

Uso de Variables Latentes para Identificar Trastornos del Sueño en Escolares y Adolescentes a partir de una Medida Sintética

Andrés Olave Quintero

Universidad del Valle
Facultad de Ingeniería, Programa Académico de Estadística
Santiago de Cali Colombia
2024 .



Uso de Variables Latentes para Identificar Trastornos del Sueño en Escolares y Adolescentes a partir de una Medida Sintética

Andrés Olave Quintero

Trabajo de grado presentado como requisito parcial para optar al título de:
Estadístico

Director: Ph.D. José Rafael Tovar Cuevas
Codirector: Ph.D. Carlos Alberto Velasco-Benítez

Universidad del Valle
Facultad de Ingeniería, Programa Académico de Estadística
Santiago de Cali, Colombia

Agradecimientos

Agradecerle Dios por guiarme siempre y en todo lugar, a mi madre Lucy, aquella mujer que me inculcó desde un inicio mi gusto por aprender sobre los pequeños detalles de este universo, sin ella mi visión sobre la vida no sería igual. También agradezco a mi abuela Francia, por la atención, amor y cuidado que siempre me ha brindado, le doy gracias a mi padre Robinson que a pesar de nuestros malos entendidos, ha mostrado interés porque yo pueda superarme académicamente. Doy gracias también al destino, tiempo, espacio o como se le quiera denominar, por ponerme en frente a personas increibles que me alientan a cumplir todos mis sueños. Le doy gracias además a mis mascotas; uno desde el cielo y las otras dos presentes, su amor incondicional va más allá de lo que un humano pueda entender.

Agradezco tanto a mi director de tesis, Rafael Tovar, como a los profesores que a lo largo de la carrera me han enseñado el potencial de esta ciencia. En especial, agradezco las enseñanzas brindadas por Roberto Behar, Álvaro Flórez, Jaime Mosquera y Andrés Ochoa, que de una u otra forma fueron vitales en el desarrollo de este trabajo.

Sin mis compañeros y amigos tampoco hubiesen sido posible muchos de mis logros, por ende, le debo un agradecimiento a Johan Sebastián, Daniela, Marcela, Julián, Juan José, Diana, Andrés y Brandon, personas que me dieron en muchas oportunidades el empujón que necesitaba.

Por último, agradezco a mi codirector Carlos Velasco porque gracias a sus esfuerzos, pude comprender un poco más sobre los constructos y cómo estos interactuan en ciencias como la medicina. Muchas gracias por realizar este tipo de trabajos en pro de la infancia.

Resumen

Estudios como los de Rabner et al (2017), Yuan et al (2023) han comprobado que los trastornos del sueño en niños guardan una estrecha relación con algunos síntomas como dolores de cabeza crónicos o problemas de salud mental, como la depresión y ansiedad. A pesar del creciente número de investigaciones sobre el tema, son muy pocos los artículos que evalúan el trastorno desde una visión de variable latente y en qué forma los constructos subyacentes aportan a la formación del trastorno del sueño como medida única.

El objetivo de este trabajo fue diseñar una medida sintética que permitiera conocer el nivel de presencia de trastornos del sueño en escolares y adolescentes. Para ello, se tomó como referencia una muestra de 428 estudiantes de un colegio ubicado en Santiago de Cali con edades que van desde los 9 a los 12 años (edad escolar) y de los 13 hasta los 18 años (edad adolescente), quienes respondieron a 7 instrumentos: PISI, ESS, PROMIS SD, PROMIS SRI, ASWS, PROMIS ANXIETY y PROMIS DEPRESSION. Adicionalmente, los padres participaron respondiendo al instrumento SHIP. Partiendo de lo anterior, se plantearon 3 alternativas para medida sintética. Para la primer alternativa se realizó un análisis factorial múltiple (AFM). En la segunda alternativa se realizó un análisis de componentes principales (ACP) secuencial. La tercera alternativa consistió en un análisis de componentes principales (ACP) aplicado a los puntajes globales de cada instrumento. Posteriormente se ponderaron los valores de los vectores propios resultantes del análisis y se ajustaron al modelo reflectivo de variables latentes como pesos asociados a los ítems. De manera agregada, se evaluaron las dos primeras alternativas transformando las puntuaciones de los ítems en puntuaciones monótonas continuas. Posteriormente se construyeron las escalas de las medidas y se evaluó el Pseudo Error Cuadrático Medio como medida de desempeño. Dados los resultados, se estableció que la medida denominada ITSNE2 (segunda alternativa con puntuaciones monótonas continuas), fue la mejor, alcanzando con un tamaño de muestra igual a 30 un Pseudo Error Cuadrático Medio de 10,79, inferior a los de las demás escalas.

Finalmente se establecieron los puntos de corte para determinar los grados de severidad del trastorno del sueño presente en el escolar y adolescente, definiendo como leves, puntuaciones T_i tal que $0 \leq T_i \leq 19{,}40$, como moderados, $19{,}40 \leq T_i \leq 44{,}30$ y severos, $44{,}30 \leq T_i \leq 100$. Se concluyó que existe una asociación significativa entre el sexo y el grado de severidad ($\chi^2 = 14{,}17$, $valor - p = 0{,}0008$). Además, que el escolar o adolescente presente por lo menos un desorden gastrointestinal funcional está asociado con el grado de severidad ($\chi^2 = 8{,}51$, $valor - p = 0{,}0141$).

(Palabras Clave): Trastornos del Sueño, Modelo Reflectivo, Variable Latente, Análisis Factorial Múltiple, Análisis de Componentes Principales, Pseudo Error Cuadrático Medio, Escolares, Adolescentes.

Abstract

Studies such as Rabner et al (2017) and Yuan et al (2023) have shown that sleep disorders in children are closely related to symptoms such as chronic headaches or mental health problems such as depression and anxiety. Despite the growing number of research on the topic, very few articles assess the disorder from a latent variable perspective and how underlying constructs contribute to the formation of sleep disorders as a single measure.

The objective of this work was to design a synthetic measure that would allow us to know the level of presence of sleep disorders in schoolchildren and adolescents. To do this, a sample of 428 students from a school located in Santiago de Cali was taken as a reference, with ages ranging from 9 to 12 years (school age) and from 13 to 18 years (adolescent age), who responded to 7 instruments: PISI, ESS, PROMIS SD, PROMIS SRI, ASWS, PROMIS ANXIETY and PROMIS DEPRESSION. Additionally, parents participated by responding to the SHIP instrument. Based on the above, 3 alternatives for synthetic measurement were proposed. For the first alternative, a multiple factor analysis (MFA) was performed. In the second alternative, a sequential principal component analysis (PCA) was performed. The third alternative consisted of a principal component analysis (PCA) applied to the global scores of each instrument. The values of the eigenvectors resulting from the analysis were then weighted and adjusted to the reflective model of latent variables as weights associated with the items. In an aggregated manner, the first two alternatives were evaluated by transforming the item scores into continuous monotonic scores. The scales of the measures were then constructed and the Pseudo Mean Square Error was evaluated as a performance measure. Given the results, it was established that the measure called ITSNE2 (second alternative with continuous monotonic scores) was the best, reaching a Pseudo Mean Square Error of 10,79 with a sample size equal to 30, lower than those of the other scales.

Finally, the cut-off points were established to determine the degrees of severity of the sleep disorder present in schoolchildren and adolescents, defining as mild, scores T_i such that $0 \leq T_i \leq 19{,}40$, as moderate, $19{,}40 \leq T_i \leq 44{,}30$ and severe, $44{,}30 \leq T_i \leq 100$. It was concluded that there is a significant association between sex and the degree of severity ($\chi^2 = 14{,}17$, $p - value = 0{,}0008$). In addition, that the schoolchild or adolescent presents at least one functional gastrointestinal disorder is associated with the degree of severity ($\chi^2 = 8{,}51$, $p - value = 0{,}0141$).

(Keywords): Sleep Disorders, Reflective Model, Latent Variable, Multiple Factor Analysis, Principal Component Analysis, Pseudo Root Mean Square Error, School Children, Adolescents.

Contenido

1. Introducción **12**

2. Definición del problema **13**
2.1. Planteamiento del problema 13
2.2. Justificación 14
2.3. Pregunta de Investigación 15
2.4. Objetivos 15
2.4.1. Objetivo General 15
2.4.2. Objetivos Específicos 15

3. Revisión de Antecedentes **16**
3.1. Antecedentes de Diseño y Validación de Escalas para Variables Latentes 16
3.2. Antecedentes de Evaluación del Desempeño de Escalas para Variables Latentes 17

4. Marco Teórico **18**
4.1. Marco Teórico Conceptual-Clínico 18
4.1.1. Trastorno 18
4.1.2. Insomnio 18
4.1.3. Somnolencia 18
4.1.4. Vigilia 19
4.1.5. Ansiedad y Depresión 19
4.1.6. Baremos (puntos de corte) 19
4.2. Marco Teórico Estadístico 19
4.2.1. Variable Latente 19
4.2.2. Modelo Reflectivo de Variables Latentes 21
4.2.3. Teoría de la Medición en Psicología 21
4.2.4. Likert Equidistante como Suma Ponderada de Categorías de Respuesta 23
4.2.5. Análisis Factorial Múltiple (AFM) 24
4.2.6. Pseudo Error Cuadrático Medio 25
4.2.7. Algoritmo K-means 27
4.2.8. Prueba Chi-Cuadrado de asociación 2 x N 27
4.2.9. Prueba de bondad de ajuste Anderson-Darling 28

5. Metodología **29**
5.1. Análisis Bibliométrico 29
5.2. Descripción del Archivo de Datos y la Fuente de Información . 30
5.3. Pre-procesamiento de datos 31
5.4. Construcción de las Tablas para los respectivos análisis 32
5.4.1. Comparación en el Análisis, Teoría de la Medición y Likert Equidistante como Suma Ponderada de Categorías de Respuesta 34

5.5. Análisis Exploratorio de los Datos 35
5.6. Análisis Multivariado 35
5.6.1. Construcción de la medida sintética para escolares-adolescentes (alternativas I, II y III) 35
5.6.2. Construcción de la medida sintética para padres 38
5.7. Creación de las escalas 38
5.8. Modelo Estadístico 39
5.9. Correlación entre medidas sintéticas y puntajes validados . . . 40
5.10. Evaluación del Desempeño de las Medidas Sintéticas 40
5.11. Clasificación de los individuos de la muestra según sus niveles de trastorno del sueño 41
5.12. Construcción de tablas de contingencia para evaluar la dependencia entre los niveles de trastorno del sueño y características de los individuos 42

6. Resultados **43**
6.1. Fase preliminar: Análisis Exploratorio de los datos 43
6.2. Resultados del Primer Objetivo: Construcción de la Medida Sintética para escolares-adolescentes (alternativas I,II y III) . . 46
6.2.1. Construcción de la medida sintética para escolares-adolescentes por medio de AFM 46
6.2.2. Construcción de la medida sintética para escolares-adolescentes por medio ACP 48
6.2.3. Construcción de la medida sintética para escolares-adolescentes con base en los puntajes totales de las pruebas, haciendo uso de ACP 49
6.2.4. Construcción de la medida sintética para escolares-adolescentes por medio de AFM y Likert Equidistante . 50
6.2.5. Construcción de la medida sintética para escolares-adolescentes por medio ACP y Likert Equidistante . . . 52
6.2.6. Construcción de la medida sintética para padres por medio de ACP 53
6.2.7. Correlación entre las medidas sintéticas y la medida de los puntajes validados 55
6.2.8. Estadísticas Resumen de las Alternativas a Escala del Trastorno del Sueño en Escolares-Adolescentes 55
6.3. Resultados del Segundo Objetivo: Evaluación del desempeño de la medida sintética (PECM) 57
6.4. Resultados del Tercer Objetivo: Descripción de las características asociadas a los trastornos del sueño en el grupo de individuos 59

7. Conclusiones y Recomendaciones **63**
7.1. Conclusiones 63
7.2. Recomendaciones 64

8. Anexos **65**

8.1. Código en R 65
8.2. Nubes de individuos y tipologías de grupos 65
8.3. Clasificaciones según puntaje total en los instrumentos 68
8.4. Resultados pruebas Anderson-Darling 68

Lista de Figuras

1. *Representación de los Constructos y SEM teórico (Gráfico de Trayectorias)* 20
2. *Metodología Aplicada en la Construcción de Medidas Sintéticas para Identificar Transtornos del Sueño en Escolares y Adolescentes* 29
3. *Análisis por Autor en VOSviewer* 30
4. *Análisis por palabras clave en VOSviewer* 30
5. *Proceso para aproximar el valor del Pseudo Error Cuadrático Medio de los estimadores para variables latentes* 40
6. Distribución de los puntajes reescalados según la medida sintética 57
7. Distribución del estimador del pseudoparámetro asociado al constructo (trastorno del sueño mediano) 58
8. *Criterio de selección del número de agrupaciones por grado de severidad según el puntaje obtenidos en la escala de la medida ITSNE2* 59
9. *Boxplot del puntaje reescaldo de los individuos clasificados por grado de severidad del trastorno de sueño* 60
10. *Similitudes entre los trastornos medidos por los instrumento para la alternativa I (AFM)* 65
11. *Similitudes entre los trastornos medidos por los instrumento para la alternativa I (AFM) con puntajes transformados* 65
12. *Nube de individuos para la alternativa I (AFM)* 66
13. *Nube de individuos para la alternativa II (ACP)* 66
14. *Nube de individuos para la alternativa III (puntajes totales y ACP)* 66
15. *Nube de individuos para la alternativa I (AFM) con puntajes transformados* 67
16. *Nube de individuos para la alternativa II (ACP) con puntajes transformados* 67
17. *Nube de individuos para la medida de padres* 67

Índice de cuadros

1. *Descripción de las variables sociodemográficas y de salud* . . . 32
2. *Descripción de los ítems de cada instrumento evaluado para medir un tipo trastorno relacionado con el sueño en escolares y adolescentes.* . . . 33
3. *Descripción de los puntajes totales de cada instrumento evaluado para medir un tipo trastorno relacionado con el sueño en escolares y adolescentes.* . . . 34
4. *Descripción de los ítems de los instrumentos aplicados a los padres donde se evalua la higiene del sueño evidenciadas en los escolares y adolescentes* . . . 34
5. *Frecuencia de los participantes según variables sociodemográficas y de salud* . . . 43
6. *Distribución de las respuestas a los ítem de cada uno de los instrumentos evaluados en los escolares y adolescentes* . . . 45
7. *Resumen de los Valores Propios Resultante del AFM para la alternativa I* . . . 46
8. *Descomposición de la Inercia Total del AFM para la alternativa I* 47
9. *Coordenadas y Contribuciones para las dos Primeras Dimensiones del AFM para la alternativa I* . . . 47
10. *Resumen Valores Propios Resultantes del ACP para la alternativa II* . . . 48
11. *Coordenadas y Contribuciones para las dos Primeras Dimensiones del ACP para la alternativa II* . . . 48
12. *Resumen Valores Propios Resultantes del ACP para la alternativa III* . . . 49
13. *Coordenadas y Contribuciones para las dos Primeras Dimensiones del ACP para la alternativa III* . . . 49
14. *Resumen de los Valores Propios Resultante del AFM para la alternativa I con datos transformados (Likert Equidistante)* . . . 50
15. *Descomposición de la Inercia Total del AFM para la alternativa I con datos transformados (Likert Equidistante)* . . . 51
16. *Coordenadas y Contribuciones para las dos Primeras Dimensiones del AFM para la alternativa I con datos transformados (Likert Equidistante)* . . . 51
17. *Resumen Valores Propios Resultantes del ACP para la alternativa II con datos transformados (Likert Equidistante)* . . . 52
18. *Coordenadas y Contribuciones para las dos Primeras Dimensiones del ACP para la alternativa II con datos transformados (Likert Equidistante)* . . . 53
19. *Resumen Valores Propios Resultantes del ACP de la medida sintética para padres* . . . 54
20. *Coordenadas y Contribuciones para las dos Primeras Dimensiones del ACP de la medida para padres* . . . 54

21. *Relación lineal entre las escalas de las medidas sintéticas y el puntaje medio de los resultados de todas las pruebas* 55
22. *Resumen de los Rangos Teóricos y Observados de las Medidas Sintéticas* 56
23. *Resumen Estadístico de las Escalas* 56
24. *Resultados del desempeño según el PECM para las medidas sintéticas diseñadas* 58
25. *Valores de límite para cada grado de severidad* 60
26. *Relación entre los grados de severidad en los participantes y variables sociodemográficas y de salud* ($\alpha = 0{,}05$) 61
27. *Grado de severidad según el puntaje para los instrumentos aplicados en los escolares y adolescentes* 68
28. *Resultados de la prueba formal Anderson-Darling aplicada a los dos conjuntos de datos para los cuales se analizó la simetría por medio del histograma* 68

1 Introducción

Este documento es de carácter investigativo y tiene como propósito diseñar una medida sintética, que le permita a los profesionales de la salud identificar el nivel de trastorno del sueño que un niño en edad escolar o adolescente presente. Según Gállego Pérez-Larraya et al (2007) los trastornos del sueño conforman un grupo extenso de procesos heterogénos. Muchas enfermedades son consideradas síntomas de algún trastorno del sueño, la forma de clasificar estos trastornos se basa en su síntoma principal y está estipulado por convención, según la Clasificación Internacional del Trastornos del Sueño (ICSD).

Adicionalmente, en este trabajo se busca indagar sobre los factores que están mayormente relacionados con niveles altos de presencia de trastorno del sueño en los niños, ya que reconocerlos puede ser de gran ayuda para fortalecer los instrumentos con los cuales se diagnóstican este tipo de cambios desadaptativos. Además de comprender los procesos subyacentes para incentivar estrategías a nivel social, que permitan el correcto desarrollo en esta población.

Para el desarrollo de la medida sintética se cuenta con 428 registros con las respuestas de los niños en edad escolar y adolescente a 7 pruebas psicométricas que miden por medio de un puntaje, la presencia de diferentes clasificaciones de trastornos del sueño. Adicionalmente, se considera la información que los padres puedan dar sobre la higiene del sueño del niño, medida a través de una prueba psicométrica.

De acuerdo con lo anterior y dado el propósito del trabajo, se hará uso de herramientas estadísticas aplicadas en la psicología y ciencias sociales, tales y como, los modelos de variables latentes, el análisis multivariado, medidas de desempeño de estimadores, los clúster y pruebas no paramétricas, con la finalidad de no solo diseñar la medida sintética, sino también estudiar su comportamiento y poder aportar con conclusiones a nivel general.

2 Definición del problema

2.1 Planteamiento del problema

Los trastornos del sueño son la segunda causa más común de dolores de cabeza en niños y adolescentes, según Rabner et al (2017), estos dos estados parecen estar relacionadas de manera bidireccional. Además, la combinación de dolores de cabeza crónicos y algún trastorno del sueño, incrementa los niveles de ansiedad y depresión en los niños y adolescentes. En un estudio realizado en Singapur, Mahendran et al (2006) encontró que de los niños que formaron parte de la investigación, alrededor del 62 % presentó por lo menos un problema del sueño, por otro lado, los problemas psiquiátricos, el género y los antecedentes familiares, predisponen la presencia de algún trastorno.

En Colombia, aunque los estudios sobre problemas relacionados con el sueño son limitados, investigaciones como la de Vélez et al (2018) revelan que los trastornos del sueño y otros problemas asociados presentan una prevalencia superior al 20 % en niños y adolescentes de entre 9 y 17 años. Entre las dificultades más destacadas se encuentran problemas para conciliar el sueño, somnolencia diurna excesiva, despertares nocturnos, así como irregularidades en la duración y regularidad del sueño.

A día de hoy, existen diferentes instrumentos que, por medio de ciertos comportamientos característicos que presentan los niños y adolescentes, buscan explicar la presencia de algún trastorno del sueño. Debido a que muchas veces no es sencillo validar que la experiencia real del niño durante el sueño se vea reflejada por su respuesta, se han implementado instrumentos aplicables a las figuras paternas, por ende, se puede tener un contraste para observar la concordancia entre lo informado por el niño y el padre. Sin embargo, hay muchos cuestionarios y cada uno puede abordar dimensiones del sueño distintas, además, la mayoría de variables no se pueden medir de una forma directa, a diferencia de como se haría en pruebas clínicas con el peso u oxígeno en sangre, por lo tanto, se emplean múltiples escalas; en ese mismo sentido, se crea la necesidad de desarrollar una metodología que permita conocer si existe presencia de trastornos del sueño en el niño o adolescente, mediante el uso de una medida que sintetice los diferentes factores de información.

2.2 Justificación

A nivel mundial cada vez toma más importancia los problemas de salud mental, como la depresión y la ansiedad, según Yuan et al (2023), la mala calidad del sueño se asoció significativamente con 7,27 veces mayores probabilidades de depresión y 2,54 veces mayores probabilidades de ansiedad social entre niños y adolescentes. Además, los adolescentes tienen mayores probabilidades de experimentar eventos de depresión y ansiedad que los escolares.

Los trastornos del sueño no solamente representan posibles problemas mentales, sino que también, favorecen la presencia de migraña, dolores abdominales, dolores músculo-esqueléticos y trastornos reumatológicos, tal como lo plantea Oh et al (2023). Esta serie de trastornos psicológicos y dolores crónicos relacionados con el sueño, constituyen una problemática que debe ser abordada de manera oportuna para el beneficio social y de los expertos en el campo clínico y psicológico, de modo que, estos puedan generar decisiones que mejoren la calidad de vida de los niños y adolescentes.

El número de investigaciones y documentos realizados en torno a los trastornos del sueño han aumentado rápidamente desde los años noventa, sin embargo, la cantidad de documentos que tienen en cuenta aspectos, tales como, la medición, representan una proporción muy baja y la mayoría de estos, no evalúan los constructos subyacentes que generan el constructo principal denominado como trastorno del sueño en los niños o adolescentes. A nivel científico y económico, los investigadores de la salud y de las ciencias sociales pueden verse beneficiados si hay una alternativa más eficiente a realizarle diferentes tipos de cuestionarios a los niños y adolescentes, ya que esto supone altos costos de implementación y puede que las preguntas no aporten significativamente al diagnóstico de algún trastorno del sueño. Por lo mencionado anteriormente, se hace indispensable realizar la investigación, aplicando de manera rigurosa los conceptos estadísticos, con la finalidad de generar conocimiento a nivel pediátrico y científico.

