

Comparación De Técnicas Estadísticas Para La Validación De Pruebas En Psicología

Elda Carolina García Tobo

Trabajo De Grado Para Optar El Título De Especialista En Estadística

Director

Henry Sebastián Rangel Quiñonez

Magíster en Ciencias Estadísticas

Universidad Industrial de Santander

Facultad De Ciencias

Escuela De Matemáticas

Especialización En Estadística

Bucaramanga

2020

Contenido

	Pág.
Introducción	8
1. Justificación	10
2. Antecedentes	11
3. Objetivos	13
3.1 Objetivo General	13
3.2 Objetivos Específicos.....	14
4. Marco Teórico.....	14
4.1 Análisis Factorial Exploratorio (AFE).....	14
4.1.1. Objetivo de la técnica.....	15
4.1.2 Modelo	15
4.1.3 Supuestos del modelo y validación.....	16
4.1.3.1 Escala de las variables	18
4.1.3.2 Tamaño de muestra	19
4.1.3.3 Matriz de correlación.....	20
4.1.4 Ajuste del AFE.....	22
4.1.4.1 Rotación de ejes.....	23
4.2 Análisis factorial confirmatorio (AFC).....	25

4.2.1 Objetivo de la técnica.....	25
4.2.2 Modelo	26
4.2.3 Ajuste AFC	27
4.3 Teoría de respuesta al ítem	29
5. Metodología	36
5.1 Descripción de la población y muestra	36
5.2 Métodos.....	37
6. Resultados	38
6.1 Análisis factorial exploratorio AWA	38
6.3 Análisis factorial confirmatorio IECA	48
6.4 Análisis teoría de respuesta al ítem IECA	56
7. Conclusiones	60
Referencias.....	63

Lista de tablas

	Pág.
Tabla 1. <i>Criterios de comparación AFE y AFC</i>	10
Tabla 2. <i>Determinante, Bartlett y KMO</i>	39
Tabla 3. <i>Resumen del análisis factorial exploratorio del test AWA SPSS</i>	44
Tabla 4. <i>Resumen del análisis factorial exploratorio del test AWA Rstudio</i>	45
Tabla 5. <i>Resumen del análisis factorial exploratorio del test AWA Factor.</i>	47
Tabla 6. <i>Estadísticas de bondad de ajuste AMOS</i>	52
Tabla 7. <i>Estadísticas de bondad de ajuste Rstudio</i>	52
Tabla 8. <i>Estimadores de los parámetros F1</i>	57
Tabla 9. <i>Estimadores de los parámetros F2</i>	57

Lista de figuras

	Pág.
<i>Figura 1.</i> Rotación de los ejes en el espacio euclídeo	23
<i>Figura 2.</i> Curva característica del ítem.....	34
<i>Figura 3.</i> CCI comparación parámetro dificultad.....	35
<i>Figura 4.</i> Pasos de análisis y selección bajo el modelo TRI	36
<i>Figura 5.</i> Matriz de correlación de Pearson prueba AWA	41
<i>Figura 6.</i> Matriz de correlación Policórica prueba AWA	42
<i>Figura 7.</i> Gráfico de sedimentación SPSS (superior) Análisis paralelo software Factor (inferior izquierda) y software Rstudio (inferior derecha).....	43
<i>Figura 8.</i> Diagrama de flujo AFC paquete AMOS modelo dos factores (superior) modelo tres factores (inferior)	50
<i>Figura 9.</i> Diagrama de flujo AFC paquete Rstudio modelo dos factores (superior) modelo tres factores (inferior)	51
<i>Figura 10.</i> Diagrama AFC paquete Rstudio modelo dos factores (superior) modelo tres factores (inferior)	55
<i>Figura 11.</i> Boxplot y grafico de barras dificultad y discriminación modelo 2PL.....	56
<i>Figura 12.</i> ICC modelo 2PL F1 (superior) F2 (inferior)	58
<i>Figura 13.</i> Función de información de los ítems (superior) y test (inferior) modelo 2PL	59

Resumen

Título: Comparación De Técnicas Estadísticas Para La Validación De Pruebas En Psicología*

Autora: Elda Carolina García Tobo**

Palabras clave: AFC, AFE, TRI, ítems categóricos, validación, psicología.

Descripción

Esta investigación comparó al interior de las Técnicas Estadísticas AFE, AFC y TRI los resultados obtenidos al seleccionar determinados criterios, comparando entre lo que generalmente realizan los psicólogos y lo que recomienda la estadística para validar pruebas psicológicas. Se analizaron los datos obtenidos de la Escala de actitud hacia el bienestar animal AWA y el índice de empatía para niños y adolescentes test IECA. La base de datos contó con 310 participantes para el IECA y 357 para el AWA.

Utilizando los criterios Little Jiffy, el programa SPSS, la matriz de Pearson, la extracción por componentes principales y la rotación Varimax, el AFE del AWA mostró una estructura de tres factores que explico el 45% de la varianza, se debió eliminar 6 ítems. Por su lado el AWA, tras eliminar 3 ítems, mostró una estructura de dos factores que explico el 43% de la varianza, cuando se utilizó Rstudio, la matriz de correlación policórica, la extracción ULS y rotación Oblimin.

Utilizando AMOS PARA SPSS y el criterio de máxima verosimilitud, el AFC del IECA a pesar de la eliminación y reestructuración aconsejada, la estructura no alcanzo el ajuste mínimo por su parte el AFC del IECA utilizando el criterio de mínimos cuadrados no ponderados Diagonalizados y Rstudio alcanzó el ajuste esperado comprobando la estructura tanto de 2 como de 3 factores

Desde la TRI se obtuvo un buen ajuste global al modelo 2PL con la estructura de dos factores. El AFC aconsejo eliminar 4 ítems, por otro lado, el análisis de ítems desde la TRI, evidenció índices inaceptables para dos ítems se sugiere realizar ajustes al instrumento, en particular crear nuevos ítems que discriminen mejor a las personas con mayor trazo latente.

* Trabajo de grado

** Facultad De Ciencias. Escuela De Matemáticas. Especialización En Estadística. Director: Henry Sebastián Rangel Quiñonez, Magíster en Ciencias Estadísticas

Abstract

Title: Comparison of Statistical Techniques for the Validation of Tests in Psychology*

Author: Elda Carolina García Tobo**

Keywords: CFA, EFA, TRI, categorical items, validation, psychology.

Description

This investigation compared within the Statistical Techniques EFA, CFA and TRI the results obtained when selecting certain criteria, comparing between what psychologists generally perform and what statistics recommend to validate psychological tests. The data obtained from the AWA Attitude Scale towards Animal Welfare and the empathy index for children and adolescents ACEI test were analysed. The database had 310 participants for the IECA and 357 for the AWA.

Using the Little Jiffy criteria, the SPSS program, the Pearson matrix, the extraction by principal components and the Varimax rotation, the AFA of the AWA showed a three-factor structure that explained 45% of the variance, 6 items had to be eliminated. For its part the AWA, after eliminating 3 items, showed a two-factor structure that explained 43% of the variance, using Rstudio, the polychoric correlation matrix, the ULS extraction and Oblimin rotation were used.

Using AMOS PARA SPSS and the maximum likelihood criterion, the CFA of the IECA, despite the recommended elimination and restructuring, the structure did not achieve the minimum adjustment, while the CFA of the IECA using the criteria of unweighted Diagonalized least squares and Rstudio reached the expected adjustment by checking the structure of both 2 and 3 factors

The TRI obtained a good overall fit to the 2PL model with the two-factor structure. The CFA advised to eliminate 4 items, on the other hand, the analysis of items from the TRI, showed unacceptable indices for two items, it is suggested to make adjustments to the instrument, in particular to create new items that better discriminate people with a longer latent line.

* Degree work

** Science Faculty. School of Mathematics. Specialization in Statistics. Director: Henry Sebastián Rangel Quiñonez, Master of Statistical Sciences

Introducción

Con frecuencia, el analista de datos en psicología se enfrenta a situaciones que requieren del análisis de la información proveniente de instrumentos de medición con variables categóricas, ejemplo de ello son los test para medir empatía, personalidad, actitudes e intereses entre otros. Contar con instrumentos de medida robustos en sus propiedades estadísticas, es esencial a la hora de diagnosticar y medir conductas humanas, puesto que aquello permite hacer inferencias sobre el constructo medido que sean válidas y confiables.

El uso de métodos estadísticos multivariados como el análisis factorial exploratorio, de ahora en adelante AFE, análisis factorial confirmatorio, de ahora en adelante AFC y la teoría de respuesta al ítem, de ahora en adelante TRI, han sido el acercamiento para analizar los datos provenientes de los test psicológicos, dado que dichas técnicas permiten al investigador interpretar y visualizar conjuntos grandes de datos, enfocándose principalmente en las relaciones entre variables.

