna forma  | 0         | 1           | 2                           | 3                   |

FOR OFFICE CODING   0   + \_\_\_\_\_ + \_\_\_\_\_ + \_\_\_\_\_  
=Total Score: \_\_\_\_\_

Si marcó algún problema, ¿cuánto le han dificultado estos problemas realizar su trabajo, encargarse de las cosas en la casa, o llevarse bien con otras personas?

|  |  |   |  |
|--|--|---|--|
| No ha sido difícil<br><input type="checkbox"/> | Algo difícil<br><input type="checkbox"/> | Muy difícil<br><input type="checkbox"/> | Extremadamente difícil<br><input type="checkbox"/> |
|--|--|---|--|