2.3 Pregunta de Investigación

¿Cómo construir una medida sintética para evaluar el nivel de presencia de trastornos del sueño en escolares y adolescentes, partiendo del modelo reflectivo de variables latentes?

2.4 Objetivos

2.4.1 Objetivo General

Diseñar una medida sintética, que permita conocer el nivel de presencia de trastornos del sueño en escolares y adolescentes.

2.4.2 Objetivos Específicos

- Construir una medida sintética que considere todas las dimensiones de las pruebas validadas en el área de la salud para evaluar trastornos del sueño, usando los conceptos de la teoría de las variables latentes y del análisis factorial.
- Evaluar el desempeño estadístico de la medida sintética creada, utilizando conceptos teóricos de la estimación abordados por la Inferencia Estadística.
- Describir las características asociadas a los trastornos del sueño en un grupo de escolares y adolescentes.

3 Revisión de Antecedentes

Este capítulo está enfocado en presentar y discutir trabajos previos, cuyos objetivos fueron construidos en torno al desarrollado de escalas para la medición de variables latentes o constructos; mediante el uso de técnicas multivariadas como el ACP, CFA, EFA y algunos indicadores de desempeño, como el Alpha de Cronbach, el Porcentaje de Varianza Explicada y el Pseudo Error Cuadrático Medio. Esta revisión bibliográfica fue delimitada a los últimos 5 años a excepción de un artículo del año 2018, el cual se consideró como necesaria su revisión.

3.1 Antecedentes de Diseño y Validación de Escalas para Variables Latentes

En el trabajo de Moo-Estrella et al (2018), diseñaron y validaron una escala de autoinforme para trastornos del sueño en escolares, el estudio se realizó con niños participantes de 6 escuelas primarias públicas en Yucatán, México, seleccionados mediante muestreo de conveniencia. El total de participantes se dividió en dos grupos: en el primero, se desarrolló la fase inicial para el desarrollo de la escala, mientras que, en el segundo se les aplicó la versión final de la escala, para la obtención de la escala se implementó una fase de diseño y prueba piloto, luego, otra fase referente a generar la estructura factorial de la escala y análizar la prevalencia de trastornos del sueño en la muestra. El análisis factorial de Componentes principales o ACP, conjunto al Alpha de Cronbach permitieron conocer la estructura factorial y la consistencia interna de la escala, respectivamente. El análisis estadístico mostró que con 6 componentes principales se explicaba una varianza total del 47.8 %, los seis componentes y sus Alpha de Cronbach correspondientes fueron: dificultades de inicio de sueño ($\alpha = 0{,}77$), pesadillas ($\alpha = 0{,}76$), despertares nocturnos ($\alpha = 0{,}71$), somnolencia diurna ($\alpha = 0{,}69$), cansancio y dificultades para despertar ($\alpha = 0{,}69$), en último lugar, el sonambulismo ($\alpha = 0{,}69$). Finalmente lograron diseñar la escala de alteraciones del sueño para niños en edad escolar, cuyas propiedades psicométricas son apropiadas para la detección de alteraciones del sueño en los niños mexicanos hispano-hablantes.

En el estudio de Randler et al (2019), proporcionaron un instrumento confiable para la medición de somnolencia excesiva para próximos estudios en niños de habla Rusa, aplicando la Escala de Somnolencia Diurna Pediátrica (PDSS), participaron en el estudio, escuelas de seis asentamientos distintos del noroeste de Rusia en el periodo de marzo a octubre del 2018, se seleccionaron 552 estudiantes mediante un muestreo sistemático, de los cuales se tomaron 285 para el Análisis Factorial Exploratorio (EFA), 267 para el Análisis Factorial Confirmatorio (AFC) y 204 para el análisis test-retest de escuelas primarias públicas situadas en el norte de Rusia. En primer lugar, se realizó la validación semántica del instrumento por medio de un proceso de

traducción realizado en varias fases, posteriormente, la consistencia interna se midió mediante los coeficientes alfa de Cronbach y se utilizó el AFC para comprobar la validez factorial del instrumento. la validez de concurrencia usada para validar un nuevo test en contraste con otro ya validado previamente y la fiabilidad test-retest se evaluaron mediante el coeficiente intraclase. Finalmente, el análisis estadístico arrojó que la consistencia interna de la escala PDSS con un Alpha de Cronbach igual a 0,8 fue alta, la validez del constructo de la PDSS, respaldada por el AFC con cargas factoriales entre 0.438 y 0.727 referentes a los ítems ¿Suele estar alerta durante el día? y ¿Con qué frecuencia cree que necesita dormir más?, respectivamente. En ese sentido, en el estudio se concluyó que el instrumento puede utilizarse para muestras de jóvenes de habla rusa en la evaluación de la somnolencia diurna y la validez del instrumento está garantizada.

3.2 Antecedentes de Evaluación del Desempeño de Escalas para Variables Latentes

Vestin et al (2023), comprobaron las propiedades psicométricas de la escala de autoinforme de depresión MADRS-Y recientemente desarrollada para adolescentes en una muestra normativa sueca. En el estudio participaron 620 adolescentes con edades que iban desde los 12 hasta los 20 años, la muestra se dividió de manera aleatoria en dos partes iguales, con la primera parte de la muestra se realizó el Análisis de Componentes Principales ACP y con la restante se realizó el Análisis Factorial Confirmatorio AFC, además, se les solicitó a los participantes que volvieran a realizar los autoinformes al cabo de tres semanas para evaluar la fiabilidad test-retest. El resultado del ACP estableció que los 12 ítems potenciales que conforman la escala MADRS-Y son necesarios para explicar el constructo, por otro lado, el AFC apoyó la estructura unifactorial mediante buenos índices de ajuste. Además, se confirmó la propiedad de invarianza de la medida, para esto, los autores analizaron indicadores como el índice de ajuste comparativo CFI, el índice de ajuste no normativo Tucker-Lewis y la raíz del error cuadrático medio, determinando así, que los pesos o factores de carga de los ítems de la prueba son los mismos con relación a características como el género (femenino y masculino) o la edad (12-15 años y 16-20 años).

Belalcazar (2023), evaluó empíricamente el desempeño de métodos multivariados para la construcción de escalas como estrategia de expresión de variables latentes usando las bases teóricas de la teoría de la estimación en la estadística matemática, en el trabajo planteó tres estudios de simulación en los cuales contempló: el modelo unidimensional, el modelo de factores correlacionados con dos variables latentes y el modelo de factores correlacionados con tres variables latentes, respectivamente. En cada uno de los estudios de simulación, el autor especificó los modelo teóricos asociados y posteriormente

estableció los métodos de ponderación Multivariados a evaluar, además, analizó las propiedades de las escalas conformadas, mediante algunas medidas de desempeño como lo son: el pseudo error cuadrático medio, el Alpha de Cronbach y el porcentaje de varianza explicada. Una vez estructuró los escenarios de simulación, el autor realizó estudios con datos reales; adquiridos con diferentes instrumentos de medición, en los cuales existían dos o más variables latentes. En el caso específico del modelo reflectivo de variables latentes unidimensional, el estudio buscó generalizar tres tipos de instrumentos que en un contexto real representan: dependencia baja, dependencia media y finalmente la dependencia alta, siendo este último, el mejor escenario que puede tenerse en un instrumento (alto grado de validez y las variables se asumen altamente relacionadas con la variable latente).

4 Marco Teórico

En este capítulo se establece el marco teórico, considerando la importancia de definir ciertos conceptos relacionados con el área de la salud y los que están relacionados propiamente con la estadística. Se presenta el marco teórico dividido en dos secciones, la primera denominada como marco teórico conceptual y la segunda como marco teórico estadístico, en ellas, se explican los áspectos más fundamentales para la realización de la metodología posteriormente indicada.

4.1 Marco Teórico Conceptual-Clínico

4.1.1 Trastorno

Según Hamui Sutton (2019), el término trastorno se utiliza de un modo más laxo al referirse a los cambios desadaptativos, físicos o mentales, en los cuales no son muy claras las etiologías y las manifestaciones.

4.1.2 Insomnio

Según Beteta et al (2023) el insomnio es la dificultad para iniciar o mantener el sueño nocturno como consecuencia de factores psicológicos, biológicos y/o ambientales. Sin embargo, no todas las personas con problemas para dormir experimentan problemas en su funcionamiento diurno.

4.1.3 Somnolencia

La somnolencia diurna excesiva (SED) está definida en la ICSD3 (clasificación internacional de desordenes del sueño) como la incapacidad del individuo para mantener la vigilia y el estado de alerta durante todo el día, incluso quedarse dormido involuntariamente, Bironneau and Meurice (2023).

4.1.4 Vigilia

Según de define en el artículo de Friedman et al (2023), el estado de vigilia, también conocido como nivel de consciencia, es evaluado por diversas escalas descritas en términos como comatoso, estupor, obnubilación, delirio, confusión y alerta. El nivel de conciencia puede considerarse como un dial de volumen del que depende la conciencia fenoménica.

4.1.5 Ansiedad y Depresión

Dado el contexto de la investigación y la población de interés, Herskovic and Matamala (2020) en su revisión, plantea que los adolescentes con ansiedad presentan respuestas conductuales desadaptativas a la amenaza (inhibición conductual y evitación), mientras que los jóvenes deprimidos muestran respuestas conductuales desadaptativas al estrés (evitación y pobres estrategias de resolución de conflictos).

4.1.6 Baremos (puntos de corte)

Según Meneses and Barrios (2014), un baremo se trata de una herramienta en la que se establece, de manera estructurada, una conexión entre la puntuación de un individuo y el comportamiento normativo proporcionado por una muestra relevante de sujetos.

4.2 Marco Teórico Estadístico

4.2.1 Variable Latente

Bollen (2002) sostiene que "Las variables no medidas, los factores, las variables no observadas, los constructos o puntuaciones verdaderas, son solo algunos términos que los investigadores utilizan para referirse a las variables en el modelo que no están presentes en el conjunto de datos" (p. 607).

Según MacCallum and Austin (2000), las Variables latentes (LVs) son constructos hipotéticos que no pueden ser medidos directamente. En técnicas utilizadas para especificar y estimar modelos de relaciones lineales como los Modelos de Ecuaciones Estructurales (SEM), cada uno de estos constructos suele estar representado por múltiples variables medibles que sirven como indicadores del constructo.

Por otro lado, las variables latentes pueden ser modeladas con el fin de identificar las relaciones entre variables observables y la o las latentes. Según Hair Jr et al (2017), un modelo de estadístico de variables latentes; que generalmente tiene usos en el campo de la psicología y otras ciencias sociales está compuesto por los siguientes dos elementos:

I) **Modelo Interno o Estructural:** Este modelo representa las trayectorias estructurales entre constructos. Este es el que se encuentra bordeado por el recuadro de linea sólida en la figura 1.

II) **Modelo Exterior o de Medición:** Este modelo representa las relaciones entre cada variable latente y las variables observables o indicadoras asociadas. Este modelo está clasificado en Modelo reflectivo y Modelo formativo. Estos se encuentran delimitados por los recuadros punteados en la figura 1.

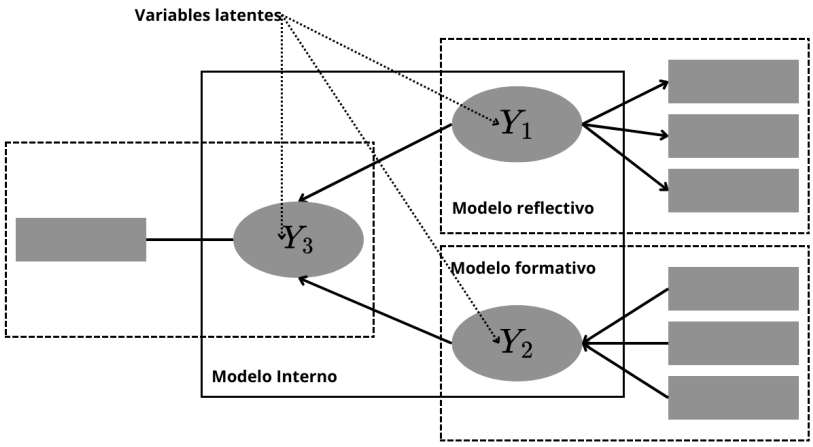

Figura 1 *Representación de los Constructos y SEM teórico (Gráfico de Trayectorias)*

En la definición formal que plantea Bartholomew et al (2011) para las variables latentes, se tratan todas las variables como continuas, por simplicidad notacional. En esta definición se tienen en cuenta dos tipos de variables. En primer lugar, se cuenta con las variables observables, denotadas como x, de manera que una colección de p variables observables pueden reescribirse como vector columna $\boldsymbol{x} = (x_1, x_2, ..., x_p)'$. Para las variables latentes, la notación es y y con subíndice q, de modo que una colección de variables latentes puede reescribirse como $\boldsymbol{y} = (y_1, y_2, ..., y_q)'$.

Debido a que tanto las variables latentes como las observables presentan variaciónes entre los individuos, estas son representadas teóricamente como variables aleatorias. Por tanto, las relaciones entre variables se deben expresar bajo la teoría de las distribuciones de probabilidad. En tal caso, después de haberse observado un conjunto de x_p variables, la información que se tiene sobre $\boldsymbol{y}$ estará contenida en la distribución condicional dado $\boldsymbol{x}$. Siendo que $\boldsymbol{x}$ es el conjunto de variables que sí se pueden observar, cualquier inferencia estará basada en su función de distribución de densidad conjunta, expresada como:

$$f(\boldsymbol{x}) = \int_{R_y} h(\boldsymbol{y}) g(\boldsymbol{x} \mid \boldsymbol{y}) dy \tag{1}$$

Donde $h(y)$ es la distribución a priori de $\boldsymbol{y}$, $g(\boldsymbol{x} \mid \boldsymbol{y})$ es la distribución condicional de $\boldsymbol{x}$ dado $\boldsymbol{y}$ y R_y es el espacio de rango de $\boldsymbol{y}$.

4.2.2 Modelo Reflectivo de Variables Latentes

En este modelo las variables observables son efectos de las variables latentes, por ejemplo, los puntajes númericos obtenidos en varias pruebas de razonamiento cuantitativo cuyas respuestas son medibles y exactas, serían indicadores de efecto de la variable latente del razonamiento cuantitativo, Bollen (2002).

Este modelo se puede expresar de la forma:

$$x_p = w_p \boldsymbol{y} + \epsilon_p \tag{2}$$

Donde:

- x_p es la variable observable p.
- w_p es un coeficiente (peso o carga) que encierra el efecto de la variable latente $\boldsymbol{y}$ sobre x_p.
- $\boldsymbol{y}$ es la variable latente.
- ϵ_p es la componente de error asociada a x_p.

Según Zapata (2012) en los modelos de medida reflectivos de variables latentes, las variables observables asociadas a una misma variable latente deben covariar, lo que implica que si existe un cambio en una de las variables observables, este cambio también se debe manisfestar en las demás.

4.2.3 Teoría de la Medición en Psicología

Según el artículo de Velleman and Wilkinson (1993), en el que se revisan las condiciones y razonamientos alrededor de la discusión sobre la aplicación del tipo de escala y su prohibición en algunos métodos estadísticos. Se asume lo siguiente:

i Sabemos qué atributos deseamos medir
ii Sabemos qué preguntas deseamos hacer sobre estos atributos
iii Hemos asignado números a estos atributos de manera que se preserven las características y relaciones relevantes entre los atributos.

Entonces,

i Las mediciones deben estar en uno de los tipos de escala reconocidos

ii Las transformaciones de los datos que no sean permitidas para ese tipo de escala de medición alterarán el significado transmitido en las mediciones y, por lo tanto, alterarán las respuestas a las preguntas que deseamos hacer
iii Por tanto, debemos seleccionar estadísticas entre aquellas apropiadas para esa escala de medición, porque su significado será invariante bajo las transformaciones permitidas de los datos.

A pesar de que estos elementos son teoremas establecidos, el argumento nace desde cómo, en la realidad, no hay una respuesta absoluta a estos enunciados, puesto que pueden surgir nuevas preguntas debidas a influencias externas, como Velleman and Wilkinson (1993) explica en su artículo. Para ello, se muestra como ejemplo la carta de Lord [Lord (1953)]. Frederic M. Lord, un destacado psicometrista, discutió en su carta cómo a veces se interpretan de manera errónea los números asignados en diferentes contextos, en este caso, los números de las camisetas de fútbol. Según Lord, estos números no miden atributos como habilidad o rendimiento, sino que son asignaciones arbitrarias utilizadas para identificar a un jugador. Según los críticos de la carta de Lord, los números de las camisetas de fútbol no miden atributos. Sin embargo, Lord pudo observar que los estudiantes que iban un año por encima se burlaban porque los de menor grado tenían números muy bajos, lo que en conclusión puede llevar a cambiar el razonamiento sobre la asignación de la escala. Esto es importante, puesto que es necesario que el analista de los datos sea flexible y considere la naturaleza de la escala.

Según Stevens (1946), el problema de las escalas de medición consiste en generar de manera explícita: a) las distintas reglas para la asignación de números, b) las propiedades matemáticas de las escalas resultantes y c) las operaciones estadísticas aplicables a las medidas realizadas con cada tipo de escala. Las escalas son posibles puesto que existe un cierto isomorfismo (correspondencia estructural exacta entre dos conjuntos de elementos, de modo que las relaciones y propiedades en un conjunto se preservan en el otro), entre lo que se puede hacer con los aspectos de los objetos y las propiedades de la serie numérica. Tal como muestra Stevens (1946), las escalas pueden catalogarse como nominal, ordinal, de intervalo o de razón, donde cada una sigue, a nivel matemático, una estructura de grupo al igual que diversas estadísticas permisibles invariantes.

Lo anterior tiene una implicación para el analista porque limita la gama de métodos estadísticos que se pueden usar. En la práctica una escala estrictamente definida como lo argumenta Stevens (1946), puede verse cuestionada por la manera en la que los datos reales no siguen las condiciones de las escalas. Según Tukey (1962), cuando las mediciones que deberían ser de escala de intervalo son tomadas con errores sistemáticos de calibración que dependen del valor medido, la diferencia entre dos valores en un extremo de la escala no serán comparables con la diferencia de dos valores al otro extremo de la escala. Por ejemplo, si se tiene un termómetro que se supone, mide la

temperatura en escala de intervalo, pero presenta un error que causa que las mediciones sean sistemáticamente más altas a medida que aumenta la temperatura, entonces esto representaría que 1 grado Celsius medido en el extremo más bajo de la escala puede ser significativamente diferente en términos reales de temperatura en comparación con la diferencia de 1 grado Celsius medido en el extremo más alto de la escala. Tukey (1962) concluyó que: "No se puede permitir que una visión demasiado simplificada y demasiado purificada de cómo son las mediciones dicte cómo se deben analizar los datos".

De acuerdo con Velleman and Wilkinson (1993), los análisis de datos deben guiarse por lo que se espera conocer de los datos, dado que los tipos de escala no son atributos fundamentales de los datos, si no que surgen de la manera en la que se miden los datos y sobre qué se quiere concluir.

- **Respuesta a ítems tipo Likert y su relación con el espacio continuo:** En la literatura de psicometría los datos observados mediante las categorías de respuesta de la escala likert se entienden como expresiones de la variable latente subyacente. Por ejemplo, cuando se modela la variable latente mediante un modelo estadístico de regresión, sucede que, en la mente del sujeto que responde a los ítems de una prueba: la escala likert es una función que por medio de un estímulo, puede enlazar una postura en el espacio continuo a un espacio en valores enteros. Boari and Nai Ruscone (2015), Belalcazar (2023).

4.2.4 Likert Equidistante como Suma Ponderada de Categorías de Respuesta

Chakrabartty (2023) plantea que debido al no cumplimiento de las propiedades de equidistancia, la suma de puntajes de elementos medidos en escala Likert puede no ser significativa. En su artículo muestra el comportamiento de la varianza obtenida a través de ACP para un ítem medido con escala Likert, utilizando muestras con diferentes distribuciones; uniforme, bipolarizada, con tendencia central, con falsos positivos (respuestas para dar buena imagen), y falsos negativos (respuestas pesimista).

La metodología para transformar ítems ordinales en ítems de escala continua implica seleccionar pesos basados en las frecuencias de las categorías de respuesta de un ítem, de tal manera que $W_i > 0$ y $\sum W_i = 1$, ante el puntaje transformado del i-ésimo individuo por elegir la j-ésima categoría de respuesta de un ítem. Las propiedades que los pesos deben seguir, son descritas continuación:

1. Ser continuos
2. Monotonicidad, un puntaje trasnformado alto debe corresponder a una respuesta al ítem en un nivel alto.

$$5W_5 > 4W_4 > 3W_3 > 2W_2 > W_1 \tag{3}$$

3. Equidistante es decir,

$$5W_5 - 4W_4 = 4W_4 - 3W_3 = 3W_3 - 2W_2 = 2W_2 - W_1 \tag{4}$$

4. Normalidad, los puntajes equidistantes deben estandarizarse de forma que $\sim Normal(0, 1)$, esta puede ser tranformada para seguir una distribución con media y varianza distinta de 0 y 1.

Un método que plantea Chakrabartty (2023) para transformar ítems ordinales (likert) en ítems de escala continua, es denominado como:

- Diferentes pesos para diferentes categorías de respuesta de diferentes ítems, considerando la frecuencia de la $(i - j)$ ésima celda.

En este caso, $W_{ij} = \frac{f_{ij}}{\sum_{j=1}^{k} j \cdot f_{ij}}$, es decir, la proporción del número de veces que se eligió la $j-$ésima categoría de respuesta del $i-$ésimo ítem y la puntuación del $i-$ésimo ítem.

Evidentemente, $W_{ij} > 0$ y $\sum_{j=1}^{5} W_{ij} = 1$ (considerando que la escala likert tenga 5 posibles respuestas). No obstante, la suma de los pesos para una categoría de respuesta de todos los ítems es diferente de uno.