El AFE el AFC y la TRI son técnicas que llevan implícitas la toma de decisiones, durante su implementación existen criterios amplios y variados en torno a los cuáles hay que decidir, que en definitiva tienen impacto en los resultados. Por lo tanto, los trabajos donde el objeto de estudio sean las técnicas en sí mismas, resultan relevantes para dilucidar las consecuencias de elegir entre diferentes criterios y contribuir al buen uso de estas herramientas de análisis en una comunidad académica o de profesionales.

Entre los criterios que deben decidirse a la hora de implementar el análisis están: el software de análisis estadístico a utilizar , la matriz de correlaciones apropiada dado la naturaleza de las

variables medidas, los métodos de rotación de ejes y los criterios para la selección de factores, entre otros.

El presente proyecto propone la comparación de las técnicas para el cálculo de variables latentes como son el AFE, AFC y TRI, en el contexto de análisis de instrumentos psicológicos de medida con datos categóricos. Al interior de cada técnica se compararán algunos métodos usados en Psicología, como por ejemplo en el AFE el uso del paquete estadístico SPSS y la técnica Little Jiffy, resaltando algunas consecuencias negativas de su aplicación. Basados en los hallazgos de las investigaciones en estadística se postularán métodos de análisis alternativos que han probado ser idóneos en el análisis de diferentes instrumentos para finalmente describir las diferencias. (Holgado-Tello, Morata-Ramírez, & Barbero-García, 2018)

A continuación, se presentan los cinco criterios que se pretende debatir, por un lado, los criterios que en Psicología de forma inapropiada han sido comúnmente seleccionados y los que la estadística ha probado apropiado para utilizar. (Domínguez-Lara, Fernández-Arata, Merino-Soto, Navarro-Loli, & Calderón De la Cruz, 2018)

Las comparaciones a continuación relacionadas, se realizarán con datos provenientes de dos instrumentos, que en común miden un constructo subyacente denominado actitud empática. La empatía en general puede conceptualizarse como la disposición afectiva y cognitiva de un individuo a participar de una realidad ajena en la que los sentimientos de otra persona o ser vivo están en el foco de la atención. (Fernández-Pinto, López-Pérez, & Márquez, 2008)

Tabla 1.

Criterios de comparación AFE y AFC

Análisis	Decisión	Uso común Little Jiffy	Recomendación Estadística
AFE-AFC	El programa estadístico	SPSS-AMOS	R studio - Factor
AFE-AFC	La matriz de correlación	Pearson	Policóricas o Tetracóricas
AFE	El método de estimación	Componentes principales (CP)	Mínimos Cuadrados no Ponderados Robustos (RULS)
AFC	El método de estimación	Máxima Verosimilitud (ML)	Mínimos Cuadrados Diagonalizados (DWLS)
AFE	Rotación de ejes	Varimax.	Promin y Oblimin

1. Justificación

El presente trabajo se enfocará en comparar los resultados de las técnicas AFE, AFC y TRI, enfocándose en las decisiones que puede tomar el investigador para llevar a cabo dichos análisis, lo que permitirá mostrar las diferencias o similitudes en los resultados al adaptarse a los criterios aconsejados en las investigaciones basadas en simulación de datos (Ferrando & Lorenzo-Seva, 2014); (Li, 2016); (Orozco, 2010), con el propósito de aportar una guía práctica, de uso para la comunidad de psicólogos que se enfrenta al análisis de datos de test con diferentes escalas, de cómo ajustarse a las recomendaciones que brinda la estadística.

Los motivos que llevan a investigar los efectos de seleccionar entre uno u otro criterio, se basan en que en ocasiones los psicólogos optan por guiarse en lo que comúnmente se hace, sin

profundizar en el componente estadístico y los requisitos que deben presentar los datos entre otros. Algunas investigaciones (Chahín-Pinzón & Briñez, 2015); (Tobar & Rubio, 2014); (Castrillón Moreno, Luna Montaña, Avendaño Prieto, & Pérez-Acosta, 2007), optan por usar el paquete estadístico SPSS para realizar el AFE y su versión AMOS para SPSS para realizar el AFC, sin tener en cuenta que la matriz de correlación utilizada por dicho software no es la recomendada y que trae como consecuencia correlaciones infraestimadas.

Otras investigaciones (Jiménez Ardila, Avendaño-Prieto, Giraldo, & Montañez, 2018); (Tobar, Prieto, López, & Sarmiento-López, 2017) no reportan o reportan tan solo parcialmente, los hallazgos encontrados en el análisis previo de los datos como son la detección y tratamiento de datos perdidos, adicional a la comprobación de los supuestos básicos de cada técnica de análisis multivariante, lo que tiene como consecuencia un ajuste incorrecto de los datos al modelo.

2. Antecedentes

Los referentes acá presentados corresponden a los análisis estadísticos hechos a los instrumentos psicológicos de medida que se abordan en este proyecto, así mismo se presentan investigaciones que presentan estudios similares al propuesto:

Escala de actitud hacia el bienestar animal (AWA): el artículo original con una muestra de 329 participantes tiene como principal conclusión del AFE, que utilizo el paquete estadístico SPSS, la extracción por componentes principales y rotación Varimax es que el AWA es una prueba multidimensional, compuesta por cuatro factores que explican el 43% de la varianza, denominados

(C1) Maltrato animal por placer o por desconocimiento conformado por los ítems 1,2,3,5,6,9 y 17, (C2) ocio con animales conformado por los ítems 4,8,10,11,13,22 y 24 , (C3) Animales de granja conformado por los ítems 7,12,15,19 y 25, finalmente el (C4) Abandono de animales conformado por los ítems 14,16,18,20,21 y 23. (Gil, 2015)

Índice de empatía para niños y adolescentes IECA, el artículo original contó con 258 participantes, en él se propone la creación del cuestionario, para el análisis de correlaciones ítem-total, se utilizó las correlaciones punto biserial y momento del producto de Pearson. Los resultados indicaron que el IECA cumplió con los requisitos mínimos de confiabilidad (test-retest entre 0,74 y 0,86). Estudios posteriores en relación a este mismo índice postulan la estructura de la escala como unidimensional F1= Actitud empática (Bryant, 1982), otros estudios, han mostrado una estructura factorial de dos factores F1= Tristeza empática, F2= Actitud empática (De Wied, y otros, 2007), o tres factores F1= Comprensión de sentimientos, F2= Sentimientos de tristeza, F3= Reacción llorosa (Del Barrio, Aluja, & García, 2004) e incluso de cuatro factores. (Calderón, 2011)

En (2014) Lloret-Segura, Ferreres-Traver, Hernandez-Baeza y Tomas-Marco, presentan la investigación que lleva como título el análisis factorial exploratorio de los ítems: una guía práctica, revisada y actualizada, en ella desaconsejan entre otros el uso del pack “Little Jiffy” dentro del AFE, paquete analizado en el presente proyecto a fin de mostrar lo inapropiado de su uso. Los autores exponen así mismo la importancia de socializar con la comunidad de psicólogos y continuar con proyectos de este tipo, puesto que como lo señalan:

Los investigadores interesados en asuntos sustantivos que utilizan rutinariamente esta técnica permanecen en muchos casos ignorantes de todo ello. Por tanto, actualizar los criterios clásicos para incorporar aquellos más adecuados es una necesidad urgente para hacer investigación de calidad. (Lloret-Segura, Ferreres-Traver, Hernández-Baeza, & Tomás-Marco, 2014, pág. 1151)

En 2013 Freiberg, Stover, de la Iglesia y Fernández, presentan una investigación que lleva como título correlaciones policóricas y tetracóricas en estudios factoriales exploratorios y confirmatorios, en ella exponen el uso inapropiado del paquete de análisis SPSS y AMOS para SPSS, puesto que en ellos se usa el coeficiente de correlación de Pearson para llevar a cabo los AFE y AFC, paquetes utilizados en el presente proyecto a fin de mostrar lo inapropiado de su uso, puesto que como lo señalan:

El incremento de estudios exploratorios y confirmatorios en psicometría obliga a revisar su metodología. Generalmente estos estudios utilizan el coeficiente de correlación r de Pearson, diseñado para variables continuas y extendido posteriormente a las categóricas (dicotómicas o politómicas). Los paquetes estadísticos actuales permiten aplicar procedimientos robustos ideados específicamente para variables categóricas, entre los que se destacan las correlaciones tetracóricas y policóricas, cuya relevancia metodológica radica en que la mayoría de las escalas psicométricas se compone de reactivos dicotómicos y politómico (Freiberg et al., 2013, pp.151)

3. Objetivos

3.1 Objetivo General

Comparar los resultados obtenidos en la validación de dos instrumentos a partir de las Técnicas Estadísticas AFE, AFC y TRI a fin de establecer pautas que faciliten el manejo adecuado de procesos de validación en psicología.

3.2 Objetivos Específicos

Comparar la conformación de los factores en un AFE en el instrumento AWA a partir de la estimación de Componentes Principales y de Mínimos Cuadrados No Ponderados.

Comparar las soluciones encontradas con el AFC para el instrumento IECA a partir de la estimación por Máxima Verosimilitud y Mínimos Cuadrados Ponderados Diagonalizados.