La puntuación del individuo que selecciona la j-ésima categoría de respuesta del i-ésimo ítem será $W_{ij} \cdot X_{ijk}$, donde X_{ijk} es el valor de la categoría $j-$ésima seleccionada por el $k-$ésimo individuo en el $i-$ésimo ítem. Es importante señalar que, para un ítem específico, los pesos aumentan a medida que aumentan los valores de anclaje asociados con las categorías de respuesta. Así, se cumplen tanto la condición de monotonicidad como la condición de equidistancia para las puntuaciones transformadas de cada ítem.

4.2.5 Análisis Factorial Múltiple (AFM)

El AFM es un método multivariado, el cual permite realizar análisis a datos que están constituidos por un conjunto de individuos descritos por varios grupos de variables, cada grupo corresponde a una tabla. En Escofier and Pagès (1998) el AFM se desarrolla con base en la siguiente notación y procedimiento:

$\boldsymbol{X}$ es la tabla completa, $\boldsymbol{I}$ todos los individuos, $\boldsymbol{K}$ todas las variables, $\boldsymbol{K_j}$ todas las variables del grupo j; $(\boldsymbol{K} = \cup_j \boldsymbol{K_j})$, $\boldsymbol{X_j}$ las tablas asociadas al grupo j.

El AFM se desarrolla en tres fases, la primera consta de asociar una nube de individuos (Nube parcial N_I^j) a cada $\boldsymbol{K_j}$ grupo de variables y realizar un análisis independiente por medio de ACP (variables cuantitativas), ACM (variables cualitativas) o ACS (tablas de contingencia).

En la segunda fase busca que la influencia de los diferentes grupos de variables este equilibrada, por lo que se asigna un peso a las variables de un mismo grupo, normalizando cada grupo o tabla. Sea $d^2(i,l)$ la suma de los cuadrados de las distancias entre dos puntos i y l en una nube parcial y $\alpha_j = \frac{1}{\lambda_1^j}$ la ponderación o peso de cada $\boldsymbol{K_j}$ calculada teniendo en cuenta el primer valor propio generado por su respectivo análisis (ACP, ACM o ACS), entonces:

$$d^2(i,l) = \sum_{j \in J} \alpha_j d^2(i^j, l^j) \tag{5}$$

Por último en la tercera fase, se quiere proyectar la nube de individuos en un subespacio de pequeña dimensión, para ello se realiza un ACP sobre $\boldsymbol{X}$ con ponderaciones sobre las variables, $\boldsymbol{X}$ está formada por todas las tablas yuxtapuestas.

4.2.6 Pseudo Error Cuadrático Medio

Tal como se plantea en la teoría de inferencia estadística, el error cuadrático medio (ECM) de un estimador W de un parámetro θ, es una función de θ definida como $E_\theta(W-\theta)^2$. El ECM mide la distancia media al cuadrado de entre el estimador W y el parámetro θ, de esa manera, se puede conocer el desempeño del estimador puntual. El ECM tiene un ventaja frente a otras posibles medidas de desempeño como el error absoluto medio, tal y como, su interpretabilidad y que es manejable analíticamente [Casella and Berger (2024) (p. 330)].

El Error cuadrático medio puede expresarse como:

$$E_\theta(W-\theta)^2 = Var_\theta(W) + (E_\theta(W) - \theta)^2 \tag{6}$$

Donde $Var_\theta(W)$ es la varianza del estimador y $(E_\theta(W) - \theta)$ el sesgo del estimador.

Desde la perspectiva estadística, un parámetro se entiende como un valor numérico que representa o describe el comportamiento característico de la variable medida en la población objeto de estudio. Además, formalmente el parámetro es una característica asociada a la función de probabilidad asumida; que mejor modela el comportamiento natural de la variable Belalcazar (2023).

Bajo la teoría de variables latentes, la función de densidad de probabilidad o de masa de probabilidad del conjunto de variables observables o medibles, puede expresarse a partir de la función de probabilidad condicional dada la variable latente, tal y como se observa en la ecuación 1. Así pues, sea $h(\boldsymbol{y})$ la función de probabilidad de densidad o masa y $h(\boldsymbol{y} \mid \boldsymbol{\theta})$ la función de

probabilidad de densidad o masa dado un parámetro θ, entonces, θ sería el valor numérico que describe el comportamiento natural de la variable latente. Identificar la distribución de probabilidad $h(\boldsymbol{y} \mid \boldsymbol{\theta})$ no es posible, ya que las características propias del constructo hacen que su parámetro no se puede saber a ciencia cierta, esto debido a que, el constructo o variable latente tiene un comportamiento aleatorio que se representa en las variables aleatorias observables. Esto implica que no hay una medición directa de la variable en la unidad observacional y no se podrían asumir los conceptos de la teoría de inferencia estadística.

Belalcazar (2023) propone una aproximación empírica de θ, él argumenta que bajo el contexto de las variables latentes, el valor que toma la escala de medición para un individuo puede ser visto como un estimador de lo que sería la expresión de la variable latente. Para un conjunto de individuos, una medida resumen como la media de los puntajes (en caso de que se presente simetría en la distribución de los puntajes de los individuos en las escalas) sería un estimador de la expresión media de la variable latente en la población objeto de estudio, convirtiendo así, a esta medida en un pseudo-parámetro. Considerando estos supuestos, el estimador W del pseudo-parámetro está definido como:

$$W = \hat{\theta} = \frac{\sum_{i=1}^{n} \psi_i}{n} \quad ; \; i = 1, 2, ..., n_l \tag{7}$$

Donde ψ_i es el puntaje del individuo i en la escala y n es el total de individuos en el conjunto.

Con la finalidad de tener una aproximación al verdadero valor del parámetro θ, Belalcazar (2023) plantea el cálculo $\bar{\theta}$ (pseudo-parámetro) como la media de los puntajes en la escala para el conjunto de m muestras con tamaño n. Evidentemente no se tiene en cuenta el tamaño de muestra, ya que se toma todo el conjunto de las m muestras, considerando todos los elementos de los $n \cdot m$ vectores de valores de puntajes en la escala (por ejemplo: $n = 30, 50, 100, 1000$ y $m = 1000$).

Dado lo anterior, no es valido pensar en ECM para el contexto de las variables latentes, por tal motivo, se plantea una aproximación denominada como Pseudo Error Cuadrático Medio (PECM). Este puede ser expresado como:

$$\begin{aligned} PECM &= Var(\hat{\theta}_m) + (E_{\bar{\theta}}(\hat{\theta}_m) - \bar{\theta}))^2 \\ &= Var(\hat{\theta}_m) + (sesgo(\hat{\theta}_m, \bar{\theta}))^2 \end{aligned} \tag{8}$$

4.2.7 Algoritmo K-means

Según el trabajo realizado por Ikotun et al (2023), el algoritmo de agrupamiento de K-means se clasifica como un algoritmo de agrupamiento particional. Dividir conjuntos de datos mediante clusters implica, en primer lugar, encontrar el error cuadrático mínimo entre los diversos puntos de datos en el conjunto de datos, además de la media de un cluster, para posteriormente, asignar cada punto de datos al centro del cluster más cercano a él.

Matemáticamente, dado un conjunto de datos $\boldsymbol{X} = x_i$ donde $i = 1, 2, ..., n$ de puntos de datos de dimensión d con tamaño n,

$\boldsymbol{X}$ es particionado en "k" grupos, $C = c_j$ donde $j = 1, 2, ..., k$ tal que

$$J(c_k) = \sum_{x_i \in c_k} \|x_i - \mu_k\|^2 \tag{9}$$

El algoritmo K-means tiene como objetivo minimizar la suma del error cuadrático para cada k grupo. Eso es, minimizar

$$J(C) = \sum_{k=1}^{K} \sum_{x_i \in c_k} \|x_i - \mu_k\|^2 \tag{10}$$

Para lograrlo, el algoritmo K-medias inicialmente, elige aleatoriamente un número k específico de centroides del conjunto de datos.

4.2.8 Prueba Chi-Cuadrado de asociación 2 x N

Mendivelso and Rodríguez (2018) presenta la prueba χ^2 como una prueba libre de distribución. En el caso de evaluar independencia, se busca observar la diferencia entre las frecuencias de una muestra bajo dos atributos distintos.

- Supuestos: Los datos provienen de una muestra aleatoria y la muestra es suficientemente grande
- Limitaciones: Si menos del 20 % de las celdas presentan frecuencias ≤ 5 no es recomendable desarrollar la prueba
- Hipótesis:
  - H_0 : No hay asociación entre las variables (independencia)
  - H_1 : Si hay asociación entre las variables (dependencia)
- Sea O_i el valor observado y E_i el valor esperado, el estadístico de prueba es:

$$T = \sum_{i=1}^{r} \sum_{j=1}^{c} \left[\frac{(O_{ij} - E_{ij})^2}{E_{ij}} \right] \qquad , donde \quad E_{ij} = \frac{R_i C_j}{N} \tag{11}$$

Si H_0 es verdadera, sigue una distribución χ^2 con (r-1)(c-1) grados de libertad, donde r es el número de filas y c el nummero de columnas.

4.2.9 Prueba de bondad de ajuste Anderson-Darling

La prueba Anderson-Darling es una prueba de bondad de ajuste utilizada para evaluar si una muestra de datos sigue una distribución específica, generalmente es usada para evaluar si los datos siguen una distribución normal. Para desarrollar esta prueba se recomienda tener en cuenta lo siguiente, (D'Agostino, 2017).

Hipótesis:

- H_0 : Los datos siguen una distribución especificada.
- H_1 : Los datos no siguen una distribución especificada.

Para decidir sobre las hipótesis, se evalúan con los siguientes pasos:

1) Se Ordena la muestra de manera ascendente, tal que:

$$X_{(1)} \leq ... \leq X_{(n)} \quad \text{, donde } n \text{ es el tamaño de la muestra}$$

2) Se calcula los valores estandarizados, $Y_{(i)}$, donde:

$$Y_{(i)} = \frac{X_{(i)} - \bar{X}}{S} \qquad para \quad i = 1, ..., n$$

3) Se calcula P_i para $i = 1, ..., n$, donde P_i es la probabilidad acumulada para el valor estandarizado $Y_{(i)}$ que sigue una función de distribución acumulada normal estándar (D'Agostino, 2017).
4) Se calcula el estadístico Anderson-Darling

$$A^2 = -\sum_{i=1}^{n} \left[\frac{2i-1}{n} \{\ln P_i + \ln(1 - P_{n+1-i})\} \right] - n$$

5) Se calcula el estadístico modificado

$$A^* = A^2(1 + 0,75/n + 2,25/n^2)$$

6) se rechaza la hipótesis de nula si A^* excede los valores 0,631, 0,752, 0,873, 1,035 y 1,159 en niveles de significancia α igual a 0,10, 0,05, 0,025, 0,01 y 0,005, respectivamente (D'Agostino, 2017).

5 Metodología

En este capítulo se define la metodología, para ello, se presenta un análisis bibliométrico donde se identificó literatura necesaria para el desarrollo del documento. Adicionalmente, se estableció una esquematización (**Figura** 2), la cual permite entender los pasos a seguir para el logro de los objetivos propuestos en el capítulo 5, posteriormente, se establece la descripción detallada de la metodología. En esta se presenta toda la estructura detrás de los datos, con los cuales se realizan todas las técnicas estadísticas para el diseño de las respectivas medidas sintéticas, permitiendo así, conocer el nivel de presencia de trastornos del sueño en escolares y adolescentes.

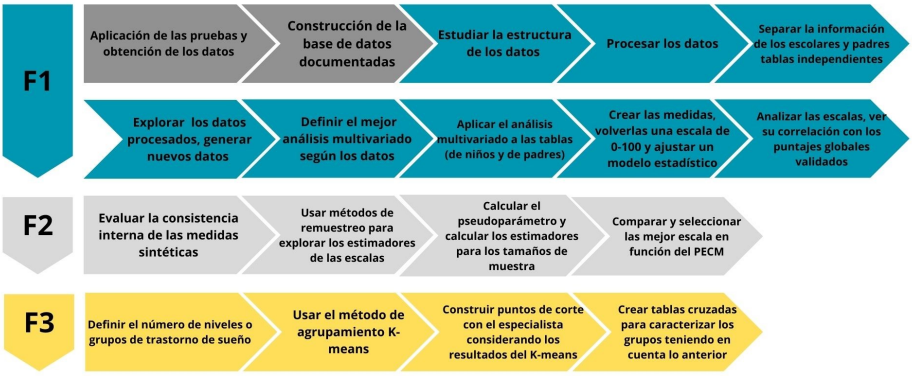

Figura 2 *Metodología Aplicada en la Construcción de Medidas Sintéticas para Identificar Transtornos del Sueño en Escolares y Adolescentes*

5.1 Análisis Bibliométrico

El análisis bibliométrico se ha convertido en una herramienta muy importante para los investigadores, pues este, facilita enfocar la búsqueda y generar conexiones entre autores. Para la realización de este trabajo, se usó Scopus, una de la más grande base de datos de citas y resúmenes de literatura científica, además, se implementó el software VOSviewer para generar el análisis de autores con más de 3 publicaciones y palabras claves asociadas a la búsqueda.

Con el fin de maximizar la capacidad de búsqueda y obtener mejores resultados, se estableció la ecuación TITLE-ABS-KEY (sleep AND disorders AND in AND children AND adolescent AND scores AND test) AND PUBYEAR $>$ 1969 AND PUBYEAR $<$ 2024 como criterio de búsqueda.

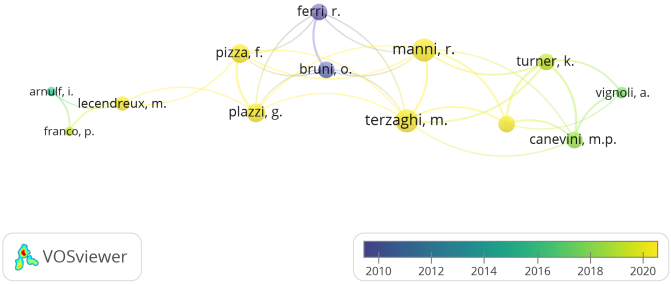

Figura 3 *Análisis por Autor en VOSviewer*

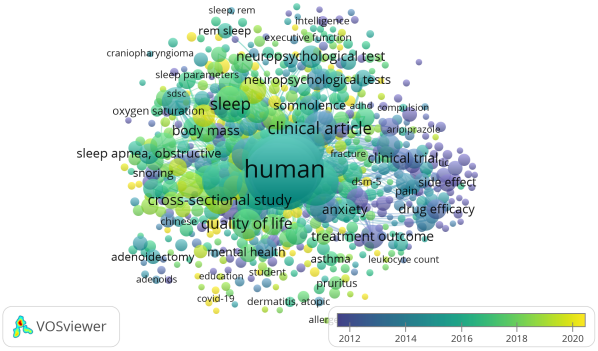

Figura 4 *Análisis por palabras clave en VOSviewer*

5.2 Descripción del Archivo de Datos y la Fuente de Información

Tal como se aprecia en la **Figura** 2, el cumplimiento de los objetivos parte de la recolección de información de una serie de pruebas diagnósticas que miden la presencia de diferentes tipos de trastornos del sueño a escolares (9 a 12 años) y adolescentes (13-18 años), en particular, estudiantes de un colegio, ubicado al oriente de Santiago de Cali. El proceso consistió en la aplicación de pruebas diagnósticas a la totalidad de los estudiantes desde sexto hasta undécimo grado del colegio. Previo a la aplicación, tanto los estudiantes como sus padres fueron informados detalladamente sobre el procedimiento, obteniéndose el consentimiento informado correspondiente. La recolección de datos fue supervisada por profesionales de la salud, quienes brindaron orientación para garantizar el correcto diligenciamiento de los cuestionarios y asegurar la calidad de los datos obtenidos. De esta manera, se construyó el archivo de datos con las respuestas a las pruebas, con un total de 827 registros y 365 variables. Como el archivo original cuenta con variables alusivas a información general, información antropométrica, respuestas a los ítems del cuestionario Roma IV (trastornos funcionales gastrointestinales) y respuestas a los ítems

de los diferentes cuestionarios para diagnosticar trastornos del sueño, solo se tomó en cuenta lo relacionado con trastornos del sueño, información sobre la presencia de por lo menos un trastorno funcional gastrointestinal y otra información como el IMC e información sociodemográfica para el análisis, disminuyendo la dimensión del archivo de datos a (827 registros y 84 variables).

Dado que el desarrollo de los pasos de aplicación de las pruebas, obtención de datos y construcción de la base da datos inicial, ya habían sido implementados, en esta Metodología no se tienen en cuenta, sin embargo, para garantizar reproducibilidad en importante ejecutarlos, esto, corresponde a las fases con color gris que aparecen en la **Figura** 2.

5.3 Pre-procesamiento de datos

Por medio del *sofware Microsoft excel* se construyó finalmente un archivo de datos con 428 registros y 84 variables, garantizando que no hubiesen datos faltantes en las variables que participan en la construcción de los puntajes de las diferentes pruebas, tanto para escolares y adolescentes, como para padres.

Por otro lado, dado que había ítems que, en lugar de indicar malas condiciones relacionadas con el sueño en los escolares o adolescentes, identificaban una buena condición, estos ítems fueron modificadas por medio del *software R*, de tal forma que las respuestas tomaran un mayor valor dentro de la escala para los escolares y adolescentes que presentaran esas buenas condiciones con menor frecuencia. Los ítems que deben ser transformados para un correcto desarrollo de la metodología, son los que cuentan con un asterisco en las tablas 2 y 4. Un ejemplo de este tipo de ítem es ASWS10 presentada en la tabla 2, cuya pregunta es la siguiente:

- ¿Cuando me despierto por la mañana, me siento descansado y alerta?

Esta pregunta indica que el escolar o adolescente tiene un correcto descanso y está preparado para desarrollar las actividades diarias. Por tanto, los individuos que de manera menos frecuente presenten este comportamiento, deben tener un mayor valor en la escala del ítem transformado. Según Chyung et al (2018), el proceso de codificación inversa requiere que el ítem del instrumento cuente con un descriptor negativo apropiado que pueda ser correctamente invertido a su contraparte. Esto quiere decir que las categorías de respuesta sean intercambiables y la perspectiva del encuestado no se vea alterada.

Dado que las categorías de respuesta de los ítems (con asterisco) reflejan la frecuencia con la que se realiza una afirmación o se experimenta una situación, y considerando que en la práctica es conveniente trabajar con los puntajes de la escala Likert para la construcción de medidas, es razonable emplear la codificación inversa. Así pues, se define el puntaje inverso x_{inv} como:

Sea x_i es el valor de la categoría a invertir en la escala original con i igual al número de categorías dentro de la escala de medición likert. Entonces

$$x_{inv} = max(\boldsymbol{x}) + 1 - x_i \qquad ;\ i = 1, ..., k \tag{12}$$

Donde $max(\boldsymbol{x})$ es el máximo valor de las categorías de respuesta.

Por ejemplo, si se quisiera invertir el ítem promis sd 2*, cuyas categorías de respuesta son:

- $1 = nunca$, $2 = casi\ nunca$, $3 = algunas\ veces$, $4 = casi\ siempre$ y $5 = siempre$

Entonces, la codificación inversa sería:

$$5 + 1 - 5 = 1$$
$$5 + 1 - 4 = 2$$
$$5 + 1 - 3 = 3$$
$$5 + 1 - 2 = 4$$
$$5 + 1 - 1 = 5$$

5.4 Construcción de las Tablas para los respectivos análisis

En la tabla 1 se observan las variables sociodemográficas y de salud de los escolares y adolescentes que se pretenden analizar para describir las características asociados a los trastornos de sueño en el grupo de escolares y adolescentes.

Puntajes Pruebas			
Información	Codif.	Descripción	Rango
Sociodemográfico y de Salud	sexo	Sexo del individuo	{0, 1}
	grupedad	Grupo al que pertenece según su edad, de 9 a 12 escolar (0) y 13 a 18 años adolescente (1)	{0, 1}
	dx imc	Diagnóstico según el Índice de Masa Corporal	{obesidad, sobrepeso, normal, delgadez, delgadez severa}
	enfermedad	El individuo sufre de alguna enfermedad	{0, 1}
	enfermedad fliar	En su familia, ¿Alguien sufre de alguna enfermedad?	{0, 1}
	dfg2	Presenta por lo menos un desorden gastrointestinal funcional	{0, 1}

Tabla 1 *Descripción de las variables sociodemográficas y de salud*

En la tabla 2 se pueden observar todos los ítems de las pruebas realizadas a los escolares y adolescentes. Las pruebas están catalogadas según el tipo de trastorno relacionado con el sueño que miden. El insomnio, la somnolencia, la alteración del sueño, el deterioro, el sueño y la vigilia, la ansiedad y la depresión conforman, por sí solos, un tipo de trastorno del sueño o, bien sea, un estado relacionado con la variable latente (nivel de trastorno del sueño) que presenta el escolar o adolescente.

En la tabla 3 se indican los puntajes totales o resultantes de cada instrumento aplicado a los escolares y adolescentes. No todos los puntajes se determinan con la misma fórmula, algunos de ellos son T-score, sumas totales o sumas ponderadas.

PISI			
Trastorno	Codif.	Pregunta/Afirmación	Rango
Insomnio	pisi 1	Me demoro más de 30 minutos para quedarme dormido después de acostarme	$0-5$
	pisi 2	En general, se me dificulta irme a acostar	
	pisi 3	Durante la noche, me despierto más de una vez	
	pisi 4	Después de despertarme durante la noche, es difícil quedarme dormido de nuevo	
	pisi 5	Me siento con sueño durante el día	
	pisi 6	¿Cuántas horas de sueño duermes en la mayoría de las noches?	
ESS			
Trastorno	Codif.	Pregunta/Afirmación	Rango
Somnolencia	ess 1	Sentarse y leer	$0-3$
	ess 2	Sentarse y ver televisión o un video	
	ess 3	Sentarse en un salón de clase durante la mañana	
	ess 4	Sentarse y viajar en un carro o bus durante aproximadamente media hora	
	ess 5	Acostarse a descansar o dormir la siesta por la tarde	
	ess 6	Sentarse y hablar con alguien	
	ess 7	Sentarse solo tranquilamente después del almuerzo	
	ess 8	Sentarse y comer una comida	
Promis sd			
Trastorno	Codif.	Pregunta/Afirmación	Rango
Alteración	promis sd 1	Tuve dificultad para poder dormirme	$1-5^*$
	promis sd 2*	Dormí toda la noche	
	promis sd 3	Tuve 1 problema con mi sueño	
	promis sd 4	Tuve problemas para dormir	
	promis sd 5	Me tomó mucho tiempo quedarme dormido	
	promis sd 6	Me preocupaba no poder quedarme dormido	
	promis sd 7	Me desperté por la noche y tuve problemas para volver a dormirme	
	promis sd 8	Daba muchas vueltas en la cama por no poder dormir	
Promis SRI			
Trastorno	Codif.	Pregunta/Afirmación	Rango
Deterioro	promis sri 1	Tuve sueño durante el día	$1-5$
	promis sri 2	Me costaba concentrarme porque tenía sueño	
	promis sri 3	Me costaba mucho hacer las cosas porque tenía sueño	
	promis sri 4	Tuve problemas durante el día porque dormí poco	
	promis sri 5	Tuve problemas para mantenerme despierto durante el día	
	promis sri 6	Era difícil divertirme porque tenía sueño	
	promis sri 7	No podía mantener los ojos abiertos durante el día	
	promis sri 8	Estaba de mal humor porque tenía sueño	
ASWS			
Trastorno	Codif.	Pregunta/Afirmación	Rango
Sueño y Vigilia	asws 1	Cuando es hora de acostarse, quiero quedarme despierto y hacer otras cosas (por ejemplo, ver televisión, jugar videojuegos o hablar por teléfono)	$1-6^*$
	asws 2*	Cuando es hora acostarse, en general, estoy listo para irme a la cama a la hora de dormir	
	asws 3	Cuando es hora de acostarse, en general, trato de aplazar o retrasar la hora de irme a la cama	
	asws 4	Cuando es hora de irse a dormir (luces apagadas), tengo problemas para sentirme cómodo	
	asws 5	Cuando es hora de irse a dormir (luces apagadas), necesito ayuda para dormirme (por ejemplo: escuchar música, ver televisión, tomar medicamentos, tener a alguien más en la cama conmigo)	
	asws 6	Cuando me despierto en la noche, tengo problemas para volver a dormirme	
	asws 7	Cuando me despierto en la noche, tengo problemas para sentirme cómodo	
	asws 8	Cuando me despierto en la noche, necesito ayuda para volver a dormir (por ejemplo: ver televisión, leer o dormir con otra persona)	
	asws 9*	Cuando me despierto por la mañana, me siento preparado para levantarme de la cama y empezar el día	
	asws 10*	Cuando me despierto por la mañana, me siento descansado y alerta	
Promis Anxiety			
Trastorno	Codif.	Pregunta/Afirmación	Rango
Ansiedad	promis anxiety 1	Sentí que algo horrible podría pasar	$1-5$
	promis anxiety 2	Me sentí nervioso	
	promis anxiety 3	Me sentí asustado o con miedo	
	promis anxiety 4	Me sentí preocupado	
	promis anxiety 5	Me preocupaba cuando estaba en casa	
	promis anxiety 6	Me asusté muy fácil	
	promis anxiety 7	Me preocupaba por lo que podría pasarme	
	promis anxiety 8	Me preocupaba cuando me iba a la cama por la noche	
Promis Depression			
Trastorno	Codif.	Pregunta/Afirmación	Rango
Depresión	promis depression 1	No podía dejar de sentirme triste	$0-4$
	promis depression 2	Me sentí solo	
	promis depression 3	Sentí que todo en mi vida salió mal	
	promis depression 4	Sentía que no podía hacer nada bien	
	promis depression 5	Me sentí aislado	
	promis depression 6	Me sentí triste	
	promis depression 7	Me sentí infeliz	
	promis depression 8	Me costaba divertirme	

Tabla 2 *Descripción de los ítems de cada instrumento evaluado para medir un tipo trastorno relacionado con el sueño en escolares y adolescentes.*

Puntajes Pruebas			
Información	Codif.	Descripción	Rango
Resultados Pruebas	pisi total	Suma de los puntajes de las preguntas de la prueba pisi	$0-30$
	ess total	Suma de los puntajes de las preguntas de la prueba ess	$0-24$
	T-score promis sd total	T-score reescalado: (Suma de puntajes de preguntas * 8) / número de preguntas respondidas)	$36{,}6-82{,}7$
	T-score promis sri total	T-score reescalado: (Suma de puntajes de preguntas * 8) / número de preguntas respondidas)	$37{,}4-84$
	asws total/10	Suma de los puntajes de asws / 10	$1-6^*$
	T-score promis anxiety total	T-score reescalado:(Suma de puntajes de preguntas * 8) / número de preguntas respondidas	$37{,}1-83{,}1$
	T score promis depression	T-score reescalado:(Suma de puntajes de preguntas * 8) / número de preguntas respondidas	$35-82$

Tabla 3 *Descripción de los puntajes totales de cada instrumento evaluado para medir un tipo trastorno relacionado con el sueño en escolares y adolescentes.*

En la tabla 3 se muestran los rangos de puntajes totales correspondientes a cada instrumento. Cada uno de estos rangos incluye clasificaciones asociadas al grado de trastorno del sueño que puede presentar el escolar o adolescente (ver en 8.3).

En la tabla 4 se aprecian los ítems del intrumento que mide la higiene del sueño de los escolares y adolescentes. Estos ítems fueron usados para evaluar las condiciones del sueño en los escolares y adolescentes a través de la percepción de los padres o figuras paternas.

SHIP			
Dimensión	Codif.	Pregunta/Afirmación	Rango
Higiene del Sueño	ship 1*	Se va a dormir a la misma hora por la noche	$1-3^*$
	ship 2*	Se duerme a los 20 minutos después de acostarse	
	ship 3*	Está listo a acostarse a la hora de irse a dormir	
	ship 4	Se resiste a acostarse a la hora de irse a dormir	
	ship 5	Duerme muy poco	
	ship 6	Duerme demasiado	
	ship 7*	Duerme lo normal	
	ship 8*	Duerme lo mismo todos los días	
	ship 9	Se queja de problemas para dormir	
	ship 10	Se despierta 1 vez en la noche	
	ship 11	Se despierta más de 1 vez en la noche	
	ship 12*	Se levanta solo en la mañana	
	ship 13	Tiene problemas para levantarse solo en la mañana	
	ship 14	Duerme siesta en el día	
	ship 15	Parece cansado	

Tabla 4 *Descripción de los ítems de los instrumentos aplicados a los padres donde se evalua la higiene del sueño evidenciadas en los escolares y adolescentes*

5.4.1 Comparación en el Análisis, Teoría de la Medición y Likert Equidistante como Suma Ponderada de Categorías de Respuesta

Antes de realizar cualquier análisis de datos, es necesario comprender las características de los datos, para este trabajo se tuvieron en cuenta tres alternativas de medida sintética para escolares y adolescentes (AFM, ACP y ACP para puntajes totales), en dos de ellas (AFM y ACP), se consideraron tanto los datos en bruto como los datos transformados mediante una técnica Likert equidistante como suma ponderada de categorías de respuesta.

La escala Likert equidistante como suma ponderada de categorías de respuesta, desarrollada en el artículo de Chakrabartty (2023), busca asignar diferentes pesos a las categorías considerando la frecuencia de respuestas, esto hace que la escala tenga en cuenta el proceso generador de los datos y el contexto de quienes responden a los items de las pruebas.

5.5 Análisis Exploratorio de los Datos

Para la obtención de resultados de este capítulo se analizaron tanto características básicas de los individuos, como también, la distribución de las respuestas a los diferentes ítems para cada prueba. La idea principal de hacer esta primera fase exploratoria es caracterizar el conjunto de datos con el fin de preestablecer conceptos asociados a los mismos, para lo cual se implementaron tablas cruzadas.

5.6 Análisis Multivariado

Las técnicas de análisis multivariado tienen gran relevancia en el campo de la psicometría, tal como se ha explicado anteriormente, el análisis factorial permite trabajar las variables latentes; presentes en una gran gama de estudios de psicología.

Haciendo uso de la librería en R "*FactoMineR*" Husson et al (2016) se crearon los resultados; análisis de componentes principales (ACP) y análisis factorial múltiple (AFM), con los cuales se construyen las diferentes medidas sintéticas e interpretación de la nube de variables e individuos. Para el análisis factorial múltiple se tuvo en cuenta las variables contenidas en las dimensiones (también denotadas como tablas) Insomnio, Somnolencia, Alteración, Deterioro, Sueño y Vigilia, Ansiedad y Depresión, así mismo, mediante componentes principales se abordó tanto de manera individual cada dimensión como el conjunto de primeros componentes. Además, para la creación de gráficos se recurrió a la libreria en R "*factoextra*" Kassambara and Mundt (2017).

5.6.1 Construcción de la medida sintética para escolares-adolescentes (alternativas I, II y III)

Con el fin de construir las distintas medidas sintéticas que representen el nivel de trastorno del sueño para realizar posteriormente su comparación, dependiendo de la tabla a ser analizada, se establecen los siguientes pasos:

Alternativa I (AFM)

i) Se debe conformar la matriz de correlación del conjunto de datos, observando si todas las correlaciones son positivas. Esta condición corresponde al concepto llamado factor tamaño (González Rojas, 2014). Contar con correlaciones mayores al 70 % permite obtener resultados más consistentes (ver estudio de simulación de Belalcazar (2023)). En caso de que no todos los ítems estén correlacionados de manera positiva, se deben transformar los ítems respectivos como se plantea en el capítulo de pre-procesamiento de datos (5.3).

ii) Definir las subtablas a tener en cuenta, estas serán equivalentes a las dimensiones y los factores serán, por ende, los ítems de las respectivas pruebas.

iii) Realizar el análisis factorial múltiple apoyándose de la librería en R "*FactoMineR*" en su función *AFM()*, para esto, se tendrá que definir k Tablas para analizar, donde $k = 1, 2, ..., 7$ y la cantidad j de ítems presente en cada Tabla k. Además, se debe establecer que los datos sean escalados a su varianza unitaria.

iv) Extraer el primer vector propio asociado al análisis global y los valores propios de cada subtabla o dimensión.

v) Crear el vector Mu, donde u será el primer vector propio extraido en el paso anterior y M será la matriz diagonal con valores $\frac{1}{\lambda_{jk}}$, donde λ_{jk} es el autovalor j asociado a la tabla k. Asegúrese que los pesos del vector Mu sumen 1, entonces, cada elemento j de la tabla k de Mu ponderado se le denomina peso α_{jk}.

vi) Sea X_{ijk} el puntaje del individuo i para el ítem j de la tabla k y α_{jk} el peso del ítem j de la tabla k. Entonces, se calcula la medida sintética para el individuo i que indique el nivel de trastorno del sueño que presenta, de la siguiente manera:

$$ITSN1_i = \sum_{k=1}^{K} \sum_{j=1}^{J} \alpha_{jk} X_{ijk} \tag{13}$$

Si se le aplica la transformación Likert Equidistante a las puntuaciones de los individuos en los ítems, entonces el cálculo de la medida sintética estaría dado por la siguiente ecuación:

$$ITSNE1_i = \sum_{k=1}^{K} \sum_{j=1}^{J} \alpha_{jk} X_{ijk}* \qquad ; \qquad X_{ijk}* = W_{ij} \cdot X_{ijk} \tag{14}$$

viii) En último lugar se definen los intervalos generales (a, b) en los cuales se puede mover la medida sintética, para ello, se debe asignar los valores mínimos y máximos de la escala tipo likert de los ítems X_{ij} o bien sea, de los ítems transformados $X_{ij}*$.

Alternativa II (ACP)

i) Desarrollar el paso *i*) planteado en la alternativa I.

ii) Realizar un análisis de componentes principales para cada dimensión, apoyándose de la librería en R "*FactoMineR*" en su función *ACP()*, por tanto, se obtendrán 7 análisis con sus respectivos resultados.

iii) Obtener el primer componente por cada dimensión k y unificarlos en una nueva tabla, de tal modo que las columnas de dicha tabla sean los ψ_{1k} con $k = 1, 2, ..., 7$.

iv) Realizar nuevamente ACP, para esta ocasión el análisis se hará a la tabla del paso anterior. Este segundo ACP puede revelar si los cuestionarios comparten estructuras comunes que permitan sintetizar mejor el constructo.

v) Extraer el primer vector propio asociado al análisis de componentes principales del paso anterior. De forma que los pesos del vector sumen 1, entonces, cada elemento j de este vector es el peso β_j.

vi) Sea Y_{ij} una matriz compuesta por los primeros vectores propios asociados a cada dimensión. Entonces, se calcula la medida sintética para el individuo i que indique el nivel de trastorno del sueño que presenta de la siguiente manera:

$$ITSN2_i = \sum_{j=1}^{J} \beta_j Y_{ij} \tag{15}$$

Si se le aplica la transformación Likert Equidistante a las puntuaciones de los individuos en los ítems, entonces el calculo de la medida sintética estaría dado por la siguiente ecuación:

$$ITSNE2_i = \sum_{j=1}^{J} \beta_j Y_{ij}* \qquad ; \qquad Y_{ij}* = g(W_{ij} \cdot X_{ij}) \tag{16}$$

Donde $Y_{ij}*$ es una función de la tabla completa o matriz de datos.

Alternativa III (puntajes totales y ACP)

i) Ejecutar los dos primeros pasos planteados en la **Alternativa II (ACP)**. En este caso si un puntaje total no está correlacionado de manera positiva con los demás puntajes, se debe aplicar transformación inversa $(1/x)$. Un puntaje que requiere esta transformación es $(asws\ total/10)$, debido a que es una prueba que originalmente está relacionada de manera negativa con las restantes.

ii) Conformar la matriz de correlación del conjunto de datos y observar si todas las correlaciones son positivas. Esta condición corresponde al concepto llamado factor tamaño (González Rojas, 2014). Contar con correlaciones mayores al 70 % permite obtener resultados más consistentes (ver estudio de simulación de Belalcazar (2023)).

iii) Realizar un análisis de componentes principales para la tabla, apoyándose de la librería en R "*FactoMineR*" en su función *ACP()*.

vi) Extraer el primer vector propio asociado al análisis de componentes principales del paso anterior. Los pesos del vector deben sumar 1, entonces así, cada elemento h de este vector es el peso γ_h.

vii) Calcular la medida sintética para el individuo i que indique el nivel de trastorno del sueño que presenta de la siguiente manera:

$$ITSN3_i = \sum_{h=1}^{H} \gamma_h Z_{ih} \tag{17}$$

5.6.2 Construcción de la medida sintética para padres

Medida sintética Padres

i) Conformar la matriz de correlación del conjunto de datos y observar si todas las correlaciones son positivas. Esta condición corresponde al concepto llamado factor tamaño (González Rojas, 2014). Contar con correlaciones mayores al 70 % permite obtener resultados más consistentes (ver estudio de simulación de Belalcazar (2023)).

iii) Realizar un análisis de componentes principales para la tabla, apoyándose de la librería en R "*FactoMineR*" en su función *ACP()*.

v) Extraer el primer vector propio asociado al análisis de componentes principales. Se debe ponderar de forma que los elementos del vector sumen 1, entonces, cada elemento l de este vector es el peso τ_l.

vii) Calcule la medida sintética para el individuo i que indique la higiene del sueño que presenta el niño según el padre, de la siguiente manera:

$$ITSP_i = \sum_{l=1}^{L} \tau_l U_{il} \tag{18}$$

viii) En último lugar defina los intervalos generales (a, b) en los cuales se puede mover la medida sintética, para ello, deberá asignar los valores mínimos y máximos de la escala tipo likert de los ítems U_{il}.

5.7 Creación de las escalas

Con el fin práctico de identificar los niveles de trastorno del sueño para la muestra y en consecuencia, para un individuo por fuera de la muestra de estudio, se escalaron de cero a cien las médidas sínteticas construidas, de manera que:

$$Escala_{ij} = \frac{Medida_{ij} - min(Medida_j)}{max(Medida_j) - min(Medida_j)} * 100 \tag{19}$$

Donde:

- $Medida_{ij}$: es el valor que toma cualquiera de las medidas j , anteriormente definidas $(ITSN1, ITSN2, ITSN3, ITSP)$ o sus variaciones respectivas para la transformación likert equidistante $(ITSNE1, ITSNE2)$, para el individuo i.
- $min(Medida_j)$: es el mínimo valor que toma la medida j evaluada en todos los individuos.
- $max(Medida_j)$: es máximo valor que toma la medida j evaluada en todos los individuos.

5.8 Modelo Estadístico

Considerando que el plantemiento del problema de investigación está enfocado en cómo construir una medida sintética para evaluar el nivel de presencia de trastornos del sueño en escolares y adolescentes, partiendo del modelo reflectivo de variables latentes, se realizó un modelo estadístico que permita modelar la escala como una variable aleatoria. Por tanto, se establecen los siguientes elementos:

- **Proceso Generador de Datos**: Aquellas condiciones psicológicas, neurológicas, de hábito del sueño, genéticas o de contexto social y familiar que puedan generar que el trastorno del sueño se presente a un nivel específico.

- **Experimento Estadístico**: Tomar un registro de un escolar-adolescente y observar el puntaje que alcanzó en la escala.

- **Evento de Interés y Variable Aletoria**: Debido a que la naturaleza continua de la variable, el evento de interés y la variable aleatoria están relacionados de manera intrínseca. Así pues, se define la variable y el intervalo en los cuales toma valores de la siguiente manera:

  Sea M el nivel del trastorno del sueño registrado para un escolar-adolescente una vez se le aplica la escala.

  Se asume que el comportamiento natural de la variable aleatoria M puede ser modelado usando la distribución de probabilidad $Beta\ (\alpha, \beta)$, entonces;

  $$M \sim Beta(\alpha, \beta) \tag{20}$$

  Para valores de $0 < M < 1$.

- **Función de Densidad de Probabilidad**:

  $$f_M(m) = \frac{m^{\alpha-1}(1-m)^{\beta-1}}{B(\alpha, \beta)} \tag{21}$$

  Donde $B(\alpha, \beta)$ es la función beta y se define para $\alpha, \beta > 0$ como

  $$B(\alpha, \beta) = \int_0^1 m^{\alpha-1}(1-m)^{\beta-1}\, dm \tag{22}$$

  Se debe tener en cuenta que por motivos práticos la $Escala_{ij}$ toma valores de 0 a 100, por tanto modelar la variable con este modelo implica dividir sus valores por 100.

5.9 Correlación entre medidas sintéticas y puntajes validados

Para determinar qué medida estaba más correlacionada con los puntajes validados de las pruebas, se observaron los coeficientes de correlaciones ρ de Spearman entre las escalas formadas por las alternativas I, II y III, incluyendo la escala para padres, con el puntaje medio derivado de los resultados globales de cada uno de los instrumentos.

5.10 Evaluación del Desempeño de las Medidas Sintéticas

Para la evaluación del desempeño de las medidas sintéticas, se realizó el cálculo del PECM. El proceso desarrollado para determinar el Pseudo Error Cuadrático Medio de los estimadores de los pseudo-parámetros (que representan la expresión media del trastorno del sueño en escolares y adolescentes) se puede observar en la Figura 5.

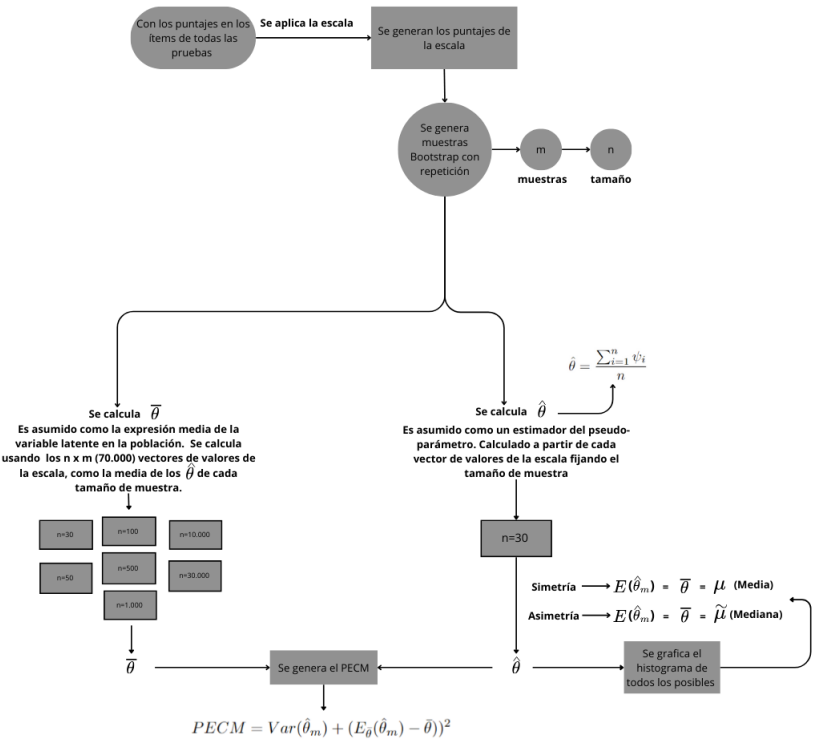

Figura 5 *Proceso para aproximar el valor del Pseudo Error Cuadrático Medio de los estimadores para variables latentes*

Tal como se observa en la Figura 5 se empleó el método de remuestreo Bootstrap para calcular el PECM. Para ello, se generaron muestras con reemplazo para el puntaje de la escala en n individuos, replicando el proceso diez mil veces. Se tomaron distintos tamaños de muestra n, con la finalidad de evaluar el PECM cuando el tamaño de muestra aumenta. Fue necesario evaluar la simetría de la distribución empírica de los $\hat{\theta}$ para definir así, una medida de tendencia central (media o mediana) como parámetro $\bar{\theta}$.