Comparar las estructuras encontradas con el AFC y la TRI dentro del instrumento IECA de dos factores, resaltando similitudes y diferencias.

4. Marco Teórico

4.1 Análisis Factorial Exploratorio (AFE)

El análisis estadístico de test y variables psicológicas que se basan en atributos medidos en escala ordinal y presentan observaciones en formato dicotómico y politómico, data en la ciencia psicológica de los trabajos de Spearman, mismo que establece los fundamentos para el cálculo estadístico de la asociación de variables teniendo en cuentas las particularidades de la medición del comportamiento humano.

El AFE es una técnica de análisis multivariada que permite examinar la interdependencia de variables, explicando a través de factores las interrelaciones entre ellas, en esta técnica las variables originales en el caso de los test psicológicos, los denominados reactivos, ítems o preguntas, son dependientes y se explican por medio de factores no observables. (Aldás & Uriel, 2005)

4.1.1. Objetivo de la técnica

La técnica de análisis factorial para el análisis de los constructos latentes, medidos por medio de variables observadas (ítems), es en la actualidad el procedimiento estadístico más usado a fin de validar o adaptar instrumentos psicológicos, no obstante, su uso se ha extendido a otras áreas. Existe dos tipos de análisis factoriales: el exploratorio y el confirmatorio, el primero se aconseja cuando hay poco o nulo conocimiento sobre la estructura de los datos, el segundo cuando existen hipótesis teóricas previas que dan cuenta de la estructura subyacente a los datos analizados (Martínez & Rondon, 2012); (Rodríguez-Jiménez, Rosero-Burbano, Sanabria, & Mateus, 2011).

Desde la época de Spearman hasta la actualidad, múltiples desarrollos se han dado (Baglin, 2014; (Barajas Marcos & Prieto Alaiz, 2015); Cuadras, 2007). Dichas actualizaciones tienen por objetivo generar soluciones estables y ajustadas a la naturaleza de los datos. No obstante, los avances en las técnicas de análisis y los softwares disponibles, las investigaciones muestran que las recomendaciones son ignoradas, y que no todos los artículos publicados en revistas particularmente no indexadas, siguen las recomendaciones en torno al uso correcto del AFE y AFC. (Lloret-Segura, Ferreres-Traver, Hernández-Baeza, & Tomás-Marco, 2014)

4.1.2 Modelo

El modelo asume variables aleatorias observables ($X_{i_1} \dots X_{i_p}$) y variables aleatorias latentes ($F_1 \dots F_k$), la representación algebraica del modelo $k \leq p$ versa así (1).

$$X_{ij} = V_{j_1} F_1 + V_{j_2} F_2 + \dots + V_{p_k} F_k + e_p. \quad (1)$$

$j =$ Individuo

$i =$ Variable observada – Item

$X_{ij} =$ Puntuación del individuo j en la variable i

$V_{pk} =$ Carga factorial – peso del factor común F asociado al ítem i

$F_k =$ Factor común – constructo latente

$e_p =$ Factor específico – término de error estocástico no observado

Para comprender los resultados de un análisis factorial y sus diferencias con otros métodos de reducción de dimensiones, conceptos como, comunalidad y unicidad son esenciales, las aproximaciones clásicas han mostrado soluciones equivalentes entre el AFE y el análisis por componentes principales, en casos donde la unicidad es mínima y el número de ítems por factor es alto. (Velicer & Jackson, 1990); (Lloret-Segura, Ferreres-Traver, Hernández-Baeza, & Tomás-Marco, 2014).

4.1.3 Supuestos del modelo y validación

El AFE permite extraer información relevante de los datos disponibles, para entender las relaciones existentes entre las variables analizadas, no obstante, lo anterior solo puede darse en el contexto del uso adecuado de la técnica, solo en dicho caso la información extraída permitirá llegar a conclusiones razonables. Por tal motivo la literatura plantea una serie de pasos y supuestos que deben llevarse a cabo para su aplicación correcta:

El investigador analiza los datos para conocer si los mismos cumplen los supuestos básicos del AFE. Entre los supuestos se deben detectar la escala de las variables, tratar casos faltantes, se

examina así mismos los valores de la matriz de correlaciones, la presencia o no de normalidad multivariante para el caso de que el método de extracción sea Máxima verosimilitud, el tamaño y homogeneidad de la muestra, presencia de linealidad, el contraste de esfericidad de Barlett y la medida de adecuación muestral de Kaiser, Meyer y Olkin KMO.

Las técnicas multivariantes AFE y AFC exigen la detección y posterior tratamiento de casos faltantes y datos atípicos, la detección debe hacerse siguiendo el análisis del caso atípico, respecto al conjunto de variables que se incorporan en el análisis y no respecto a una sola de ellas, el procedimiento recomendado es calcular la distancia de cada caso al centroide del conjunto de datos. (Aldás & Uriel, 2005)

Dentro de los análisis descriptivos univariados que son recomendados, previo al uso del AFE, AFC y TRI, se recomienda el análisis de los coeficientes de asimetría y curtosis. En el caso que la asimetría tenga un valor absoluto mayor que 1, el analista se encontrará frente a una distribución muy sesgada, si la asimetría está en valor absoluto entre 1 y 0.5, la distribución está moderadamente sesgada, finalmente si la asimetría está en valor absoluto entre 0.5 y 0, la distribución es aproximadamente simétrica. En el caso de distribuciones acusadamente asimétricas la interpretación de la curtosis no debe realizarse.

Según Lloret-Segura, et al. (2014) en relación con el análisis de la normalidad univariada de los ítems se postulan que “Algunos autores recomiendan las distribuciones con coeficientes de asimetría y curtosis en el rango (-1,1)... Otros en cambio consideran aceptables valores en el rango (-1.5, 1.5)..., o incluso el rango [-2, 2]” (p.1158).

No obstante el análisis univariado es importante, el análisis del coeficiente de asimetría y curtosis para las técnicas multivariadas, prioriza el contraste de la normalidad multivariante, en la literatura el contraste de Mardia-curtosis y Mardia-apuntalamiento (*Mardia's test*), el test Henze-

Zirkler's y el test *Royston's*, son recomendados para el análisis multivariante de la normalidad (Ho: Los datos siguen una Distribución Normal Multivariada), en la actualidad el programa estadístico SPSS no cuenta con la opción para probar normalidad por medio de un test, solo ofrece la posibilidad de realizar gráficos Q-Q, la recomendación para el análisis multivariante de normalidad indica que los gráficos deben acompañarse usualmente de la información proporcionada por los test. (Korkmaz, Goksuluk, & Zararsiz, 2014)

La información proporcionada por los descriptivos univariantes y el test de normalidad, deberá ser tomada en cuenta por el analista para decidir el método de extracción de factores a utilizar, métodos de máxima verosimilitud ML, suele requerir normalidad multivariada a fin de obtener resultados ajustados. Cuando se viola el supuesto de normalidad, uno de los métodos más empleados es el de mínimos cuadrados no ponderados (ULS) y el de mínimos cuadrados no ponderados robustos (RULS). (Lloret-Segura, Ferreres-Traver, Hernández-Baeza, & Tomás-Marco, 2014)

4.1.3.1 Escala de las variables. Los ítems son el insumo más importante del test, de su ajuste dependerá en gran medida el éxito de la estructura factorial resultante. Diversos aspectos en la creación de los ítems deben ser analizados:

El número de alternativas de respuesta, puede usualmente variar entre pruebas dicotómicas y politómicas desde tres hasta cinco opciones de respuesta. De manera clásica, entre las pruebas politómicas es común utilizar la alternativa de respuesta neutral (no sabe/no responde, ni en desacuerdo/ni de acuerdo). Este aspecto es definitivo para escoger el tipo de matriz de correlaciones que ha de usarse dentro del análisis factorial, matriz de correlaciones de Pearson, Policóricas o Tetracóricas. (Barajas Marcos & Prieto Alaiz, 2015).

Algunos test suelen utilizar preguntas de control, ítems que miden lo mismo con ligeros cambios en su redacción, por otro lado, la recomendación clásica, implica tener como mínimo tres ítems por factor. Estos aspectos suelen influir en la estabilidad de la solución, el número de factores resultantes del análisis y la comunalidad expresada por los ítems. (Kline, 1997)

Dependiendo del número de opciones cinco como mínimo y la normalidad de los ítems (curtosis y asimetría aproximadamente cero), de manera clásica se han igualado las soluciones de la matriz de Pearson con la Policórica o Tetracórica. Este punto de análisis es crucial y es uno de los criterios de decisión a los que ha de enfrentarse el investigador. (Carretero-Dios & Pérez, 2005)

4.1.3.2 Tamaño de muestra. El cálculo del tamaño de muestra no sigue en la mayoría de los casos un procedimiento único, algunos estudios de validez usan un tamaño de muestra mínimo $n \geq 300$ que se considera bueno, otros por su lado utilizan el criterio n/p siendo n el número de sujetos y p el número de ítems, el criterio postula medir como mínimo 10 sujetos por cada ítem del cuestionario $n \geq 10 \text{ veces } p$, si el resultado de este procedimiento da menor a 100 sujetos el criterio obliga a medir como mínimo 100 personas $n \geq 100$. Finalmente, la matriz de correlación utilizada en el análisis es uno de los criterios que analizan otros investigadores para determinar el tamaño de muestra. (Carretero-Dios & Pérez, 2005)

Con matrices de correlaciones Policóricas, el tamaño de muestra debe ser mayor que el necesario para la matriz de correlación de Pearson. Entre más homogénea la muestra usada para el análisis en la versión original del instrumento, menor el tamaño necesario para su proceso de adaptación, en relación con la comunalidad, cuando las saturaciones $>.70$ un tamaño de muestra de 100 a 150 casos es suficiente, comunalidades de $.40$ a $.70$ requerirán tamaños de muestra de 200, finalmente, comunalidades de $.30$ e inferiores requerirán como mínimo muestras de 400

casos. Se explicará a continuación la importancia de la selección adecuada de la matriz de correlación en el análisis factorial.