Se estableció como criterio para seleccionar la mejor medida que el Pseudo Error Cuadrático Medio (PECM) sea mínimo. Por lo tanto, la medida que presente consistentemente un menor sesgo y varianza, independientemente del tamaño de la muestra, será la que mejor se adapte a los objetivos.

5.11 Clasificación de los individuos de la muestra según sus niveles de trastorno del sueño

Posterior a la selección de la mejor medida sintética, considerando el Pseudo Error Cuadrático Medio más bajo, se preestablecieron tres niveles para clasificar a los escolares y adolescentes en función del puntaje alcanzado; leve, moderado y severo. Para esto, se utilizó el algoritmo K-means para agrupar los individuos en un nivel específico y posteriormente se establecieron los puntos de corte, los cuales definen bajo qué puntajes se pertenece a los niveles leve, moderado o severo de trastorno del sueño.

Los pasos de la metodología para la clasificación y definión de puntos de corte para diagnosticar el nivel de presencia de trastorno del sueño en escolares y adolescentes es el siguiente:

i) Identificar el número de clúster óptimos, seleccionando el menor número de agrupaciones de individuos similares, que a su vez minimicen la suma de cuadrados totales.
ii) Generar muestras bootstrap de los puntajes de la medida sintética reescalada.
iii) Inicializar el algoritmo K-means teniendo en cuenta el número de clúster seleccionado y que debe ejecutarse 10 veces en diferentes posiciones iniciales.
iv) Extraer los valores mínimos y máximos para cada clúster.
v) Replicar diez mil veces el proceso a partir del segundo paso y extraer la media de los mínimos y máximos de cada clúster resultante. Estas medias serán los puntos de corte para cada uno de las clasificaciones definidas previamente.

Finalmente, se define que un puntaje igual a cero indica que el escolar o adolescente no presenta trastorno del sueño.

5.12 Construcción de tablas de contingencia para evaluar la dependencia entre los niveles de trastorno del sueño y características de los individuos

Como último paso de la metodología, se crearon tablas cruzadas para la muestra de escolares y adolescentes. Para ello, se tuvo en cuenta las frecuencias observadas de los individuos que presentaban cierta categoría en la variable sociodemográfica o de salud, y a la vez eran clasificados es un determinado nivel de trastorno del sueño.

A estas tablas se les aplicó la prueba Chi-Cuadrado para asociación. El umbral usado para determinar si la relación entre una variable y otra era significativa es que el valor-p sea menor que el nivel de significancia $\alpha = 0{,}05$.

6 Resultados

En este capítulo se presentan los resultados generados a partir de la apliación de la metodología definida en el capítulo anterior. En primer lugar, se realiza un análisis exploratorio de los datos, el cual consta de un análisis de las características sociodemográficas y de salud de los escolares y adolescentes, además de un análisis de las respuestas registradas en cada uno de los instrumentos aplicados a los mismos.

Posteriormente, se presentan los resultados de cada una de las medidas sintéticas diseñadas a través de las metodologías de análisis multivariado; ACP y AFM. También se realiza la comparación entre las medidas, seleccionando la que cummpla con los criterios establecidos del PECM.

Finalmente, se clasifican los individuos bajo tres niveles; leve, moderado y severo, con base a los puntajes que consiguieron en la medida sintética reescalada.

6.1 Fase preliminar: Análisis Exploratorio de los datos

Al realizar el procesamiento de los datos, se encontró que de los 827 escolares y adolescentes a quienes se les aplicó las diferentes pruebas, 428 (51,7 %) de ellos contaban con la información completa con respecto a las puntuaciones en los diferentes ítems analizados. Con el objetivo de comprender las características sociodemográficas y de salud de los escolares y adolescentes, se analizan las variables contenidas en la tabla 5.

Variable		Frecuencia	Porcentaje
Sexo	Masculino	186	43,5 %
	Femenino	242	56,5 %
Grupo de edad	Escolares	99	23,1 %
	Adolescentes	329	76,9 %
Diagnóstico IMC	Delgadez	4	0,9 %
	Normal	314	73,4 %
	Obesidad	16	3,7 %
	Sobrepeso	91	21,3 %
	No responde	3	0,7 %
Presenta alguna enfermedad	Si	121	28,3 %
	No	271	63,3 %
	No responde	36	8,4 %
Alguien en la familia presenta alguna enfermedad	Si	157	36,7 %
	No	235	54,9 %
	No responde	36	8,4 %
Trastorno gastrointestinal	Si	138	32,2 %
	No	254	59,3 %
	No responde	36	8,4 %

Tabla 5 *Frecuencia de los participantes según variables sociodemográficas y de salud*

En la tabla 5 se puede observar que la proporción entre el sexo femenino y masculino es similar, siendo el primero el más predominante con un 56,5 %.

Con relación a la variable grupedad, el 76,9 % de la muestra son adolescentes, mientras que el porcentaje restante escolares, esto debido a que solo se pudo recolectar información de estudiantes de sexto grado hasta undécimo grado y en estos grados es más frecuente encontrar niños en edad adolescente.

Al observar la variable Diagnóstico IMC, se destaca que la mayoría de escolares y adolescentes se encuentran en una condición normal en cuanto a la relación masa corporal-altura. No obstante, hay un porcentaje bastante importante de niños que sufren de sobrepeso (21,3 %), sumado a los casos extremos como la delgadez y la obesidad, con un 0,9 % y un 3,7 % respectivamente.

Con relación a las enfermedades, la mayoría de niños informan no presentar ninguna (63,3 %) y también afirman que algún integrante de su familia presenta alguna enfermedad (54,9 %). Además de las anteriores variables, también se considera el diagnóstico de trastorno gastrointestinal conseguido por medio de la aplicación del Cuestionario de Síntomas Gastrointestinales Pediátricos Roma IV. A partir de esto, se puede ver que la prevalencia asociada a presentar por lo menos un trastorno gastrointestinal en la muestra de niños en edad escolar y adolescente es del 35,2 %.

Para concluir con el análisis exploratorio, se procede a realizar el análisis de las respuesta que cada uno de los escolares y adolescentes plasmó en los diferentes ítems de los instrumentos (PISI, ESS, PROMIS SD, PROMIS SRI, ASWS, PROMIS ANXIETY y PROMIS DEPRESSION). Para ello, se presentan las frecuencias y las respectivas proporciones de las categorías de cada uno de los ítem de las pruebas en la tabla 6.

PISI												
Ítem	1		2		3		4		5		6	
Punt.	frec.	%	frec.	%	frec.	%	frec.	%	frec.	%	frec.	%
0	97	22,7	174	40,7	164	38,3	226	52,8	80	18,7	28	6,5
1	136	31,8	104	24,3	177	41,4	106	24,8	115	26,9	87	20,3
2	66	15,4	60	14,0	46	10,7	38	8,9	55	12,9	108	25,2
3	37	8,6	27	6,3	17	4,0	16	3,7	47	11,0	100	23,4
4	29	6,8	30	7,0	13	3,0	15	3,5	45	10,5	87	20,3
5	63	14,7	33	7,7	11	2,6	27	6,3	86	20,1	18	4,2

ESS																
Ítem	1		2		3		4		5		6		7		8	
Punt.	frec.	%	frec.	%	frec.	%	frec.	%	frec.	%	frec.	%	frec.	%	frec.	%
0	203	47,4	103	24,1	208	48,6	112	26,2	61	14,3	373	87,1	188	43,9	340	79,4
1	143	33,4	174	40,7	118	27,6	140	32,7	132	30,8	35	8,2	123	28,7	63	14,7
2	40	9,3	82	19,2	56	13,1	70	16,4	74	17,3	11	2,6	59	13,8	17	4,0
3	42	9,8	69	16,1	46	10,7	106	24,8	161	37,6	9	2,1	58	13,6	8	1,9

PROMIS SD																
Ítem	1		2		3		4		5		6		7		8	
Punt.	frec.	%	frec.	%	frec.	%	frec.	%	frec.	%	frec.	%	frec.	%	frec.	%
1	140	32,7	175	40,9	192	44,9	166	38,8	130	30,4	231	54,0	217	50,7	147	34,3
2	92	21,5	125	29,2	93	21,7	105	24,5	107	25,0	68	15,9	80	18,7	96	22,4
3	120	28,0	98	22,9	90	21,0	97	22,7	96	22,4	68	15,9	88	20,6	95	22,2
4	46	10,7	0	0,0	31	7,2	35	8,2	56	13,1	35	8,2	24	5,6	40	9,3
5	30	7,0	30	7,0	22	5,1	25	5,8	39	9,1	26	6,1	19	4,4	50	11,7

PROMIS SRI																
Ítem	1		2		3		4		5		6		7		8	
Punt.	frec.	%	frec.	%	frec.	%	frec.	%	frec.	%	frec.	%	frec.	%	frec.	%
1	51	11,9	120	28,0	155	36,2	171	40,0	186	43,5	263	61,4	260	60,7	178	41,6
2	62	14,5	88	20,6	87	20,3	106	24,8	110	25,7	80	18,7	88	20,6	83	19,4
3	150	35,0	134	31,3	112	26,2	91	21,3	81	18,9	60	14,0	55	12,9	76	17,8
4	96	22,4	59	13,8	52	12,1	46	10,7	35	8,2	15	3,5	20	4,7	40	9,3
5	69	16,1	27	6,3	22	5,1	14	3,3	16	3,7	10	2,3	5	1,2	51	11,9

ASWS																				
Ítem	1		2		3		4		5		6		7		8		9		10	
Punt.	frec.	%	frec.	%	frec.	%	frec.	%	frec.	%	frec.	%	frec.	%	frec.	%	frec.	%	frec.	%
1	66	15,4	77	18,0	142	33,2	239	55,8	237	55,4	222	51,9	236	55,1	285	66,6	63	14,7	54	12,6
2	101	23,6	71	16,6	110	25,7	84	19,6	59	13,8	99	23,1	91	21,3	63	14,7	50	11,7	39	9,1
3	80	18,7	40	9,3	66	15,4	34	7,9	35	8,2	46	10,7	48	11,2	26	6,1	49	11,4	43	10,0
4	72	16,8	87	20,3	49	11,4	22	5,1	23	5,4	25	5,8	18	4,2	17	4,0	85	19,9	76	17,8
5	61	14,3	102	23,8	36	8,4	24	5,6	32	7,5	13	3,0	16	3,7	13	3,0	98	22,9	121	28,3
6	48	11,2	51	11,9	25	5,8	25	5,8	42	9,8	23	5,4	19	4,4	24	5,6	83	19,4	95	22,2

PROMIS ANXIETY																
Ítem	1		2		3		4		5		6		7		8	
Punt.	frec.	%	frec.	%	frec.	%	frec.	%	frec.	%	frec.	%	frec.	%	frec.	%
1	217	50,7	133	31,1	203	47,4	137	32,0	253	59,1	231	54,0	222	51,9	281	65,7
2	67	15,7	77	18,0	88	20,6	73	17,1	73	17,1	86	20,1	73	17,1	62	14,5
3	99	23,1	128	29,9	86	20,1	109	25,5	58	13,6	64	15,0	78	18,2	44	10,3
4	29	6,8	56	13,1	31	7,2	63	14,7	29	6,8	24	5,6	34	7,9	21	4,9
5	16	3,7	34	7,9	20	4,7	46	10,7	15	3,5	23	5,4	21	4,9	20	4,7

PROMIS DEPRESSION																
Ítem	1		2		3		4		5		6		7		8	
Punt.	frec.	%	frec.	%	frec.	%	frec.	%	frec.	%	frec.	%	frec.	%	frec.	%
0	189	44,2	166	38,8	183	42,8	145	33,9	196	45,8	139	32,5	195	45,6	211	49,3
1	68	15,9	66	15,4	65	15,2	62	14,5	66	15,4	71	16,6	81	18,9	87	20,3
2	87	20,3	99	23,1	91	21,3	114	26,6	83	19,4	103	24,1	76	17,8	77	18,0
3	49	11,4	57	13,3	48	11,2	61	14,3	53	12,4	68	15,9	47	11,0	30	7,0
4	35	8,2	40	9,3	41	9,6	46	10,7	30	7,0	47	11,0	29	6,8	23	5,4

Tabla 6 *Distribución de las respuestas a los ítem de cada uno de los instrumentos evaluados en los escolares y adolescentes*

Con base en la tabla 6, se pudo evidenciar algunas proporciones que dan información importante sobre los niños participantes. El 20,1 % de los escolares y adolescentes se sienten siempre con sueño durante el día (PISI, ítem 6), el 37,6 % afirman que en el último mes, se han quedado dormidos al sentarse en un salón de clases por la mañana (ESS, ítem 5), el 11,7 % siempre daban vueltas en la cama por no poder dormir (PROMIS SD, ítem 8). Además de esto, el 22,4 % de los niños casi siempre tienen sueño durante el día (PROMIS SRI, ítem 1) y 22,2 % de ellos nunca se despertaban por la mañana y se sentían descansados o alertas (ASWS, ítem 10). Finalmente, con respecto a su estado de ánimo, el 14,7 % casi siempre se sentían preocupados y el 15,9 % casi siempre se sentían tristes (PROMIS ANXIETY, ítem 4 y PROMIS DEPRESSION, ítem 6).

6.2 Resultados del Primer Objetivo: Construcción de la Medida Sintética para escolares-adolescentes (alternativas I,II y III)

En esta sección del capítulo se presentan los resultados del análisis multivariado y las respectivas medidas sintéticas diseñadas a partir de los pesos asociados a los vectores propios resultantes. Además, se presenta un breve análisis sobre los ítems o puntajes de ítems en función de sus pesos.

6.2.1 Construcción de la medida sintética para escolares-adolescentes por medio de AFM

En la tabla 7 se puede apreciar que la primer dimensión explica el mayor porcentaje de la variabilidad de los datos con un 29,85 %. Como el objetivo de este trabajo no es reducir el número de factores, sino capturar la mayor información posible en la primer dimensión, para así definir los pesos asociados a los ítems en la ecuación de la medida sintética, no se hace tan indispensable considerar más dimensiones que aumenten ese porcentaje de varianza.

Dim.	Valor propio	% Varianza	% Acumulado	Dim.	Valor propio	% Varianza	% Acumulado
1	4.37	29.85	29.85	29	0.11	0.79	88.87
2	0.99	6.80	36.65	30	0.11	0.74	89.61
3	0.75	5.14	41.79	31	0.10	0.70	90.32
4	0.60	4.08	45.87	32	0.10	0.66	90.98
5	0.52	3.53	49.40	33	0.09	0.64	91.62
6	0.50	3.39	52.78	34	0.09	0.61	92.23
7	0.43	2.95	55.73	35	0.09	0.60	92.83
8	0.39	2.64	58.37	36	0.08	0.55	93.38
9	0.34	2.33	60.70	37	0.08	0.52	93.90
10	0.33	2.29	62.99	38	0.07	0.49	94.39
11	0.32	2.18	65.17	39	0.07	0.46	94.85
12	0.31	2.15	67.32	40	0.07	0.45	95.30
13	0.30	2.02	69.34	41	0.06	0.43	95.73
14	0.29	1.99	71.33	42	0.06	0.42	96.15
15	0.24	1.67	73.00	43	0.06	0.38	96.54
16	0.24	1.64	74.64	44	0.05	0.37	96.91
17	0.22	1.50	76.14	45	0.05	0.35	97.26
18	0.21	1.45	77.58	46	0.05	0.34	97.60
19	0.19	1.33	78.91	47	0.05	0.31	97.91
20	0.18	1.24	80.15	48	0.04	0.30	98.21
21	0.17	1.19	81.34	49	0.04	0.28	98.49
22	0.17	1.18	82.52	50	0.04	0.28	98.76
23	0.16	1.08	83.60	51	0.04	0.25	99.01
24	0.14	0.99	84.58	52	0.04	0.24	99.25
25	0.14	0.95	85.53	53	0.03	0.22	99.48
26	0.13	0.89	86.42	54	0.03	0.20	99.68
27	0.12	0.85	87.27	55	0.02	0.16	99.84
28	0.12	0.82	88.09	56	0.02	0.16	100.00

Tabla 7 *Resumen de los Valores Propios Resultante del AFM para la alternativa I*

Tal como se observa en la tabla 8, la metodología aplicada en la construcción de la medida sintética por medio de la primer alternativa, dió como resultado que los instrumentos que más contribuyen en la explicación de la primer dimensión, son aquellos que miden el imsomnio, el deterioro relacionado con el sueño y la alteración del sueño con inercias de 0,73 (16,7 %), 0,71 (16,2 %) y 0,70 (16 %), respectivamente. No obstante, todos los tratornos medidos contribuyen balanceadamente a la inercia total de la primer dimensión.

Trastorno	Dim.1	Dim.2	Dim.3
Insomnio	0.73	0.23	0.03
Somnolencia	0.42	0.25	0.25
Alteración	0.70	0.16	0.02
Deterioro	0.71	0.09	0.01
Sueño y vigilia	0.66	0.21	0.04
Ansiedad	0.60	0.02	0.16
Depresión	0.54	0.04	0.25
Inercia total	4.37	0.99	0.75

Tabla 8 *Descomposición de la Inercia Total del AFM para la alternativa I*

Los resultados de la tabla 9 muestran que todos los ítems están altamente asociados a la primer dimensión, pues sus coordenadas son positivas, lo cual indica que todos son claves en la comprensión de la variable latente (trastorno del sueño en escolares y asolescentes). También se puede ver que los 5 ítems que más contribuyen a la primer dimensión son: pisi 2, pisi 5, promis sd 1, asws 7 y promis sd 4, con valores de 4,04, 3,55, 2,72, 2,70 y 2,66 respectivamente. Estos ítems están asociados a los instrumentos que miden insomnio, alteración del sueño, sueño y vigilia.

	Coordenadas		Contribuciones			Coordenadas		Contribuciones	
Ítem	Dim 1	Dim 2	Dim 1	Dim 2	Ítem	Dim 1	Dim 2	Dim 1	Dim 2
pisi 1	0.56	-0.29	2.62	3.21	promis sri 7	0.57	0.07	1.59	0.11
pisi 2	0.69	-0.26	4.04	2.45	promis sri 8	0.61	0.20	1.78	0.83
pisi 3	0.51	-0.31	2.22	3.56	asws 1	0.44	-0.06	1.19	0.10
pisi 4	0.55	-0.49	2.59	9.04	asws 2	0.20	-0.07	0.26	0.15
pisi 5	0.65	0.34	3.55	4.40	asws 3	0.57	-0.01	2.01	0.00
pisi 6	0.46	0.16	1.79	0.91	asws 4	0.58	-0.29	2.11	2.29
ess 1	0.31	0.30	0.92	3.64	asws 5	0.56	-0.25	1.95	1.67
ess 2	0.38	0.19	1.36	1.49	asws 6	0.64	-0.47	2.54	6.11
ess 3	0.52	0.22	2.50	1.94	asws 7	0.66	-0.37	2.70	3.81
ess 4	0.24	0.13	0.54	0.70	asws 8	0.53	-0.40	1.76	4.36
ess 5	0.42	0.46	1.68	8.90	asws 9	0.25	0.23	0.39	1.48
ess 6	0.22	0.09	0.45	0.36	asws 10	0.20	0.18	0.25	0.87
ess 7	0.39	0.42	1.44	7.32	promis anxiety 1	0.55	0.10	1.41	0.21
ess 8	0.29	0.09	0.79	0.33	promis anxiety 2	0.62	0.18	1.76	0.66
promis sd 1	0.73	-0.29	2.72	1.86	promis anxiety 3	0.63	0.13	1.81	0.35
promis sd 2	0.13	-0.13	0.08	0.35	promis anxiety 4	0.67	0.17	2.08	0.61
promis sd 3	0.69	-0.26	2.37	1.51	promis anxiety 5	0.66	0.08	1.97	0.12
promis sd 4	0.73	-0.28	2.66	1.75	promis anxiety 6	0.58	0.10	1.57	0.21
promis sd 5	0.67	-0.39	2.27	3.29	promis anxiety 7	0.56	0.09	1.46	0.17
promis sd 6	0.62	-0.20	1.94	0.92	promis anxiety 8	0.60	-0.00	1.64	0.00
promis sd 7	0.60	-0.43	1.83	4.17	promis depression 1	0.64	0.20	1.58	0.67
promis sd 8	0.66	-0.29	2.21	1.85	promis depression 2	0.64	0.17	1.58	0.51
promis sri 1	0.62	0.38	1.84	3.01	promis depression 3	0.62	0.21	1.48	0.78
promis sri 2	0.70	0.23	2.34	1.12	promis depression 4	0.62	0.15	1.50	0.40
promis sri 3	0.69	0.29	2.27	1.82	promis depression 5	0.68	0.13	1.77	0.28
promis sri 4	0.72	0.13	2.49	0.33	promis depression 6	0.66	0.22	1.67	0.83
promis sri 5	0.68	0.24	2.22	1.21	promis depression 7	0.61	0.12	1.44	0.26
promis sri 6	0.59	0.14	1.67	0.43	Promis depression 8	0.59	0.13	1.37	0.27

Tabla 9 *Coordenadas y Contribuciones para las dos Primeras Dimensiones del AFM para la alternativa I*

Formulación de la Medida Sintética (primera alternativa)

Por último, se construyó la medida sintética de la primer alternativa, usando los pesos ponderados generados con el vector Mu, el cual fue conseguido por medio del análisis factorial múltiple. A esta medida se le denominó **ITSN1** y su ecuación está dada por:

$$
\begin{aligned}
ITSN1_i = {} & 0{,}0120x_{i11} + 0{,}0417x_{i21} + 0{,}0367x_{i31} + 0{,}0520x_{i41} + 0{,}0659x_{i51} + 0{,}0842x_{i61} \\
& + 0{,}0004x_{i12} + 0{,}0014x_{i22} + 0{,}0022x_{i32} + 0{,}0013x_{i42} + 0{,}0026x_{i52} + 0{,}0024x_{i62} \\
& + 0{,}0005x_{i72} + 0{,}0003x_{i82} + 0{,}0006x_{i13} + 0{,}0001x_{i23} + 0{,}0004x_{i33} + 0{,}0004x_{i43} \\
& + 0{,}0145x_{i53} + 0{,}0374x_{i63} + 0{,}0432x_{i73} + 0{,}0623x_{i83} + 0{,}0631x_{i14} + 0{,}1280x_{i24} \\
& + 0{,}0009x_{i34} + 0{,}0026x_{i44} + 0{,}0029x_{i54} + 0{,}0033x_{i64} + 0{,}0035x_{i74} + 0{,}0067x_{i84} \\
& + 0{,}0006x_{i15} + 0{,}0002x_{i25} + 0{,}0004x_{i35} + 0{,}0004x_{i45} + 0{,}0003x_{i55} + 0{,}0004x_{i65} \\
& + 0{,}0141x_{i75} + 0{,}0319x_{i85} + 0{,}0178x_{i95} + 0{,}0188x_{i105} + 0{,}0564x_{i16} + 0{,}1135x_{i26} \\
& + 0{,}0008x_{i36} + 0{,}0024x_{i46} + 0{,}0028x_{i56} + 0{,}0033x_{i66} + 0{,}0034x_{i76} + 0{,}0066x_{i86} \\
& + 0{,}0008x_{i17} + 0{,}0006x_{i27} + 0{,}0005x_{i37} + 0{,}0004x_{i47} + 0{,}0004x_{i57} + 0{,}0004x_{i67} \\
& + 0{,}0131x_{i77} + 0{,}0358x_{i87}
\end{aligned} \tag{23}
$$

El rango de valores que puede adoptar la medida sintética ITSN1 se determina por el intervalo **(0,64; 5,01)**.