4.1.3.3 Matriz de correlación. Para que el análisis factorial sea razonable, es importante que la matriz de correlación contenga grupos de variables que se correlacionen $r > 0,3$ y al menos un coeficiente significativo $p < 0,05$. Una matriz de correlaciones próxima a la matriz identidad indica que el análisis factorial conduciría a una solución deficiente.

Dentro del paquete estadístico SPSS por defecto el análisis factorial utiliza la matriz de correlación de Pearson, el programa estadístico Factor y Rstudio, permite al usuario seleccionar la matriz que sea apropiada dada la naturaleza de los datos analizados.

La Matriz de correlación de Pearson es la más conocida, el coeficiente de correlación asociado, se utiliza cuando las variables a analizar son continuas, los datos son obtenidos en forma de pares independientes.

$$r = \frac{\sum xy}{N\sigma_x\sigma_y} \quad (2)$$

x = Desviación de la puntuación X respecto a su media

y = Desviación de la puntuación Y respecto a su media

σ_x = Desviación estandar de la distribución de puntuaciones X

σ_y = Desviación estandar de la distribución de puntuaciones Y

N = Numero de participantes

La matriz de correlación tetracórica, se utiliza cuando las variables a analizar son dicotómicas, los datos son obtenidos en forma de pares independientes.

$$r_t = \cos \left[\frac{180^\circ \sqrt{bc}}{\sqrt{ab} + \sqrt{bc}} \right];$$
$$r_t = \cos \left[\frac{180^\circ \sqrt{ad}}{\sqrt{ad} + \sqrt{bc}} \right] \quad (3)$$

En la fórmula las notaciones son conocidas, a,b,c,d son las frecuencias de la tabla de contingencia. En el caso en el que el producto bc sea superior al producto ad, entonces la expresión a utilizar será aquella presentada después del punto y coma. (Palmer, Jiménez, & Montaña, 2000)

La matriz de correlación policórica, que constituye una generalización de la correlación tetracórica, se utiliza cuando las variables a analizar son ordinales, los datos son obtenidos en forma de pares independientes.

A fin de estimar el coeficiente de correlación policórica, con las variables ordinales se construye una tabla de contingencia, para la tabla de contingencia la correlación policórica es una medida que se obtiene a partir de la distribución acumulada normal bivariada, dicha distribución posteriormente se deriva y se estiman simultáneamente los parámetros, debido a la complejidad de estimar los parámetros para calcular el coeficiente de correlación policórico, su cálculo se resuelve por medio de la implementación de software computacionales.

Los resultados en simulación de datos, han concluido que al trabajar con datos que incluyan variables categóricas, a fin de robustecer las estimaciones y obtener el máximo porcentaje de explicación de la varianza, es recomendable usar matrices de correlación policóricas cuando las variables son ordinales y tetracóricas cuando son dicotómicas.

Una vez se ha estimado la matriz de correlación apropiada para el análisis de las relaciones entre las variables, la matriz obtenida debe ser analizada para ello debe presentarse el determinante de la matriz, el contraste de esfericidad de Bartlett y índice de medida de suficiencia de muestreo KMO, mismos que se explican a continuación.

En relación con el valor del determinante de la matriz de correlaciones como herramienta que permite clasificar los sistemas de ecuaciones lineales según sus soluciones, la literatura postula que para continuar con el análisis el mismo debe aproximarse a cero más no debe ser cero, lo que supondría la existencia de multicolinealidad.

En relación con el test Bartlett cuyo propósito consisten en saber si podemos factorizar las variables originales de forma eficiente, la literatura postula que se debe rechazar la hipótesis nula, la matriz de correlaciones es igual a la matriz identidad, puesto que de no ser así, no existiría correlaciones significativas entre las variables y el modelo factorial no sería pertinente.

Finalmente, en la literatura el índice de medida de suficiencia de muestreo KMO, que debe ser analizado de manera global, así como para cada variable, debe presentar valores iguales o superiores a 0,7 para sugerir que cada variable puede ser predicha con un mínimo de error por las otras variables que conforman la matriz.

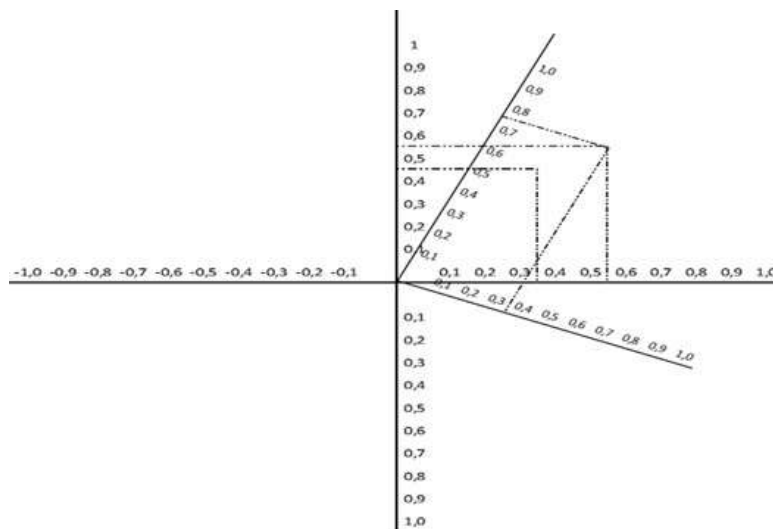
4.1.4 Ajuste del AFE

Una vez se ha verificado que los datos no violan los supuestos necesarios para llevar a cabo un análisis factorial, el análisis de datos se enfrenta ahora a la tarea de seleccionar el tipo de rotación de factores.

4.1.4.1 Rotación de ejes. El propósito del análisis de factores es explicar y dar cuenta de las relaciones observadas en los datos, se trata que los factores comunes tengan una interpretación e identificación clara, a fin de poder hacerlo los factores deben ser rotados a partir de la solución inicial.

Cuando se estiman los factores en el análisis y los mismos son rotados, su interpretación puede entenderse si se conceptualizan gráficamente. Los factores y sus cargas pueden representarse en un espacio euclidiano, las variables (ítems) pueden caer en cualquiera de los cuatro cuadrantes. Los ejes pueden ser rotados en infinitas posiciones, dicha posibilidad da cuenta de las diferentes soluciones que pueden estimarse en el análisis factorial. (Kline, 1997)

Figura 1. Rotación de los ejes en el espacio euclídeo



Adaptada de (Kline, 1997)

La rotación de los factores, cambia las cargas factoriales, al hacer que cada una de las variables tenga una correlación lo más próxima a uno, con uno de los factores y próximas a cero con los demás. No obstante, lo anterior cabe destacar que la varianza explicada es la misma,

independiente de la rotación, más aún, una solución rotada reproduce las correlaciones originales tal como lo hace una solución no rotada. (Aldás & Uriel, 2005)

La solución es rotada con el propósito de lograr la mayor simplicidad e interpretabilidad. Existen diversos métodos de rotación descritos en la literatura que se agrupan en dos formas básicas, la ortogonal y la rotación oblicua, cada uno de ellos presenta características particulares en torno a la forma en que las variables son rotadas, ellos tienen implicaciones importantes en la interpretación de la solución obtenida.

La rotación ortogonal minimiza el número de variables que tienen cargas altas en cada factor, maximizando la suma de varianzas de las cargas factoriales al cuadrado dentro de cada factor, los ejes se rotan de forma tal que siempre formen ángulos rectos unos con otros, lo que en términos generales implica que los factores no están correlacionados entre sí. (Ferrando & Lorenzo-Seva, 2014)

Algunos ejemplos de rotaciones ortogonales utilizados con mayor frecuencia en la literatura son, la rotación Varimax y la Quartimax, como regla general, cuando se selecciona el criterio de rotación Varimax, el mismo se acompaña de la normalización de Kaiser, lo que evita que las variables con mayores comunalidades, tengan mayor influencia en la solución final, la rotación Varimax se aconseja utilizar cuando no hay un factor dominante, por otro lado cuando se espera un solo factor general se aconseja utilizar la rotación Quartimax (Kline, 1997); (Lloret-Segura, Ferreres-Traver, Hernández-Baeza, & Tomás-Marco, 2014).