6.2.2 Construcción de la medida sintética para escolares-adolescentes por medio ACP

La tabla 10 presenta los valores propios asociados a las dimensiones conformadas al realizar el análisis de componentes principales posterior a la reducción de dimensionalidad mediante un ACP individual por instrumento. En esta tabla se puede evidenciar que en la primer dimensión se explica el 61,93 % de la variabilidad de los datos.

Dim.	Valor Prop.	% Varianza	% Varianza Acum.
1	4.34	61.93	61.93
2	0.79	11.24	73.17
3	0.71	10.12	83.30
4	0.38	5.44	88.74
5	0.31	4.40	93.14
6	0.30	4.23	97.37
7	0.18	2.63	100.00

Tabla 10 *Resumen Valores Propios Resultantes del ACP para la alternativa II*

Los datos resumidos en la tabla 11 indican que los nuevos factores generados por medio de la reducción de la dimensionalidad derivada del primer ACP a cada instrumento, están altamente relacionados con la primer dimensión. Se podría entender cada uno de estos factores como una representación de la información más relevante de cada uno de los instrumentos que miden los diferentes tipos de trastorno definidos en la primer columna de la tabla.

	Coordenadas		**Contribuciones**	
Factor-Trastorno	Dim 1	Dim 2	Dim 1	Dim 2
Insomnio	0.86	-0.27	16.92	9.15
Somnolencia	0.64	0.56	9.46	39.69
Alteración	0.84	-0.38	16.23	18.71
Deterioro	0.84	0.24	16.16	7.62
Sueño y vigilia	0.81	-0.36	15.06	16.19
Ansiedad	0.77	0.20	13.64	4.84
Depresión	0.74	0.17	12.53	3.79

Tabla 11 *Coordenadas y Contribuciones para las dos Primeras Dimensiones del ACP para la alternativa II*

Otra información que se puede destacar de la tabla 11 es que de los nuevos factores, los tres que más contribuyen a la primer dimensión son los que representan la información más relevante de insomnio, alteración del sueño y deterioro relacionado con el sueño.

Formulación de la Medida Sintética (segunda alternativa)

Posteriormente, se construyó la segunda medida sintética, usando los pesos ponderados derivados del primer vector propio, el cual fue conseguido por medio del análisis de componentes principales. A esta medida se le denominó **ITSN2** y su ecuación está dada por:

$$ITSN2_i = 0{,}1561y_{i1} + 0{,}1167y_{i2} + 0{,}1529y_{i3} + 0{,}1526y_{i4} + 0{,}1473y_{i5} + 0{,}1402y_{i6} + 0{,}1343y_{i7} \quad (24)$$

El rango de valores que puede adoptar la medida sintética ITSN2 se determina por el intervalo **(1,67; 12,45)**.

6.2.3 Construcción de la medida sintética para escolares-adolescentes con base en los puntajes totales de las pruebas, haciendo uso de ACP

Tal como se observa en la tabla 12, la primer dimensión explica más de la mitad de la variabilidad de los datos, puntualmente el 59,29 %, mientras que el resto pasan a explicar una menor proporción. Este alto porcentaje de varianza explicada genera una buena representación del trastorno del sueño.

Dim.	Valor Prop.	% Varianza	% Varianza Acum.
1	4.15	59.29	59.29
2	0.77	11.04	70.33
3	0.71	10.08	80.41
4	0.48	6.79	87.20
5	0.37	5.30	92.51
6	0.30	4.35	96.86
7	0.22	3.14	100.00

Tabla 12 *Resumen Valores Propios Resultantes del ACP para la alternativa III*

Observando la tabla 13 se puede analizar que los puntajes globales de las pruebas están altamente relacionados con la primer componente. Además, los puntajes globales que más aportan a la construcción de la primer dimensión son: el puntaje alcanzado en el instrumento pisi, en promis sri y en promis sd.

	Coordenadas		**Contribuciones**	
Puntaje Trastorno	Dim 1	Dim 2	Dim 1	Dim 2
pisi total	0.85	-0.27	17.52	9.25
ess total	0.66	0.46	10.50	27.27
T-score promis sd total	0.79	-0.37	15.14	17.97
T-score promis sri total	0.84	0.17	17.15	3.89
asws total/10	0.71	-0.44	12.05	24.84
T-score promis anxiety total	0.76	0.28	14.01	10.11
T score promis depression	0.75	0.23	13.64	6.67

Tabla 13 *Coordenadas y Contribuciones para las dos Primeras Dimensiones del ACP para la alternativa III*

Formulación de la Medida Sintética (tercera alternativa)

En último lugar, se construyó la medida sintética asociada a los puntajes globales, de igual forma, se usaron los valores del primer vector propio conseguido por medio del análisis de componentes principales y se ponderaron para generar los pesos correspondiente al modelo de variables latentes. A esta medida sintética se le denominó **ITSN2** y su ecuación está dada por:

$$ITSN3_i = 0{,}1587z_{i1} + 0{,}1229z_{i2} + 0{,}1476z_{i3} + 0{,}1571z_{i4} + 0{,}1317z_{i5} + 0{,}1420z_{i6} + 0{,}1401z_{i7} \tag{25}$$

El rango de valores que puede adoptar la medida sintética ITSN3 se determina por el intervalo **(21,47; 56,52)**.

6.2.4 Construcción de la medida sintética para escolares-adolescentes por medio de AFM y Likert Equidistante

En la tabla 14 se muestra el resumen de los valores propios para cada una de las dimensiones resultantes del análisis factorial múltiple con las puntuaciones transformadas en puntuaciones monotónicos continuas. En esta tabla se observa que la primer dimensión explica el 25,92 % de la variabilidad de los datos. Sin embargo, es importante considerar que la segunda dimensión explica un 12,43 % de esa variabilidad. A diferencia de la medida sin realizar la transformación de los puntajes, esta explica un mayor porcentaje de varianza en un número menor de dimensiones.

Dim.	Valor propio	% Varianza	% Acumulado	Dim.	Valor propio	% Varianza	% Acumulado
1	3.45	25.92	25.92	29	0.07	0.49	93.54
2	1.66	12.43	38.35	30	0.06	0.45	93.99
3	1.01	7.61	45.97	31	0.06	0.42	94.41
4	0.70	5.29	51.26	32	0.05	0.40	94.81
5	0.50	3.73	54.98	33	0.05	0.37	95.18
6	0.48	3.63	58.62	34	0.05	0.35	95.53
7	0.42	3.17	61.78	35	0.05	0.35	95.88
8	0.40	2.98	64.76	36	0.04	0.33	96.21
9	0.39	2.94	67.71	37	0.04	0.32	96.54
10	0.35	2.62	70.33	38	0.04	0.29	96.83
11	0.31	2.31	72.65	39	0.04	0.28	97.12
12	0.29	2.17	74.82	40	0.04	0.27	97.39
14	0.26	1.96	78.80	42	0.03	0.24	97.88
15	0.23	1.72	80.52	43	0.03	0.21	98.09
16	0.20	1.53	82.05	44	0.03	0.20	98.30
17	0.19	1.40	83.44	45	0.03	0.20	98.50
18	0.17	1.31	84.75	46	0.03	0.19	98.68
19	0.16	1.20	85.95	47	0.02	0.17	98.86
20	0.16	1.18	87.13	48	0.02	0.17	99.02
21	0.15	1.09	88.22	49	0.02	0.16	99.19
22	0.12	0.92	89.14	50	0.02	0.15	99.33
23	0.11	0.86	90.00	51	0.02	0.14	99.47
24	0.10	0.77	90.77	52	0.02	0.13	99.60
25	0.09	0.67	91.44	53	0.02	0.12	99.72
26	0.07	0.56	92.00	54	0.01	0.10	99.82
27	0.07	0.54	92.54	55	0.01	0.10	99.92
28	0.07	0.51	93.05	56	0.01	0.08	100.00

Tabla 14 *Resumen de los Valores Propios Resultante del AFM para la alternativa I con datos transformados (Likert Equidistante)*

Se puede apreciar en la tabla 14 los tres instrumentos que más contribuyen a la explicación de primer dimensión son aquellos que miden trastornos como deterioro relacionado con el sueño, alteración del sueño y ansiedad con inercias de 0,74 (21,4 %), 0,73 (21,1 %) y 0,72 (20,9 %), respectivamente. Los problemas de ánimo y funcionamiento en función del tiempo de descanso, los problemas al momento de conciliar el sueño y los problemas de ansiedad y preocupación ante las situaciones son los aspectos medidos en los instrumentos cuyos puntajes logran las mayores inercias.

Trastorno	Dim.1	Dim.2	Dim.3
Insomnio	0.26	0.60	0.13
Somnolencia	0.15	0.08	0.67
Alteración	0.73	0.07	0.03
Deterioro	0.74	0.11	0.04
Sueño y vigilia	0.19	0.65	0.13
Ansiedad	0.72	0.15	0.01
Depresión	0.66	0.00	0.01
Inercia total	3.45	1.66	1.01

Tabla 15 *Descomposición de la Inercia Total del AFM para la alternativa I con datos transformados (Likert Equidistante)*

Al igual que en los resultados de la medida sintética con los puntajes sin transformar, se presenta en la tabla 16 las coordenadas y contribuciones resultantes del análisis multivariado. Se observa que todos los ítems son clave para representar la variable latente, puesto que todas están altamente relacionados con la primer dimensión al tener coordenadas positivas. Además, los ítems promis sri 3, promis sd 3 son los primeros tres en cuanto a las constribuciones, con valores de 3,40, 3,34 y 3,33, respectivamente.

	Coordenadas		Contribuciones			Coordenadas		Contribuciones	
Ítem	Dim 1	Dim 2	Dim 1	Dim 2	Ítem	Dim 1	Dim 2	Dim 1	Dim 2
pisi 1	0.40	0.53	1.66	6.08	promis sri 7	0.71	-0.34	2.80	1.31
pisi 2	0.44	0.69	2.04	10.49	promis sri 8	0.76	-0.23	3.25	0.63
pisi 3	0.31	0.55	1.04	6.72	asws 1	0.33	0.29	0.66	1.05
pisi 4	0.31	0.66	1.00	9.45	asws 2	0.08	0.04	0.04	0.02
pisi 5	0.39	0.39	1.58	3.28	asws 3	0.37	0.69	0.80	5.97
pisi 6	0.14	-0.02	0.21	0.01	asws 4	0.39	0.68	0.89	5.77
ess 1	0.20	0.24	0.51	1.49	asws 5	0.38	0.63	0.88	4.99
ess 2	0.20	0.10	0.50	0.26	asws 6	0.36	0.75	0.77	6.98
ess 3	0.29	0.19	1.05	0.90	asws 7	0.37	0.75	0.80	7.04
ess 4	0.15	0.05	0.26	0.06	asws 8	0.31	0.66	0.59	5.44
ess 5	0.19	-0.00	0.46	0.00	asws 9	0.08	-0.20	0.04	0.51
ess 6	0.11	0.20	0.16	1.00	asws 10	0.03	-0.32	0.01	1.29
ess 7	0.27	0.21	0.89	1.11	promis anxiety 1	0.74	-0.39	2.59	1.53
ess 8	0.18	0.09	0.42	0.22	promis anxiety 2	0.71	-0.22	2.43	0.47
promis sd 1	0.75	-0.13	3.12	0.21	promis anxiety 3	0.76	-0.39	2.74	1.48
promis sd 2	0.08	0.17	0.04	0.35	promis anxiety 4	0.77	-0.17	2.86	0.30
promis sd 3	0.78	-0.27	3.34	0.82	promis anxiety 5	0.77	-0.34	2.84	1.13
promis sd 4	0.77	-0.24	3.25	0.66	promis anxiety 6	0.75	-0.33	2.67	1.10
promis sd 5	0.70	-0.07	2.69	0.05	promis anxiety 7	0.71	-0.38	2.42	1.42
promis sd 6	0.77	-0.31	3.31	1.11	promis anxiety 8	0.71	-0.38	2.38	1.47
promis sd 7	0.69	-0.28	2.65	0.88	promis depression 1	0.72	-0.06	2.41	0.03
promis sd 8	0.70	-0.15	2.76	0.27	promis depression 2	0.72	-0.03	2.43	0.01
promis sri 1	0.13	0.27	0.10	0.86	promis depression 3	0.71	-0.04	2.36	0.01
promis sri 2	0.68	-0.10	2.60	0.11	promis depression 4	0.72	-0.00	2.42	0.00
promis sri 3	0.78	-0.17	3.40	0.33	promis depression 5	0.76	-0.11	2.70	0.11
promis sri 4	0.77	-0.23	3.33	0.61	promis depression 6	0.73	-0.01	2.50	0.00
promis sri 5	0.76	-0.28	3.19	0.89	promis depression 7	0.68	-0.08	2.17	0.06
promis sri 6	0.71	-0.38	2.82	1.65	promis depression 8	0.68	-0.05	2.19	0.02

Tabla 16 *Coordenadas y Contribuciones para las dos Primeras Dimensiones del AFM para la alternativa I con datos transformados (Likert Equidistante)*

Formulación de la Medida Sintética (primera alternativa) con datos transformados

Finalmente, se construyó la medida sintética usando los pesos ponderados generados con el vector Mu y habiendo transformado los puntajes de cada uno de los ítems en puntuaciones monotónicas continuas. En este caso, a esta medida sintética se le denominó **ITSNE1** y su ecuación está dada por:

$$\begin{aligned} ITSNE1_i = {} & 0{,}0103x^*_{i11} + 0{,}0286x^*_{i21} + 0{,}0327x^*_{i31} + 0{,}0367x^*_{i41} + 0{,}0504x^*_{i51} + 0{,}0299x^*_{i61} \\ & + 0{,}0003x^*_{i12} + 0{,}0008x^*_{i22} + 0{,}0018x^*_{i32} + 0{,}0010x^*_{i42} + 0{,}0015x^*_{i52} + 0{,}0014x^*_{i62} \\ & + 0{,}0004x^*_{i72} + 0{,}0002x^*_{i82} + 0{,}0007x^*_{i13} + 0{,}0001x^*_{i23} + 0{,}0006x^*_{i33} + 0{,}0005x^*_{i43} \\ & + 0{,}0180x^*_{i53} + 0{,}0502x^*_{i63} + 0{,}0722x^*_{i73} + 0{,}0842x^*_{i83} + 0{,}0172x^*_{i14} + 0{,}1445x^*_{i24} \\ & + 0{,}0012x^*_{i34} + 0{,}0030x^*_{i44} + 0{,}0047x^*_{i54} + 0{,}0051x^*_{i64} + 0{,}0055x^*_{i74} + 0{,}0097x^*_{i84} \\ & + 0{,}0005x^*_{i15} + 0{,}0001x^*_{i25} + 0{,}0003x^*_{i35} + 0{,}0003x^*_{i45} + 0{,}0003x^*_{i55} + 0{,}0003x^*_{i65} \\ & + 0{,}0095x^*_{i75} + 0{,}0203x^*_{i85} + 0{,}0087x^*_{i95} + 0{,}0039x^*_{i105} + 0{,}0957x^*_{i16} + 0{,}1509x^*_{i26} \\ & + 0{,}0009x^*_{i36} + 0{,}0023x^*_{i46} + 0{,}0036x^*_{i56} + 0{,}0041x^*_{i66} + 0{,}0042x^*_{i76} + 0{,}0090x^*_{i86} \\ & + 0{,}0008x^*_{i17} + 0{,}0006x^*_{i27} + 0{,}0005x^*_{i37} + 0{,}0005x^*_{i47} + 0{,}0004x^*_{i57} + 0{,}0004x^*_{i67} \\ & + 0{,}0177x^*_{i77} + 0{,}0445x^*_{i87} \end{aligned} \tag{26}$$

El rango de valores que puede adoptar la medida sintética ITSNE1 se determina por el intervalo **(0,05; 1,20)**.

6.2.5 Construcción de la medida sintética para escolares-adolescentes por medio ACP y Likert Equidistante

En tabla 17 se presentan los valores propios correspondientes a las dimensiones obtenidas tras realizar el análisis de componentes principales, aplicado después de reducir la dimensionalidad a través de un ACP individual para cada instrumento. Lo más relevante es que en la primera dimensión se logra explicar el 48,77 % de la variabilidad total, lo que representa una disminución de 13,16 % comparado con lo que se lograba explicar en el análisis sin transformar los datos. Sin embargo, se observó un aumento en la varianza explicada por la segunda dimensión, con un incremento del 11,54 %.

Dim.	Valor Prop.	% Varianza	% Varianza Acum.
1	3.41	48.77	48.77
2	1.59	22.78	71.55
3	0.89	12.69	84.25
4	0.45	6.36	90.61
5	0.28	3.96	94.57
6	0.21	3.05	97.62
7	0.17	2.38	100.00

Tabla 17 *Resumen Valores Propios Resultantes del ACP para la alternativa II con datos transformados (Likert Equidistante)*

Es importante analizar el porcentaje de varianza explicada por la segunda dimensión, ya que si se compara con los resultados de otras medidas, es la que mayor porcentaje presenta. Esto puede indicar que es apropiado fijarse

de manera complementaria en los factores que aportan a esta dimensión, no obstante no es esencial para la representación del constructo latente (trastorno del sueño).

	Coordenadas		Contribuciones	
Factor-Trastorno	Dim 1	Dim 2	Dim 1	Dim 2
Insomnio	0.49	0.78	7.09	38.39
Somnolencia	0.36	0.24	3.76	3.52
Alteración	0.86	-0.24	21.79	3.55
Deterioro	0.86	-0.28	21.70	4.96
Sueño y vigilia	0.41	0.81	4.89	40.84
Ansiedad	0.85	-0.37	21.40	8.62
Depresión	0.81	-0.04	19.36	0.12

Tabla 18 *Coordenadas y Contribuciones para las dos Primeras Dimensiones del ACP para la alternativa II con datos transformados (Likert Equidistante)*

Como se muestra en la tabla 18, los nuevos factores que más contribuyen a la primera dimensión son los que representan la alteración del sueño, el deterioro relacionado con el sueño la ansiedad y la depresión. Además, todos estos factores están altamente relacionados con la primer dimensión, representando adecuadamente el trastorno de sueño como constructo.

Formulación de la Medida Sintética (segunda alternativa) con datos transformados

Una vez se obtuvieron los pesos ponderados derivados del primer vector propio, el cual fue conseguido por medio del análisis de componentes principales, se construyó la medida sintética con base a los puntajes transformados. A esta medida se le denominó **ITSNE2** y su ecuación está dada por:

$$ITSNE2_i = 0{,}1058y^*_{i1} + 0{,}0771y^*_{i2} + 0{,}1855y^*_{i3} + 0{,}1851y^*_{i4} + 0{,}0879y^*_{i5} + 0{,}1838y^*_{i6} + 0{,}1748y^*_{i7} \tag{27}$$

El rango de valores que puede adoptar la medida sintética ITSNE2 se determina por el intervalo **(0,11; 3,73)**.

6.2.6 Construcción de la medida sintética para padres por medio de ACP

Con la finalidad de evaluar si existe una relación entre los puntajes de los diferentes instrumentos autoadministrados para los niños en edad escolar y adolescente, entonces se diseñó una medida sintética con la información que los padres pueden proveer sobre la higiene del sueño del niño.

Lo primero que se puede analizar es el resultado de los valores propios asociados a las dimensiones generadas por el análisis de componentes principales. Como se observa en la tabla 19, el porcentaje de varianza explicado por la primer dimensión es del 22,59 %.

Dim.	Valor Prop.	% Varianza	% Varianza Acum.
1	3.39	22.59	22.59
2	1.52	10.15	32.74
3	1.27	8.49	41.23
4	1.21	8.09	49.32
5	1.11	7.37	56.68
6	0.90	5.98	62.66
7	0.82	5.48	68.14
8	0.77	5.11	73.26
9	0.71	4.72	77.98
10	0.64	4.30	82.27
11	0.60	3.99	86.26
12	0.58	3.89	90.15
13	0.54	3.60	93.75
14	0.50	3.32	97.07
15	0.44	2.93	100.00

Tabla 19 *Resumen Valores Propios Resultantes del ACP de la medida sintética para padres*

Al igual que sucede con los resultados de los análisis multivariados para las medidas de los niños, se presenta una alta relación entre ítems y la primera dimensión. Esto se puede observar en la tabla 20, donde cada uno de los valores de las coordenadas para la primer dimensión son positivos. También se puede evidenciar que los tres ítems que más contribuyen a esta dimensión son: ship 7, ship 1 y ship 8.

	Coordenadas		**Contribuciones**	
Ítem	Dim 1	Dim 2	Dim 1	Dim 2
ship 1	0.62	-0.30	11.23	5.97
ship 2	0.40	-0.16	4.69	1.65
ship 3	0.56	-0.35	9.22	7.86
ship 4	0.39	-0.14	4.51	1.35
ship 5	0.53	0.22	8.26	3.14
ship 6	0.16	0.50	0.72	16.72
ship 7	0.62	-0.33	11.49	7.26
ship 8	0.59	-0.37	10.19	9.09
ship 9	0.50	0.22	7.35	3.29
ship 10	0.49	0.48	7.21	15.15
ship 11	0.41	0.47	4.94	14.81
ship 12	0.32	-0.20	2.95	2.68
ship 13	0.42	0.06	5.15	0.23
ship 14	0.29	0.34	2.53	7.55
ship 15	0.57	0.22	9.57	3.25

Tabla 20 *Coordenadas y Contribuciones para las dos Primeras Dimensiones del ACP de la medida para padres*

Formulación de la Medida Sintética (segunda alternativa) con datos transformados

Se construyó la medida sintética asociada a los puntajes de los padres al responder el instrumento que mide higiene del sueño del niño. Para ello, se usaron los valores del primer vector propio conseguido por medio del análisis de componentes principales y se ponderaron para generar los pesos correspondiente al modelo de variables latentes. A esta medida sintética se le denominó **ITSP** y su ecuación está dada por:

$$ITSP_i = 0{,}0899u_{i1} + 0{,}0581u_{i2} + 0{,}0815u_{i3} + 0{,}0570u_{i4} + 0{,}0771u_{i5} + 0{,}0227u_{i6} + 0{,}0909u_{i7} + 0{,}0856u_{i8} + 0{,}0728u_{i9} + 0{,}0720u_{i10} + 0{,}0596u_{i11} + 0{,}0461u_{i12} + 0{,}0609u_{i13} + 0{,}0427u_{i14} + 0{,}0830u_{i15} \quad (28)$$

El rango de valores que puede adoptar la medida sintética ITSP se determina por el intervalo **(1; 3)**.

6.2.7 Correlación entre las medidas sintéticas y la medida de los puntajes validados

Luego de la construcción de las medidas sintéticas, se generaron las escalas para cada medida sintética, tal como se definió en la metodología. Para determinar la correlación de las distintas escalas de las medidas sintéticas que buscan identificar el nivel de presencia de trastorno del sueño en escolares y adolescentes con los puntajes obtenidos en los instrumentos que ya han pasado por procesos de validación de consistencia interna y externa, se observó el coeficiente ρ de Spearman que existe entre cada escala y el promedio de puntajes globales de todos los instrumentos.

Correlaciones de Spearman						
Medida Escalada	$ITSN1$	$ITSN2$	$ITSN3$	$ITSNE1$	$ITSNE2$	$ITSP$
ρ	0.9476	0.9791	0.9997	0.5803	0.5936	0.3474

Tabla 21 *Relación lineal entre las escalas de las medidas sintéticas y el puntaje medio de los resultados de todas las pruebas*

Tal como se observa en la tabla 21, la escala de la medida más correlacionada con el promedio de puntajes globales de todos los instrumentos es la escala de ITSN3, esto debido a que está compuesta por la información de los puntajes globales. Las escalas de las medidas ITSN2 e ITSN1 también cuentan con una correlación positiva fuerte frente al promedio de los puntajes globales, a diferencia de sus versiones con los puntajes de los ítems transformados. Esto debido a que la transformación contempla la frecuencia con la cual se selecciona un puntaje dentro del ítem y puede darle más o menos peso a los ítems. A pesar de esto, las escalas de las medidas ITSNE1 e ITSNE2 conservan un moderado grado de correlación positiva.

Para dar respuesta a si la información que proveen los padres está altamente relacionada con la que indican los escolares y adolescentes, se puede ver en la tabla 21 que la escala de la medida ISTP presenta una correlación positiva baja. Esto se puede deber a que los padres no tienen control sobre los hábitos de sueño en los niños con edad escolar y adolescente, como si lo podrían tener en edades más prematuras.

6.2.8 Estadísticas Resumen de las Alternativas a Escala del Trastorno del Sueño en Escolares-Adolescentes

En la tabla 22 se observan los rangos teóricos y observados de las medidas sintéticas previas al escalado de 0 a 100. En el caso de las medidas

ITSNE1 e ITSNE2, los intervalos teóricos dependen de la muestra, por lo que se determinaron mediante remuestreo.

Rango	$ITSN1$	$ITSN2$	$ITSN3$	$ITSNE1$	$ITSNE2$	$ITSP$
Observado	(0.64, 4.69)	(1.76, 11.72)	(21.75, 52.43)	(0.05, 1.20)	(0.11, 3.73)	(1, 2.85)
Teórico	(0.64, 5.01)	(1.67, 12.45)	(21.47, 56.52)	(0.05, 1.20)	(0.10, 3.78)	(1, 3)

Tabla 22 *Resumen de los Rangos Teóricos y Observados de las Medidas Sintéticas*

Después de la construcción de las medidas sintéticas y al observar que la escala de la medida sintética ITSP presenta una baja correlación con lo que informan los escolares y adolescentes sobre los diferentes trastornos subyacentes relacionados con el sueño, se considera pertinente analizar en estudios futuros esta relación con mayor profundidad. Estos resultados concuerdan con la literatura, tal como lo dice Moo-Estrella et al (2018) son escasos los instrumentos de autoreporte sobre el sueño para niños que han sido validados en español. Sin embargo, estos muestran una mejor correlación con las medidas objetivas del sueño en comparación con aquellos que involucran la participación de los padres, como lo evidencian las técnicas actigráficas.

Escalas	Min.	1er Q.	Mediana	Media	3er Q.	Max.	Desv. Est.
$ITSN1$	0.00	22.07	34.95	37.37	50.64	100.00	19.98
$ITSN2$	0.00	16.98	31.67	33.76	47.12	100.00	20.79
$ITSN3$	0.00	25.26	40.78	40.69	54.46	100.00	20.01
$ITSNE1$	0.00	13.85	24.65	29.84	40.95	100.00	20.35
$ITSNE2$	0.00	9.58	17.45	22.35	31.28	100.00	17.82

Tabla 23 *Resumen Estadístico de las Escalas*

Al analizar la tabla 23, se puede observar que al aplicar la escala referente a la medida ITSN3, esta presenta la mediana más grande $(40, 78)$, por lo que, la mitad de los escolares y adolescentes obtuvieron un puntaje de trastorno del sueño superior a 40 puntos sobre 100 posibles. Además que las escalas de las medidas ITSN1 e ITSNE2 son las que menor variabilidad presentan, con valores de 19,98 y 17,82, respectivamente.

En los gráficos presentados en la figura 6 se observa cómo cambia la distribución de los puntajes según la escala. La escala asociada a la medida sintética ITSN3 sigue una distribución normal, según la prueba de Anderson-Darling, con un valor-p de 0,07, el cual es mayor al nivel de significancia de 0,05, a diferencia de los histogramas de las escalas de las medidas ITSN1 e ITSN2, en los cuales se evidencia que hay mayoritariamente individuos con valores de puntajes bajos. Este comportamiento es aún más marcado en las escalas ITSNE1 e ITSNE2, para las cuales muy pocos individuos presentan puntajes altos (subestiman la presencia de trastorno del sueño en la muestra). Esto se ve reflejado en las pruebas Anderson-Darling cuyos valores-p son todos menores al nivel de significancia $\alpha = 0{,}05$ (ver en 8.4).

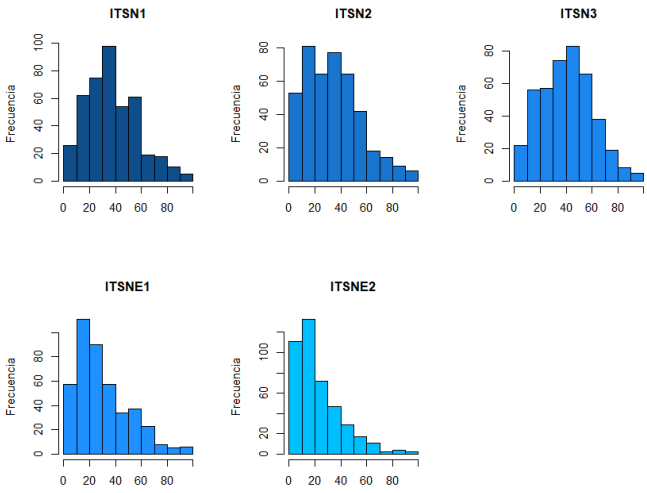

Figura 6 Distribución de los puntajes reescalados según la medida sintética

6.3 Resultados del Segundo Objetivo: Evaluación del desempeño de la medida sintética (PECM)

Como se pudo evidenciar en la sección anterior, los histogramas indicaban la existencia de asimetría a la derecha, según la prueba Anderson-Darling las escalas de las medidas: ITSN1, ITSN2, ISTNE1 e ITSNE2 no siguen una distribución normal ya que los valores-p resultantes son menores al nivel de significancia $\alpha = 0,05$ (ver 8.4). Con relación a la escala de la medida ITSNE3, se asume que su distribución es aproximadamente simétrica. Dado estos comportamientos, se seleccionó como el mejor estimador para representar la tendencia central de los puntajes a la mediana, exceptuando el caso de la escalas de la medida ITSN3, para la cual se seleccionó la media.

Mediante la técnica de remuestreo Bootstrap se calcularon los valores del estadístico de interés (mediana o media) para diferentes submuestras con reemplazo. De esta manera se busca aproximar la distribución muestral real del estadístico en cuestión y explorar la posibilidad de aplicar conceptos de la teoría de la estimación abordados en la inferencia estadística. En la figura 7 se muestra la distribución de los estimadores según la medida sintética escalada, cuando se tienen tamaños de muestra iguales a 30. La línea roja indica el valor esperado del estimador o de otra forma dicho; el nivel típico de trastorno del sueño esperado en la población de escolares y adolescentes.

Como este análisis se realizó con un tamaño de muestra igual a 428, es de interés analizar lo que sucede al variar ese tamaño, por lo que se observó el comportamiento de la medida sintética escalada en función de la varianza y el

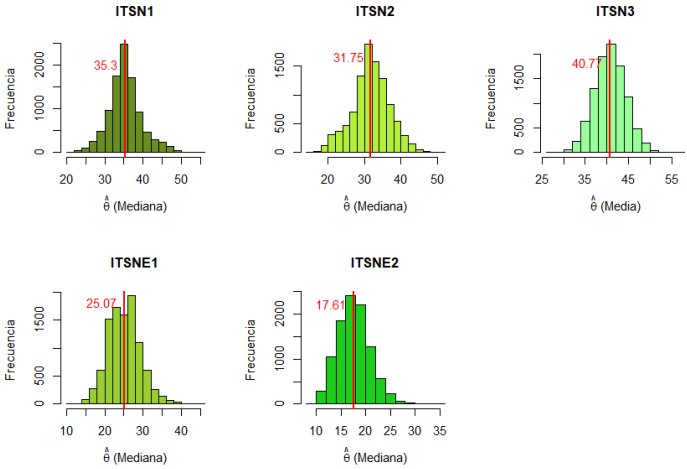

Figura 7 Distribución del estimador del pseudoparámetro asociado al constructo (trastorno del sueño mediano)

sesgo. Inicialmente, se asume que el estimador es insesgado dado que el análisis de componentes principales y el análisis factorial múltiple requiere de que el conjunto de variables sea por lo menos centrado. No obstante es posible que el estimador esté desviado ligeramente del pseudo parámetro y se sobreestime o subestime el nivel típico de trastorno del sueño. Para ello se construyó la tabla 24, donde se muestra cómo el Pseudo Error Cuadrático medio disminuye en función del tamaño de muestra, ocasionado por la disminución en el valor de la varianza y el sesgo.

Escala de la Medida Sintética	Métrica	Tamaño de muestra						
		30	50	100	500	1000	10000	30000
ITSN1	PECM	19,1745	10,5395	4,0797	0,6353	0,2891	0,0249	0,0093
	$Var(\hat{\theta}_m)$	19,0613	10,5022	4,0740	0,6327	0,2880	0,0241	0,0084
	$sesgo(\hat{\theta}_m)^2$	0,1132	0,0373	0,0057	0,0026	0,0011	0,0009	0,0009
ITSN2	PECM	26,2962	15,9042	6,9878	1,2831	0,6702	0,0531	0,0109
	$Var(\hat{\theta}_m)$	26,2908	15,8968	6,9737	1,2815	0,6701	0,0530	0,0107
	$sesgo(\hat{\theta}_m)^2$	0,0054	0,0073	0,0141	0,0016	0,0000	0,0001	0,0003
ITSN3	PECM	13,1161	7,9788	3,9292	0,8016	0,4080	0,0401	0,0139
	$Var(\hat{\theta}_m)$	13,1135	7,9773	3,9291	0,8011	0,4071	0,0395	0,0133
	$sesgo(\hat{\theta}_m)^2$	0,0027	0,0015	0,0001	0,0005	0,0009	0,0006	0,0005
ITSNE1	PECM	17,7530	10,8779	5,8182	1,6742	0,9615	0,1710	0,0974
	$Var(\hat{\theta}_m)$	17,6606	10,8320	5,8071	1,6647	0,9517	0,1701	0,0967
	$sesgo(\hat{\theta}_m)^2$	0,0924	0,0459	0,0111	0,0096	0,0098	0,0009	0,0007
ITSNE2	PECM	10,7964	6,5116	3,2416	0,7897	0,4762	0,0436	0,0045
	$Var(\hat{\theta}_m)$	10,7700	6,5057	3,2408	0,7894	0,4755	0,0434	0,0045
	$sesgo(\hat{\theta}_m)^2$	0,0264	0,0058	0,0008	0,0003	0,0007	0,0002	0,0000

Tabla 24 *Resultados del desempeño según el PECM para las medidas sintéticas diseñadas*

Al comparar el Pseudo Error Cuadrático Medio, la tabla 24 revela que para tamaños de muestra pequeño, la escala de la medida sintética ITSN3 tiene el menor sesgo, seguida de la escala de ITSNE2. Sin embargo, a medida que

el tamaño de muestra aumenta se puede observar que el sesgo para la escala de ITSNE2 disminuye de manera constante hasta ser practicamente cero, a diferencia del sesgo de la escala de ITSN3, el cual a pesar de que decrece, a partir de un tamaño de muestra igual a 100 fluctua. Otro comportamiento interesante por analizar es que a partir de un tamaño de muestra igual a 500, la varianza del estimador para la escala de ITSN1 disminuye rápidamente en comparación a las demás medidas sintéticas escaladas.

Según lo descrito en la metodología, la mejor medida sintética y su respetiva escala, será la que presente el menor Pseudo Error Cuadrático Medio. Dados los resultados de la tabla 24, la medida a seleccionar es la denominada como ITSNE2, derivada de la aplicación de Análisis de Componentes Principales sobre nuevos factores determinados por los primeros componentes resultantes de un primer análisis de componentes principales sobre cada instrumento individualmente, cuya muestra de puntajes fueron transformados en puntuaciones monotónicas continuas que safistacen la propiedad de equidistancia.

6.4 Resultados del Tercer Objetivo: Descripción de las características asociadas a los trastornos del sueño en el grupo de individuos

Con el fin de clasificar a los escolares y adolescentes, se presenta una propuesta que agrupa a los individuos según sus puntajes en la escala de la medida ITSNE2, utilizando el algoritmo K-means. Además, se describen algunas características de la población y su relación con las diferentes clasificaciones.

En primer lugar, se estableció el número óptimo de agrupaciones que indican el grado de severidad del trastorno del sueño presente en los escolares y adolescentes. Tal como se observa en la figura 8, el número de clusters que minimizan la suma de cuadrados total es igual a 3, esto implica que con esa cantidad de cluster, se logra la menor variabilidad entre los datos sin complicar innecesariamente el modelo de agrupamiento. Por ello, se propone dividir los puntajes de la muestra en tres grados de severidad: leve, moderado y severo.

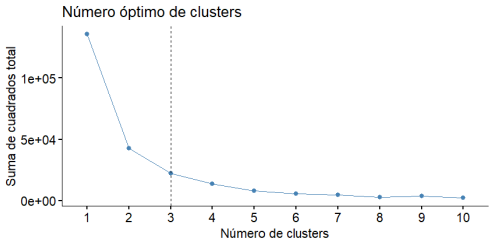

Figura 8 *Criterio de selección del número de agrupaciones por grado de severidad según el puntaje obtenidos en la escala de la medida ITSNE2*

Una vez se definieron la cantidad de agrupaciones, se establecieron los puntos de corte para cada uno de los grados de severidad. Los resultados presentados en la tabla 25 muestran que, mediante un remuestro bootstrap, los individuos que se ubicaban en un grado de trastorno del sueño leve presentaron, en promedio, puntajes mínimos en la medida sintética escalada de 0,05. Con relación a los individuos clasificados con presencia de trastorno del sueño moderado, estos presentaron, en promedio puntajes mínimos de 19,39, mientras que los clasificados en un grado severo obtuvieron puntajes de 44,3.

Límite promedio	Leve	Moderado	Severo
mínimo	0.05	19.39	44.3
máximo	19.01	42.56	96.31

Tabla 25 *Valores de límite para cada grado de severidad*

Con base a los límites de cada cluster que representa un grado de severidad, se propusieron unos puntos de corte. Sea T_i el puntaje en la escala de 0 a 100 para el individuo i, dado su puntaje en la medida $ITSNE2_i$, entonces:

- Leve: Si y solo sí;

$$0 < T_i \leq 19{,}40$$

- Moderado: Si y solo sí;

$$19{,}40 < T_i \leq 44{,}30$$

- Severo: Si y solo sí:

$$44{,}30 < T_i \leq 100$$

Con fines de explorar el comportamiento de la muestra frente a las clasificaciones por su grado de severidad, se presenta la figura 9 el digrama de cajas para cada uno de los grados.

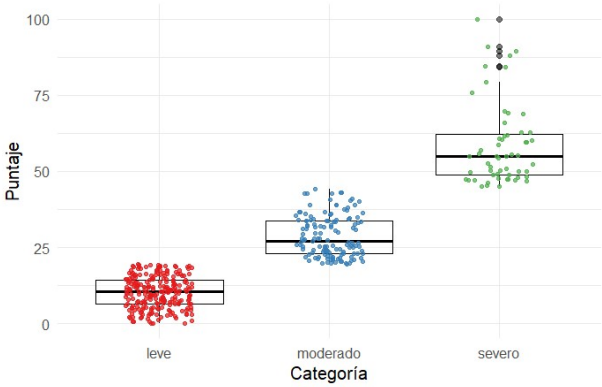

Figura 9 *Boxplot del puntaje reescaldo de los individuos clasificados por grado de severidad del trastorno de sueño*

En la figura 9 se observa que los escolares y adolescentes que fueron clasificados con presencia leve de trastorno del sueño, son poco variables entre si, además que la distribución de estos es bastante simétrica, indicando que estos individuos presentan características semejantes en cuanto a imsonio, somnolencia, alteración del sueño, deterioro relacionado con el sueño, sueño y vigilia, ansiedad y depresión. Dado los pesos asociados a estos factores, es muy probable que presenten los menores valores especialmente en alteración, deterioro y ansiedad.

En lo que respecta al grado de severidad moderado, se observa en la figura 9 que existe un leve sesgo hacia la derecha, en este caso, esto implica que hay diferencias en la variabilidad. Lo que se puede interpretar de esta particularidad, es que los escolares y adolescentes con un grado de severidad moderado no es tan bajo condiciones similares frente al imsonio, somnolencia, alteración del sueño, deterioro relacionado con el sueño, sueño y vigilia, ansiedad y depresión, ya que algunos pueden estar más cercanos a características de los leves y otros más encaminados a un grado severo.

Finalmente, los escolares y adolescentes clasificados con grado severo, son los que mayor variabilidad presentan, tal como se observa en la figura 9, la mayoría de individuos se encuentran entre un puntaje de 50 y un poco más de 60, no obstante, hay un grupo de individuos que se distancia bastante del resto. Estos son los que mayores problemas están presentando, sobre todo en cuanto a la alteración del sueño, deterioro relacionado con el sueño y ansiedad.

Variable		Frecuencia	% Leve	% Moderado	% Severo	Prueba χ^2
Sexo	Masculino	186	28.3 %	11.9 %	3.3 %	$\chi^2 = 14.1790$
	Femenino	242	27.6 %	19.2 %	9.8 %	gl = 2
	Total	428	55.8 %	31.1 %	13.1 %	valor-p = 0.0008
Grupedad	Escolares	99	14,0 %	6,8 %	2.3 %	$\chi^2 = 1.5312$
	Adolescentes	329	41.8 %	24.3 %	10.7 %	gl = 2
	Total	428	55.8 %	31.1 %	13.1 %	valor-p = 0.4651
Diagnóstico IMC	Normal	314	42.5 %	22.3 %	9.7 %	$\chi^2 = 1.7105$
	Sobrepeso/Obesidad	111	12.8 %	9.0 %	3.6 %	gl = 2
	Total	425	55.3 %	31.4 %	13.3 %	valor-p = 0.4252
Presenta alguna enfermedad	Si	121	16.3 %	10.5 %	4.1 %	$\chi^2 = 0.8792$
	No	271	40.1 %	20.7 %	8.4 %	gl = 2
	Total	392	56.4 %	31.1 %	12.5 %	valor-p = 0.6443
Alguien en la familia presenta enfermedad	Si	157	23.0 %	11.7 %	5.4 %	$\chi^2 = 0.4821$
	No	235	33.4 %	19.4 %	7.1 %	gl = 2
	Total	392	56.4 %	31.1 %	12.5 %	valor-p = 0.7858
Desorden gastrointestinal funcional	Si	138	16.6 %	14.0 %	4.6 %	$\chi^2 = 8.5194$
	No	254	39.8 %	17.1 %	7.9 %	gl = 2
	Total	392	56.4 %	31.1 %	12.5 %	valor-p = 0.0141

Tabla 26 *Relación entre los grados de severidad en los participantes y variables sociodemográficas y de salud ($\alpha = 0{,}05$)*

Tal como se observa en la tabla 26, con un valor-p igual a 0,0008 menor al nivel de significancia $\alpha = 0{,}05$ se rechaza la hipótesis nula (H_0 : no existe asociación entre las variables), por tanto, existe asociación significativa entre la variable sexo y el grado de severidad del trastorno del sueño. Lo que implica que el sexo femenino y masculino presenta distribuciones diferentes en cuanto al grado de severidad del trastorno del sueño. En particular se observa una

mayor prevalencia de escolares y adolescentes con sexo femenino cuyo grado de severidad es moderado o severo, en comparación a los de sexo masculino. Esto podría estar influenciado por los múltiples cambios hormonales que experimentan las mujeres en la adolescencia.