En la rotación oblicua los ejes pueden tomar cualquier posición en el espacio, este tipo de rotación se enfoca en maximizar la simplicidad de la solución rotada, esto posibilita tener en cuenta la posible relación de dependencia entre los factores. No obstante, si los factores son en realidad

de naturaleza independiente, la solución factorial rotada por rotación oblicua mostrará correlaciones entre factores cercanos a cero. (Ferrando & Lorenzo-Seva, 2014)

Según Martínez y Rondón (2012), algunos ejemplos de rotaciones oblicuas utilizados con mayor frecuencia son la oblimin, promax y promin, la oblimin permite establecer relaciones jerárquicas entre los factores, la promax modifica los resultados de una rotación ortogonal hasta crear una solución con cargas factoriales lo más próximas posible a la ‘estructura ideal’, finalmente la promin se aconseja cuando un pequeño número de factores tiene saturaciones mayores a la unidad” (p.205).

4.2 Análisis factorial confirmatorio (AFC)

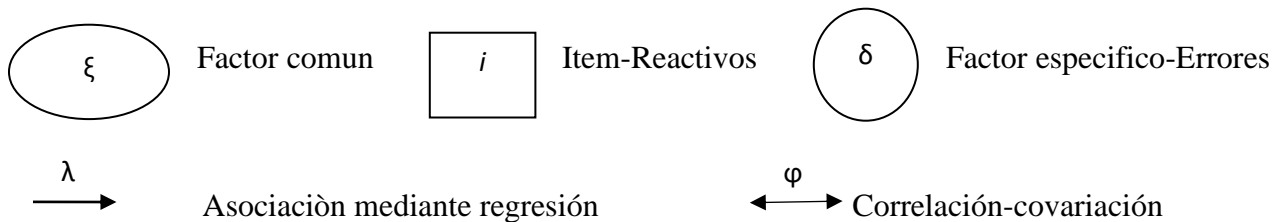
El AFC es una técnica de análisis multivariado desarrollado por Joreskog 1973, es una técnica de análisis estadístico que posibilita probar hipótesis. Basándose en la información teórica o en trabajos previos, el constructor del test hipotetiza las cargas factoriales que son puestas a prueba, dichas hipótesis se ajustan en una matriz de datos que debe superar una serie de indicadores críticos para lograr la adecuación entre los datos observados y la matriz esperada (Pilatti, Godoy & Brussino, 2011; Ronnback et al., 1999).


4.2.1 Objetivo de la técnica


El AFC tiene por objetivo probar hipótesis en relación a la carga y estructura factoriales de un instrumento de medida, dichas hipótesis se ajustan en una matriz de datos que debe lograr una adecuación entre los datos observados y la matriz esperada, el AFC utiliza métodos iterativos para obtener el mejor ajuste posible. (Kline, 2014; Pilatti, Godoy & Brussino, 2011; Ronnback et al., 1999).


4.2.2 Modelo


El AFC generalmente se presenta por medio de un diagrama de flujo o *pathdiagram*, que presenta los siguientes símbolos:



 Variable independiente o factor (constructo latente): son la fuente de causalidad en el modelo, no tienen flechas dirigidas hacia ellos, sus flechas de conexión son siempre hacia afuera, unidireccionales cuando van hacia la variable dependiente, la dirección de la flecha indica que es el factor latente, aquel que genera las puntuaciones en el ítem.

 Variable dependiente o ítem (variables observadas): son causalmente dependientes, tienen flechas apuntando directamente a ellas, los ítems comparten algo de un constructo latente cuando todos tienen una flecha que los conecta con un determinado factor.

 Todas las variables dependientes están interconectadas por líneas de correlación.

 Las variables dependientes, nunca están interconectadas por líneas de correlación, si esta correlación existe, será indicada por líneas de correlación entre los residuales-errores, que dirigen una flecha en sentido izquierda a derecha hacia un ítem.

4.2.3 Ajuste AFC

Según Ruiz, Pardo y San Martín (2010), existen diversos indicadores para probar el modelo de ajuste a los datos observados, cuatro de éstos que son:

La razón de chi-cuadrado sobre los grados de libertad (CMIN/DF), el indicador debe ser cercano a uno para modelos con ajuste excelente, valores inferiores a cinco muestran un modelo aceptable, valores entre los rangos de uno a tres indican modelos óptimos y valores inferiores a dos y cercanos a uno son indicativos de una adecuación excelente.

El índice de ajuste comparativo (CFI), indica la proporción de covariación entre los ítems explicada por el modelo, presentan valores de cero hasta uno, valores cercanos o iguales a 0 indican ausencia de ajuste, valores cercanos o superiores a 0,95 se consideran óptimos, y cercanos a 0,90 indican ajuste aceptable e iguales a uno ajuste es excelente.

El índice de bondad del ajuste global (GFI), presentan valores de cero hasta uno, valores cercanos o iguales a 0 indican ausencia de ajuste, valores cercanos o superiores a 0,95 se consideran óptimos, y cercanos a 0,90 indican ajuste aceptable e iguales a uno ajuste es excelente.

El error cuadrado de aproximación (RMSEA), error de aproximación del modelo, valores cercanos a 0.05 o menores indican ajuste óptimo, cercanos a 0.08 muestran errores de aproximación indicando que el modelo no debe ser usado para generar ninguna estimación.

El AFC consisten en varias etapas, una de ellas es la fase de estimación del modelo, la función de ajuste, nos indica que de un infinito número de matrices estimadas que cumplen las restricciones del modelo planteadas en la lista anterior, si matriz de varianzas-covarianza estimada Σ es la más próxima a la matriz de varianzas y covarianzas observada S , entonces las estimaciones de los parámetros contenidas en Σ serán razonables en el sentido de ser consistentes con los datos muestrales S

$$F = [S; \Sigma(p)]' \quad (4)$$

F = Funcion de ajuste

S = Matriz de varianzas y covarianzas observada

Σ = Matriz de varianzas y covarianzas estimada

p = Orden de la matriz input

Dada una matriz de varianzas y covarianzas muestrales, se estiman los parámetros del modelo factorial hipotetizado, se evalúa el ajuste del modelo, determinando en qué medida la matriz estimada se aproxima o ajusta a la muestra. En esta función de ajuste, $[S - \Sigma(p)]$ la matriz residual se obtiene después de calcular la diferencia entre dichas dos matrices: la matriz de varianzas y covarianzas S de la muestra y la matriz de varianzas y covarianzas $\Sigma(p)$ estimada, a partir de los parámetros del modelo”. (Holgado-Tello, Morata-Ramírez, & Barbero-García, 2018)

Existen diversos procedimientos de estimación de la función de ajuste, se describirán el de uso común, ML y el recomendado por estudios de simulación de data para datos categóricos RULS y DWLS.

La estimación por máxima verosimilitud ML implica minimizar la siguiente función.

$$F_{ML} = tr(S\Sigma^{-1}) + [\log|\Sigma| - \log|S|] \quad (5)$$

La verosimilitud parte de una muestra observada, o un resultado obtenido y nos da una medida de lo probable que es observar en una muestra aleatoria lo que se ha obtenido, si la muestra proviniera de una familia paramétrica (distribución) con unos determinados parámetros, es decir

una medida de lo creíble que es observar, lo observado si la muestra proviene de una población que se distribuye según unos parámetros determinados. (Aldás & Uriel, 2005)

El ML asume que las variables observadas están medidas en una escala de medida de intervalo, basan sus cálculos en la matriz de correlación de Pearson, los datos analizados deben cumplir el requisito de normalidad multivariante y funciona mejor en muestras de tamaño grande.

La estimación por Mínimos Cuadrados no Ponderados Diagonalizados DWLS implica minimizar la siguiente función

$$F_{DWLS} = [S - \Sigma(p)]' (W_D)^{-1} [S - \Sigma(p)] \quad (6)$$

Es una variante robusta de WLS, trabaja con una matriz de correlaciones policóricas que es la base para obtener la matriz AC a partir de entonces participa en la matriz W distribuida libremente, ya que las variables observadas no necesariamente deben cumplir la asunción de normalidad multivariante.

4.3 Teoría de respuesta al ítem

En términos generales, se asume que la experiencia de acontecimientos (como el individuo siente y experimenta el mundo psíquico) es continua. No obstante, en el momento en que se nota dicha experiencia y se intenta caracterizarla, lo general es hacerlo aludiendo a categorías, ejemplo: si ese hombre es o no es ($x = 0,1$); si la experiencia es muy fuerte, moderadamente fuerte o muy débil ($x = 0,1,2$), o finalmente si algo es fuertemente agradable, agradable, desagradable o fuertemente desagradable ($x = 1,2,3,4$). Como puede notarse, si deseamos describir las

particularidades de una experiencia, generalmente se hace de forma categórica, nominal u ordinalmente.

Considerando lo anterior, el procedimiento de asignar números a las categorías es lo común. Cabe destacar que, aunque la naturaleza matemática de los números enteros en efecto comprueba que la distancia entre cada par de números, siempre es la misma, para el caso concreto, los números son tan solo una representación que cumple la función de una etiqueta, entonces jamás se deberá entender dichos números como igualmente distribuidos, por tal motivo cálculos estadísticos como la media, desviación estándar, el coeficiente de correlación entre otros, no son adecuados de estimar usando dichos números.

La familia de los modelos TRI provee las bases para construir medidas de intervalo para el tipo de datos anteriormente descritos. A fin de elaborar dicha métrica, los datos categóricos, representados por medio de etiquetas numéricas, deben someterse a un proceso estocástico, que caracterizará matemáticamente dichas variables aleatorias, y su evolución, para que los datos puedan una vez transformados, servir para realizar procesos de inferencia en torno a ellos. (Wright & Stone, 2004)

Dentro de los modelos TRI se encuentra el que es denominado por muchos como el modelo más parsimonioso, el análisis Rasch que presenta asimismo estimaciones para respuestas dicotómicas y politómicas, dentro del modelo dicotómico, la probabilidad de una respuesta es modelada como una función logística de la diferencia entre el parámetro de ítem y la persona. El modelo presupone varios aspectos a saber, unidimensional, independencia local e invariabilidad y objetividad específica de la estimación. (Orozco, 2010). A continuación, se presenta este modelo:

$$P\{x_{ni}|B_n, D_i\} = \frac{e^{(B_n - D_i)}}{1 + e^{(B_n - D_i)}}$$

$i = 1, 2, \dots, k$,

$x = 0, 1$, o su equivalente,

$$\ln \left[\frac{P_{ni}}{1 - P_{ni}} \right] = B_n - D_i \quad (7)$$

x_{ni} = Respuesta de un sujeto ante una pregunta

$P_{ni}(x = 0, 1)$ = Probabilidad de una respuesta incorrecta ($x = 0$) o correcta ($x = 1$)

B_n = Habilidad de la persona

D_i = Dificultad del ítem, denotada en algunos textos como β

El análisis del patrón de respuestas para un sujeto X_{ni} se usa como evidencia (replicas) de la medida que se desea abstraer B_n , dicha medida depende de la dificultad del ítem D_i , el modelo recopila los fracasos y éxitos, entendidos como eventos gobernados por un proceso probabilístico. (Wright & Mok, 2004)

El análisis de un ítem contemplará dentro del modelo dicotómico diversas asunciones, relacionadas todas con el análisis de las probabilidades de éxitos y fracasos en las respuestas de los sujetos frente a las preguntas en los test.

Si la probabilidad de éxito P_{ni1} es igual a la probabilidad de fracaso P_{ni0} entonces la habilidad y la dificultad serán iguales $B_n = D_i$ (González, 2008); (Linacre, 2012).

$$\begin{aligned}
 P_{ni1} &= P_{ni0} = 0,5 \\
 \log e\left(\frac{P_{ni1}}{P_{ni0}}\right) &= \log e\left(\frac{0,5}{0,5}\right) \\
 \log_e(1) &= 0 = B_n - D_I \\
 B_n &= D_i \qquad (8)
 \end{aligned}$$

Si la probabilidad de éxito P_{ni1} es uno y la de fracaso P_{ni0} es 0, entonces la habilidad en lógitos tenderá a más infinito en relación a cualquier ítem de dificultad finita y la diferencia entre la habilidad y la dificultad será positiva y mayor a 0.5 (González, 2008); (Linacre, 2012).

$$\begin{aligned}
 P_{ni1} &= 1, P_{ni0} = 0 \\
 \log e\left(\frac{P_{ni1}}{P_{ni0}}\right) &= \log e\left(\frac{1}{0}\right) \\
 \log_e(\infty) &= \infty = B_n - D_I \\
 B_n &= \infty \qquad (9)
 \end{aligned}$$

Si la probabilidad de éxito P_{ni1} es 0 y la probabilidad de fracaso P_{ni0} es 1, la habilidad en lógito tenderá a menos infinito en relación a cualquier ítem de dificultad finita, y entonces la diferencia entre la habilidad y la dificultad será negativa y menor a 0,5 (González, 2008); (Linacre, 2012).

$$\begin{aligned}
 P_{ni1} &= 0, P_{ni0} = 1 \\
 \log e\left(\frac{P_{ni1}}{P_{ni0}}\right) &= \log e\left(\frac{0}{1}\right) \\
 \log_e(0) &= -\infty = B_n - D_I \\
 B_n &= -\infty \qquad (10)
 \end{aligned}$$

Existen diferentes aproximaciones del modelo entre las que se encuentran, dentro del modelo dicotómico los llamados 1PL, 2PL y 3PL, al interior del modelo politómico con variable categórica se encuentran los modelos de crédito parcial, el crédito parcial generalizado y el modelo de respuesta gradual, entre otros. Dada las características de la prueba a analizar, el test IECA prueba dicotómica, con opciones de respuesta 0 y 1, se desarrollará teóricamente el modelo dicotómico de dos parámetros.

Modelo dicotómico 2PL

La primera generalización del modelo Rasch de un parámetro se conoce como modelo 2PL o modelo de dos parámetros (dificultad D_i y discriminación α_i).

$$P\{x_{ni}|B_n, D_i\} = \frac{e^{\alpha_i(B_n - D_i)}}{1 + e^{\alpha_i(B_n - D_i)}}$$

$$i = 1, 2, \dots, k,$$

$$x = 0, 1, \text{ o su equivalente,}$$

$$\ln \left[\frac{P_{ni}}{1 - P_{ni}} \right] = B_n - D_i \quad (11)$$

x_{ni} = Respuesta de un sujeto ante una pregunta

$P_{ni}(x = 0, 1)$ = Probabilidad de una respuesta incorrecta o correcta

B_n = Habilidad de la persona, denotada en algunos textos como θ

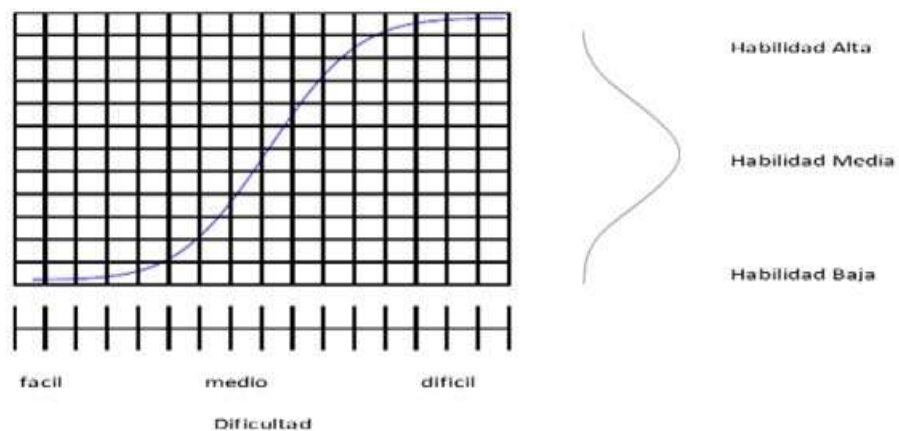
D_i = Dificultad del ítem, denotada en algunos textos como β

α_i = Discriminación del ítem

Curva característica del ítem

La curva característica del ítem es una representación gráfica de la discriminación y dificultad del ítem combinadas, generalmente dicha grafica muestra la forma de una S, su orientación creciente refleja el cambio en la probabilidad de respuesta a medida que la habilidad de quien responde cambia D_i , se asume bajo el modelo que a mayor habilidad la dificultad del ítem se torna menor, pues la probabilidad de respuesta aumenta al aumentar la habilidad.

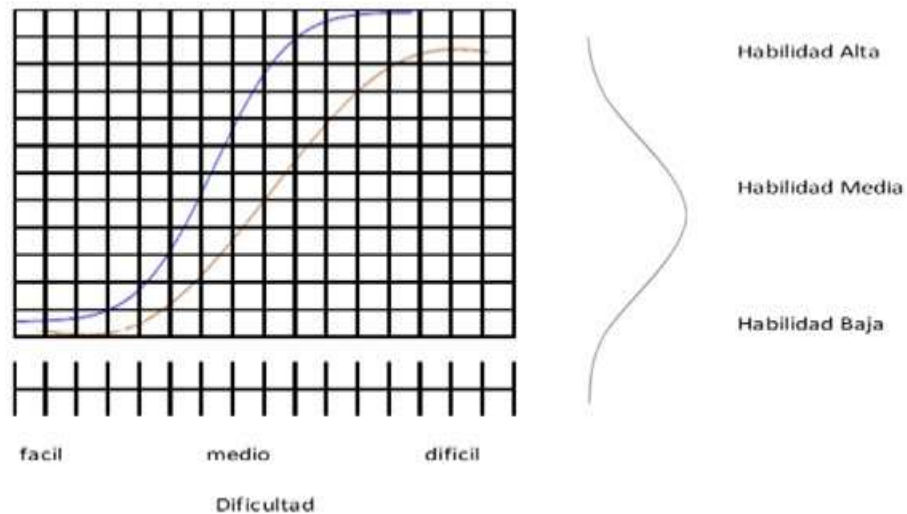
Figura 2. Curva característica del ítem



Adaptado de Tristán, (1998)

En el modelo 2PL, la discriminación es observada por la pendiente, la dificultad entonces corresponderá a la localización, en la gráfica se observan dos ítems que presentan la misma discriminación (misma pendiente) pero dificultad (localización) diferente.