Como se aprecia en la tabla 26, el desorden gastrointestinal funcional y el grado de severidad del trastorno del sueño están asociados significativamente, con un valor-p igual a 0,0141 menor al nivel de significancia $\alpha = 0{,}05$, los resultados señalan que la distribución de los escolares y adolescentes que presentan o no, por lo menos un desorden gastrointestinal es distinta y no se debe solo al azar. La proporción de escolares y adolescentes que tienen un grado de severidad del trastorno del sueño moderado o severo es mayor para los presentan desorden gastrointestinal (39,8 % y 13,06 %), que los que no (26,4 % y 12,2 %). A pesar de la falta de estudios suficientes sobre las causas, estos resultados subrayan la relevancia del eje cerebro-intestino, una red bidireccional que conecta el sistema nervioso central con el aparato digestivo y cómo esta puede influir sobre el comportamiento relacionado con el sueño en escolares y adolescentes. En este sentido, un estudio realizado por Santucci et al (2024) señala que los niños con trastornos funcionales del dolor abdominal (FAPD) presentaron una peor calidad del sueño, mayor gravedad del insomnio, alteraciones en el sueño, deterioro relacionado con el sueño, somnolencia diurna, deficiencias en la higiene del sueño, jadeo y pesadillas, en comparación con sus pares saludables.

Por otro lado, variables como grupedad, el diagnóstico del IMC, si presenta alguna enfermedad o si es su familia alguien presenta enfermedad, no guardan relación con el grado de severidad del trastorno del sueño.

7 Conclusiones y Recomendaciones

7.1 Conclusiones

Con base a los resultados obtenidos, se puede concluir que la mejor alternativa para construir una medida sintética que permita conocer el nivel de presencia de trastornos del sueño en escolares y adolescentes, partiendo del modelo reflectivo de variables latentes, es la alternativa II (ACP) con una transformación monótona continua de los ítems. Los factores más relevantes para la medida son: alteración del sueño, deterioro relacionado con el sueño, ansiedad y depresión, traduciendose a que los escolares y adolescentes que tengan ciertos problemas para dormir, conciliar el sueño, realizar actividades con normalidad por tener sueño o experimentar síntomas de ansiedad o depresión, tienden a tener un mayor grado de severidad de trastorno del sueño.

Al evaluar la mediana de los puntajes obtenidos para cada una de las escalas como un estimador del nivel típico del trastorno del sueño que presenta la población de escolares y adolescentes, se puede concluir que, a pesar de que el estimador es aproximadamente insesgado para todos los casos, las estimaciones pueden variar considerablemente si la muestra es pequeña. Sin embargo, los estimadores son aparentemente consistentes, ya que al aumentar el tamaño de la muestra, la varianza y el sesgo del estimador disminuye.

Las medidas más correlacionadas positivamente con el promedio de puntajes globales de todos los instrumentos fueron las conseguidas por medio de las alternativas I (AFM), II (ACP) y III (puntajes totales y ACP). Esto se debe a que, a diferencia de las alternativas con datos transformados, estas técnicas no alteran la distribución original de los datos, lo que permite mantener la relación con la variabilidad de los puntajes globales. Por otro lado, las escalas para estas alternativas tienden a posicionar a la mayoría de individuos al rededor de 40 puntos de 100 posibles, mientras que las escalas para la alternativa I (AFM) y II (ACP) con una transformación monótona continua de los ítems, tienden a ajustar los puntajes según la influencia que tenga la percepción general sobre determinado ítem o factor-trastorno, ubicando la mayoría de individuos al rededor de 20 puntos de 100 posibles. Esto corresponde a la naturaleza misma del método de transformación, ya que se le da mayor o menor ponderación a las respuestas más frecuentes.

Tomando como base los resultados de las correlaciones entre las medidas y el promedio de puntajes globales de todos los instrumentos, parece ser que los padres no tienen mucho conocimiento sobre las dificulades relacionadas con el sueño que los escolares y adolescentes presentan durante la escuela y en las noches. Conocer esto puede ser de ayuda para los profesionales de la salud a la hora de generar estrategias encaminadas a reducir los niveles del trastorno del sueño presente en el escolar o adolescente, haciendo parte del proceso a

los padres.

Al final de los resultados se pudo observar que existe asociación entre el sexo y el grado de severidad, por lo que se concluye que las escolares y adolescentes tienden a presentar trastornos del sueño moderados y severos. Por lo tanto, se vuelve una necesidad entender las condiciones experimentadas por el sexo femenino que propician esas dificultades relacionadas con el sueño. Por otro lado, efectivamente los desordenes gastrointestinales están asociados con el grado de severidad del trastorno del sueño, sugiriendo que tratar este tipo de desordenes puede influir positivamente en el sueño de los escolares y adolescentes.

7.2 Recomendaciones

Se recomienda generar nuevas propuestas de medidas sintéticas considerando la matriz de correlación policórica, con la finalidad de comparar los resultados con la metodología implentada. Además, se recomienda evaluar otros modelos de variables latentes, como lo es el modelo formativo, para el que se necesitarían variables observables, por ejemplo, indicadores derivados de polisomnografías, de resistencia física, indicadores de estudios clínicos, entre otros. Al Optar por un modelo formativo se llegaría a diferentes resultados, ya que el constructo sería una composición de esas variables observables, contrario a lo que sucede con el modelo reflexivo. De tal manera que se pueda nutrir la discusión del trastorno de sueño como constructo.

Se considera importante que se estudie sobre los efectos que tiene la muestra sobre la escala final, dada el contexto social, cultural, económico, de seguridad, educativo y familiar del escolar y adolescente; los resultados pueden variar. Para esto se recomienda probar la escala ITSNE2 en diferentes escenarios y comparar sus resultados con escalas o instrumentos validados.

Por último se recomienda generar nuevos trabajos de investigación que exploren la relaciones que bajo los resultados obtenidos fueron significativas, como los son el sexo y el desorden gastrointestinal. Esto podría desarrollarse ajustando un modelo Gamma tomando la medida sintética escalada como la variable respuesta y como covarables, el sexo y el desorden gastrointestinal ($ITSNE2 \sim Sexo + Desorden$), ya que este tipo de modelo es útil cuando la variable de respuesta es continua, positiva y presenta una distribución sesgada hacia la derecha, como es el caso de la medida sintética. Con esto, se podrían complementar los resultados de la pruebas chi cuadrado finales.

8 Anexos

8.1 Código en R

Se puede acceder al código en el software R utilizado a lo largo de este trabajo mediante este link: Script

8.2 Nubes de individuos y tipologías de grupos

A continuación se presentan a manera de anexo algunos gráficos de los resultados de los análisis (ACP y AFM) planteados.

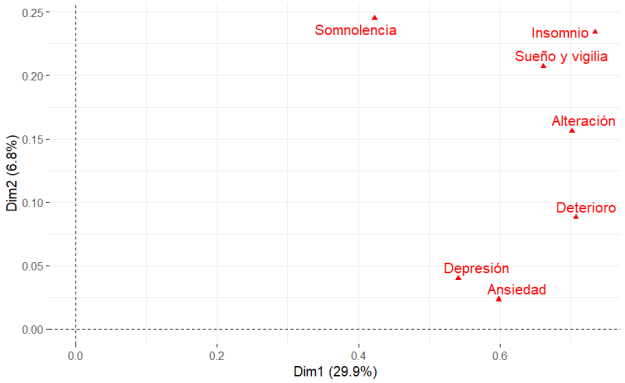

Figura 10 *Similitudes entre los trastornos medidos por los instrumento para la alternativa I (AFM)*

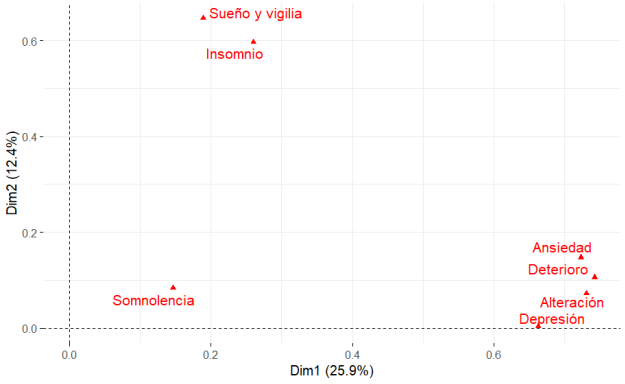

Figura 11 *Similitudes entre los trastornos medidos por los instrumento para la alternativa I (AFM) con puntajes transformados*

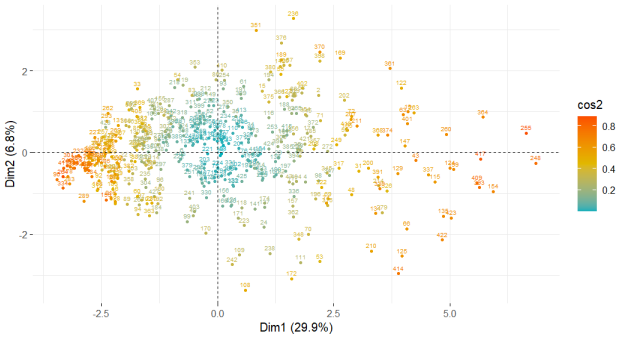

Figura 12 *Nube de individuos para la alternativa I (AFM)*

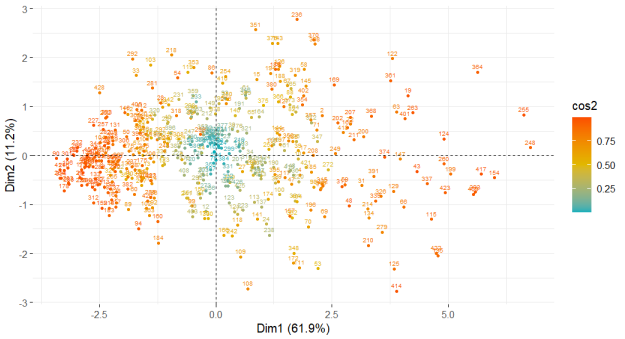

Figura 13 *Nube de individuos para la alternativa II (ACP)*

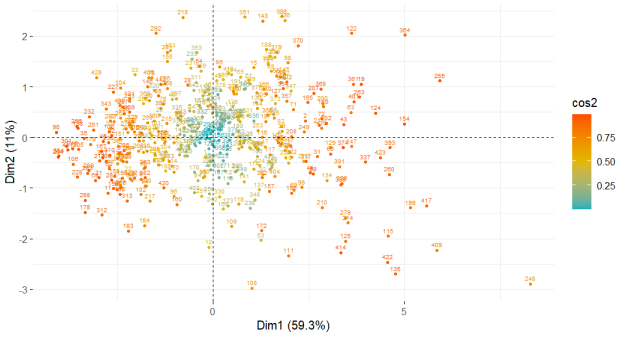

Figura 14 *Nube de individuos para la alternativa III (puntajes totales y ACP)*

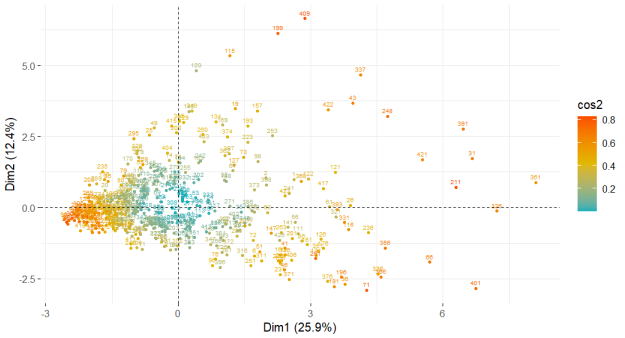

Figura 15 *Nube de individuos para la alternativa I (AFM) con puntajes transformados*

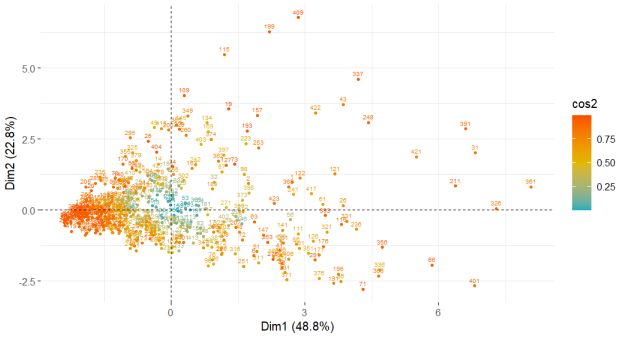

Figura 16 *Nube de individuos para la alternativa II (ACP) con puntajes transformados*

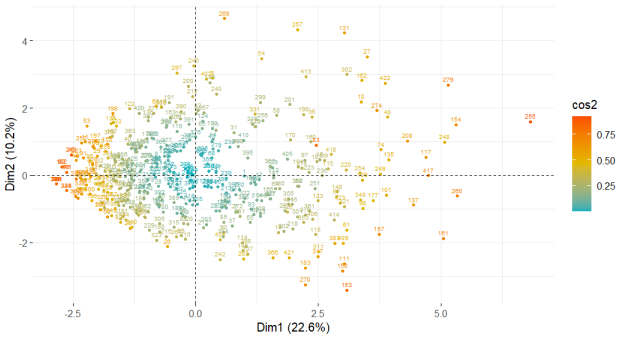

Figura 17 *Nube de individuos para la medida de padres*

8.3 Clasificaciones según puntaje total en los instrumentos

Instrumento	Clasificación según puntaje
PISI	**Leve:** puntajes entre 1 y 10 **Moderado:** puntajes entre 11 y 20 **Severo:** puntajes entre 21 y 30
ESS	**Somnolencia diurna excesiva:** puntajes mayores a 10 **Alto nivel de somnolencia diurna excesiva:** puntajes mayores o iguales a 16
PROMIS SD	**Ninguno a leve:** puntajes menores a 55 **Leve:** puntajes entre 55 y 59,9 **Moderado:** puntajes entre 60 y 69,9 **Grave:** puntajes mayores o iguales a 70
PROMIS SRI	**Ninguno a leve:** puntajes menores a 55 **Leve:** puntajes entre 55 y 59,9 **Moderado:** puntajes entre 60 y 69,9 **Grave:** puntajes mayores o iguales a 70
ASWS	Sin puntajes de corte específicos. Puntajes en la subescala cercanos al puntaje máximo indica mejor calidad del sueño, exceoto cuando esta se invierte
PROMIS Anxiety	**Ninguno a leve:** puntajes menores a 55 **Leve:** puntajes entre 55 y 59,9 **Moderado:** puntajes entre 60 y 69,9 **Grave:** puntajes mayores o iguales a 70
PROMIS Depression	**Ninguno a leve:** puntajes menores a 55 **Leve:** puntajes entre 55 y 59,9 **Moderado:** puntajes entre 60 y 69,9 **Grave:** puntajes mayores o iguales a 70

Tabla 27 *Grado de severidad según el puntaje para los instrumentos aplicados en los escolares y adolescentes*

8.4 Resultados pruebas Anderson-Darling

Prueba Anderson-Darling, datos: puntajes de las medidas escaladas		
Escalas	A^*	valor-p
ITSN1	2.5674	0.00
ITSN2	3.8327	0.00
ITSN3	0.69257	0.07
ITSNE1	10.43	0.00
ITSNE2	14.751	0.00
Prueba Anderson-Darling, datos: estimadores de los pseudoparámetros		
ITSN1	66.027	0.00
ITSN2	21.041	0.00
ITSN3	6.7776	0.00
ITSNE1	16.965	0.00
ITSNE2	14.268	0.00

Tabla 28 *Resultados de la prueba formal Anderson-Darling aplicada a los dos conjuntos de datos para los cuales se analizó la simetría por medio del histograma*

Referencias

Bartholomew DJ, Knott M, Moustaki I (2011) Latent variable models and factor analysis: A unified approach. John Wiley & Sons

Belalcazar M (2023) Evaluación del desempeño estadístico de escalas construidas para representar la expresión de variables latentes mediante análisis multivariado. Master's thesis, Universidad del Valle

Beteta FG, Cristóbal ID, Alegre AN, et al (2023) Protocolo diagnóstico y terapéutico del insomnio. Medicine-Programa de Formación Médica Continuada Acreditado 13(86):5102–5106

Bironneau V, Meurice JC (2023) Somnolence diurne excessive au cours du syndrome d'apnées hypopnées obstructives du sommeil (sahos). Médecine du Sommeil 20(3):135–146

Boari G, Nai Ruscone M (2015) A procedure simulating likert scale item responses. Electronic journal of applied statistical analysis 8(3):288–297

Bollen KA (2002) Latent variables in psychology and the social sciences. Annual review of psychology 53(1):605–634

Casella G, Berger R (2024) Statistical inference. CRC Press

Chakrabartty SN (2023) Likert equidistante como suma ponderada de categorías de respuesta. CULTURA EDUCACIÓN Y SOCIEDAD 14(1):75–92

Chyung SY, Barkin JR, Shamsy JA (2018) Evidence-based survey design: The use of negatively worded items in surveys. Performance Improvement 57(3):16–25

D'Agostino RB (2017) Goodness-of-fit-techniques. Routledge

Escofier B, Pagès J (1998) Analyses factorielles simples et multiples. Dunod, Paris 284

Friedman G, Turk KW, Budson AE (2023) The current of consciousness: Neural correlates and clinical aspects. Current Neurology and Neuroscience Reports 23(7):345–352

Gállego Pérez-Larraya J, Toledo J, Urrestarazu E, et al (2007) Clasificación de los trastornos del sueño. In: Anales del Sistema Sanitario de Navarra, SciELO Espana, pp 19–36

González Rojas VM (2014) Análisis conjunto de múltiples tablas de datos mixtos mediante pls

Hair Jr JF, Matthews LM, Matthews RL, et al (2017) Pls-sem or cb-sem: updated guidelines on which method to use. International Journal of Multivariate Data Analysis 1(2):107–123

Hamui Sutton L (2019) La noción de "trastorno": entre la enfermedad y el padecimiento. una mirada desde las ciencias sociales. Revista de la Facultad de Medicina (México) 62(5):39–47

Herskovic V, Matamala M (2020) Somatization disorder, anxiety and depression in children and adolescents. Revista Medica Clinica Las Condes 31:183–7

Husson F, Josse J, Le S, et al (2016) Package 'factominer'. An R package 96(96):698

Ikotun AM, Ezugwu AE, Abualigah L, et al (2023) K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. Information Sciences 622:178–210

Kassambara A, Mundt F (2017) Package 'factoextra'. Extract and visualize the results of multivariate data analyses 76(2)

Lord FM (1953) On the statistical treatment of football numbers.

MacCallum RC, Austin JT (2000) Applications of structural equation modeling in psychological research. Annual review of psychology 51(1):201–226

Mahendran R, Subramaniam M, Cai Y, et al (2006) Survey of sleep problems amongst singapore children in a psychiatric setting. Social psychiatry and psychiatric epidemiology 41:669–673

Mendivelso F, Rodríguez M (2018) Prueba chi-cuadrado de independencia aplicada a tablas 2xn. Revista Médica Sanitas 21(2):92–95

Meneses J, Barrios M (2014) Psicometría. Psicometría pp 1–273

Moo-Estrella JA, Valencia-Flores M, Arankowsky-Sandoval GM (2018) Development and validation of the sleep disturbances scale for school-age children. Acta pediátrica de México 39(2):121–133

Oh A, Koehler A, Yonker M, et al (2023) Sleep disorders and chronic pain syndromes in the pediatric population. In: Seminars in Pediatric Neurology, Elsevier, p 101085

Rabner J, Kaczynski KJ, Simons LE, et al (2017) The sleep hygiene inventory for pediatrics: Development and validation of a new measure of sleep in a sample of children and adolescents with chronic headache. Journal of child neurology 32(13):1040–1046

Randler C, Kolomeichuk S, Morozov A, et al (2019) Psychometric properties of the russian version of the pediatric daytime sleepiness scale (pdss). Heliyon 5(7)

Santucci N, Velasco-Benitez CA, King C, et al (2024) Children with functional abdominal pain have more sleep disturbances. first of a kind school study. The Journal of Pain 25(4):67–68

Stevens SS (1946) On the theory of scales of measurement. Science 103(2684):677–680

Tukey JW (1962) The future of data analysis. The annals of mathematical statistics 33(1):1–67

Vélez RR, Zamora LH, Bautista JEC, et al (2018) Confiabilidad y validez del cuestionario de trastornos de sueño bears en niños y adolescentes escolares de bogotá, dc, colombia: Estudio fuprecol. Retos: nuevas tendencias en educación física, deporte y recreación (34):89–93

Velleman PF, Wilkinson L (1993) Nominal, ordinal, interval, and ratio typologies are misleading. The American Statistician 47(1):65–72

Vestin M, Åsberg M, Wiberg M, et al (2023) Psychometric validity of the montgomery and åsberg depression rating scale for youths (madrs-y). Nordic Journal of Psychiatry 77(5):421–431

Yuan W, Chen L, Wu Y, et al (2023) Sleep time and quality associated with depression and social anxiety among children and adolescents aged 6–18 years, stratified by body composition. Journal of Affective Disorders

Zapata EDM (2012) Comparación de la eficiencia en la estimación de parámetros entre análisis de estructura de covarianzas, modelo de rutas mediante mínimos cuadrados parciales y análisis generalizado de componentes estructurales un estudio de simulación. PhD thesis, Universidad Autónoma de Madrid