Figura 3. CCI comparación parámetro dificultad



Adaptada de Tristán, (1998)

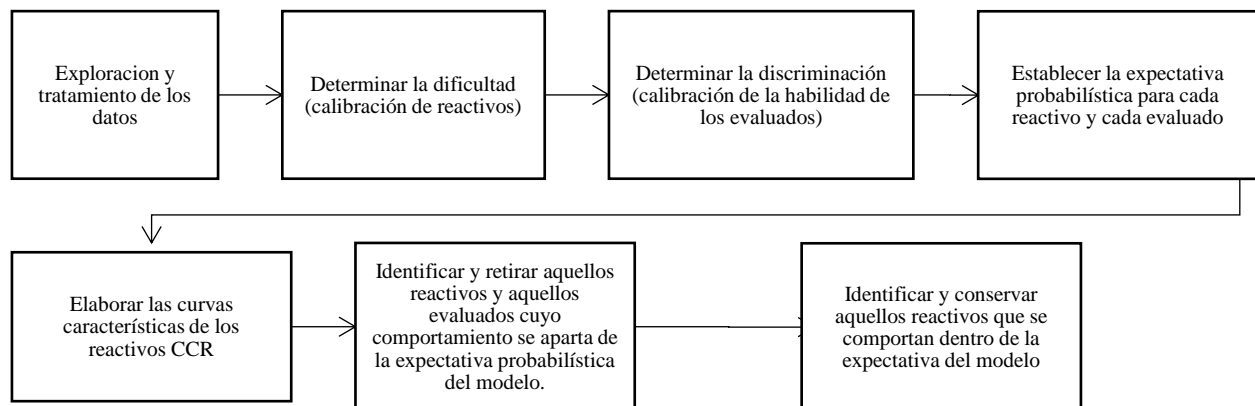
Función de información

La función de información, confiabilidad en términos clásicos, expresa la cantidad de información que proporciona el ítem y el test respecto al rasgo latente, indica que tan bien está estimada la habilidad a lo largo de todos los niveles del ítem, el máximo nivel de información para el modelo de dos parámetros se observa cuando la habilidad y la dificultad son iguales.

El modelo TRI, asume ciertos criterios que permiten evaluar el ajuste de los ítems en la evaluación del rasgo latente, en relación con la discriminación α_i se espera que la misma sea siempre positiva y diferente de cero. Una discriminación de cero implica que la habilidad no juega un papel importante para responder el ítem, valores negativos son muestra que el ítem está funcionando inapropiadamente, pues a medida que decrece la habilidad hay más probabilidad de contestar el ítem. En relación con la dificultad del ítem se espera que la misma fluctúe en el rango de -3 a 3, lo que indicaría un ajuste adecuado.

A manera de síntesis y de acuerdo con lo expuesto por González (2008), el procedimiento para análisis TRI corresponde a una sucesión de pasos que de forma organizada permitirán seleccionar la forma final del instrumento analizado.

Figura 4. Pasos de análisis y selección bajo el modelo TRI



5. Metodología

5.1 Descripción de la población y muestra

El presente proyecto descriptivo, contó con una muestra recogida durante el I y II semestre del 2019, incluyó un marco poblacional depurado de 6.992 participantes entre estudiantes y padres de familia, se descartó incluir en las mediciones a los trabajadores de las instituciones educativas. Un total de 1.402 participantes de las provincias Comunera y Guanentá fueron escogidos por medio de muestreo aleatorio, se redujo posteriormente a 667, pues no todos consintieron la aplicación o diligenciaron los cuestionarios. La muestra final quedó constituida por un total de

participantes distribuidos así, 310 participantes conforman la base de datos de la prueba IECA y 357 participantes la de la prueba AWA.

5.2 Métodos

Las bases de datos que se usaran para el análisis corresponden a los resultados obtenidos de tres instrumentos de medida psicológicos ellos son:

IECA constituido por 22 ítems con escala de respuesta dicotómica (sí/no), el 50% de los ítems se califican en orden directo (1/0) y 50% en orden inverso (0/1). La prueba está diseñada para ser aplicada a niños y adolescentes con un rango de edad comprendido entre 6 a 13 años.

AWA constituido por 25 ítems con escala de respuesta tipo Likert con cinco opciones de respuesta bajo el estilo criterio adhesión (Muy de acuerdo, De acuerdo, Indiferente, En desacuerdo, Muy en desacuerdo), 10 de los ítems se califican en orden directo (1, 2, 7, 8, 9, 10, 13, 21, 22 y 24) y 15 en orden inverso (3, 4, 5, 6, 4, 11, 12, 14, 15, 16, 17, 19, 20, 23 , 25). La prueba está diseñada para ser aplicada a adolescentes y adultos con un rango de edad comprendido entre 11 a 41 años.

Los investigadores, entre los que se incluye la autora de este proyecto, contaron con el permiso de cada institución educativa, de los padres o representantes de los menores participantes y de cada participante mayor de 18 años, para llevar a cabo las mediciones en 13 instituciones educativas de la provincia Guantán y comunera. Las instituciones que se mostraron inicialmente interesadas en participar en la investigación son de carácter tanto público como privado y pertenecen algunas al sector urbano y otras al rural.

Una vez obtenido los permisos de cada institución y de los participantes, se procedió a aplicar los consentimientos informados, una vez se obtuvo consentimiento, se procedió a aplicar

la encuesta sociodemográfica y diferentes pruebas entre las que se encuentran las que constituyen las dos bases de datos a utilizar en el presente proyecto.

Se realizará el análisis univariado de los ítems de cada uno de los instrumentos AWA e IECA, posteriormente se realizará la comparación propuesta en los objetivos así: AFE para la prueba AWA, AFC, luego TRI se utilizará para la prueba IECA. Se comparará los resultados obtenidos de los programas SPSS, Factor y Rstudio, la matriz de correlación de Pearson frente a las policóricas y tetracóricas, el ajuste bajo el método de componentes principales en contraposición al de mínimos cuadrados ponderados robustos y finalmente el método de rotación Varimax en contraposición con el Promin.

6. Resultados

6.1 Análisis factorial exploratorio AWA

En esta primera sección el análisis se enfoca en las diferencias y similitudes encontradas en el AFE del instrumento AWA, se presenta los resultados del software estadístico SPSS usando de forma clásica en psicología y que permite el análisis utilizando para ello la matriz de correlación de Pearson, en contraste con lo recomendado por la estadística que corresponde al uso de software Rstudio y Factor que posibilitan los análisis con matrices policóricas y tetracóricas, para que el lector pueda observar lo inapropiado del primer enfoque a la hora de analizar instrumentos de medida en psicología.

Los resultados presentados corresponden a las respuestas dadas por una muestra final de 357 participantes a las 25 preguntas del cuestionario AWA, dentro del análisis previo de los datos se resalta que la base final, no contenía valores perdidos, puesto que los cuestionarios de aplicación en línea contaban con la opción de obligatoriedad de selección en las preguntas que conformaban el test AWA, en caso contrario el formulario no fue tenido en cuenta para conformar la base de datos.

Los datos, determinante de la matriz, prueba de Bartlett y estadístico KMO se estimaron para las matrices resultantes de los programas estadísticos SPSS y Rstudio.

Tabla 2.

Determinante, Bartlett y KMO

Test	Programa	Estadístico	gl	p valor	Programa	Estadístico	gl	p valor
Determinante	SPSS	0,014			Rstudio	0,0001		
Bartlett	SPSS	1514,040	300	0,000	Rstudio	3082,022	300	0,000
KMO	SPSS	0,785			Rstudio	0,870		

Nota: KMO=Prueba Kaiser Meyer Olkin

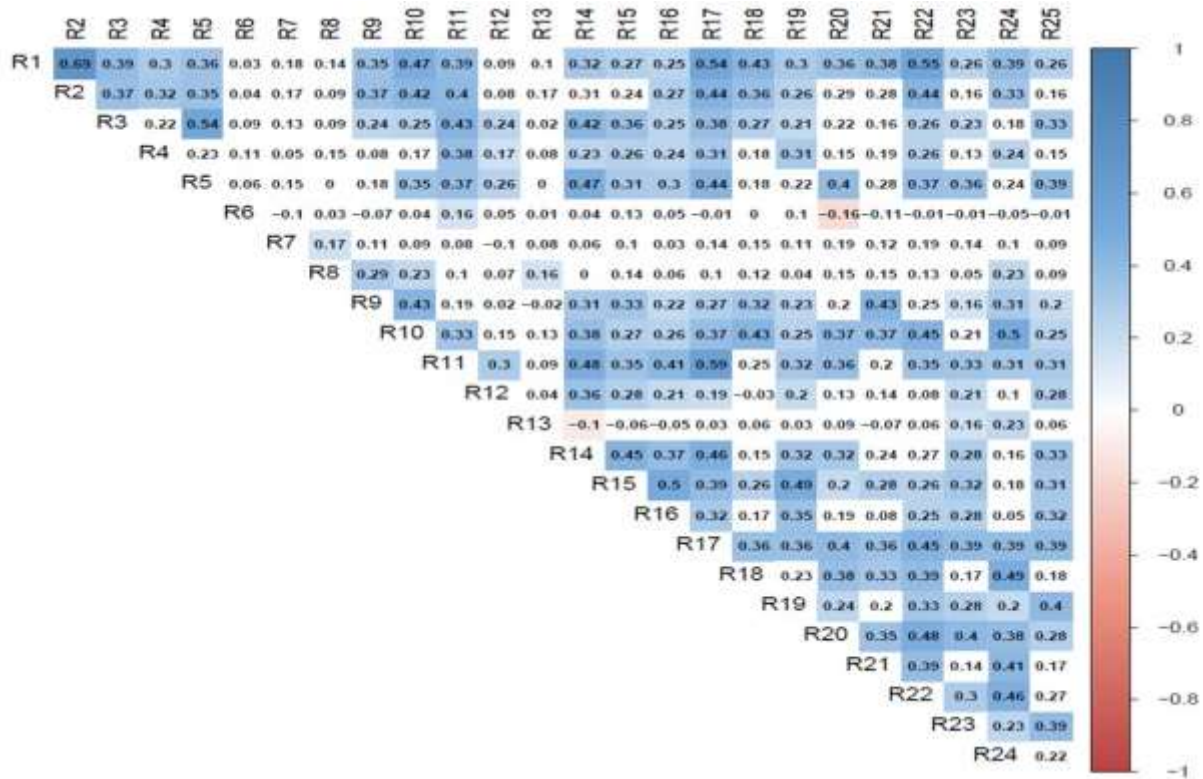
El determinante de la matriz para cada caso es cercano a cero, lo indica alta multicolinealidad entre las variables, el valor para la matriz policórica se ajusta más a la norma que establece que los valores más adecuados son aquellos más cercanos a 0. En este caso del test de Bartlett, tanto la matriz policórica como la de Pearson muestran presencia de multicolinealidad en el conjunto de variables ($p < 0,05$), finalmente el KMO, coeficiente sirve para comparar las magnitudes de los coeficientes de correlación general o simple con respecto a las magnitudes de los coeficientes de correlación parcial, nos indica un aceptable nivel de multicolinealidad entre las

variables para la matriz de Pearson estimada por el programa SPSS y un buen nivel de multicolinealidad entre las variables para la matriz estimada por el programa Rstudio.

La matriz de correlación de Pearson estimada por medio del programa estadístico SPSS muestra coeficientes que oscilan entre -0,16 y 0,47, un total de 20 ítems presentan alguna correlación negativa con otro, 10 de los ítems presentan correlaciones de cero, la totalidad de los ítems presentan correlaciones con otros ítems menores a 0,3, tan solo 19 presentan alguna correlación mayor que 0,3 y ninguno de los ítems presenta correlaciones por encima de 0,5. El análisis de la matriz de Pearson indicaría que debería analizarse la opción de eliminar de la estructura los ítems 6,7,8,12,13 y 23 que presentan correlaciones que no superan el 0,3.

La matriz de correlación policórica estimada por medio del programa Rstudio, presenta coeficientes de correlación que oscilan entre -0,16 y 0,69. Un total de 16 ítems presentan correlaciones negativas con otros ítems de la prueba, 6 ítems presentan correlaciones de cero, la totalidad de los 25 ítems presentan correlaciones con otros ítems menores a 0,3, un total de 21 ítems presentan correlaciones mayores que 0,3 y finalmente, 11 ítems presenta correlaciones con otros ítems por encima de 0,5. El análisis indicaría que debería analizarse la opción de eliminar de la estructura los ítems 6, 7, 8 y 13 que presentan correlaciones que no superan el 0,3.

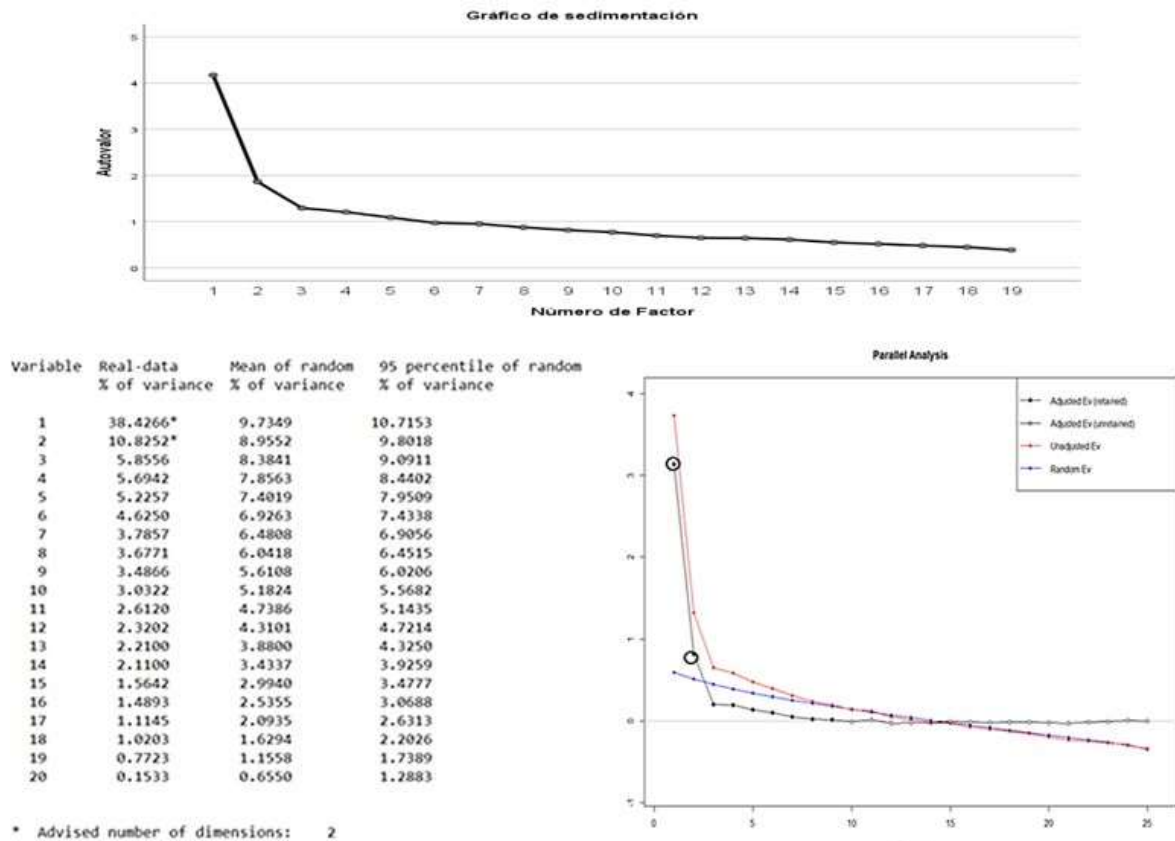
Figura 6. Matriz de correlación Policórica prueba AWA



Se eliminó en SPSS los ítems 6,7,8,12,13 y 23, en Rstudio y en Factor los ítems 6, 7, 8 y 13 a razón del análisis de la matriz de correlación, adicional, se eliminaron en Rstudio los ítems 4,9 y 12 por presentar cargas $< 0,3$.

En los softwares se configuro el análisis así: se usó la matriz de correlación Pearson para SPSS y policórica para los otros, el método seleccionado para la extracción de factores fue componentes principales CP, residuales mínimos ULS para Rstudio y mínimos cuadrados no ponderados robustos RULS para factor, el método de rotación Varimax para SPSS y Oblimin para Rstudio y Promin para factor, solicitando una solución de 4 factores ajustándonos al referente teórico del AWA, no obstante, se analiza la recomendación del análisis paralelo para determinar el número de factores a retener.

Figura 7. Gráfico de sedimentación SPSS (superior) Análisis paralelo software Factor (inferior izquierda) y software Rstudio (inferior derecha).



SPSS difiere en sus resultados, recomienda extraer 5 factores, gráfico de sedimentación criterio Kaiser, el quinto factor estaba compuesto por solo dos ítems, entonces se retienen 4 factores, Rstudio y Factor producen la misma recomendación extraer 2 factores, el gráfico de autovalor para el criterio de contraste de caída y la recomendación textual.

En el programa SPSS, tras la eliminación de los ítems antes relacionados, la matriz de correlación Pearson presentó valores que variaron en el rango entre -0,11 y 0,47, no todos los ítems presentaron correlaciones significativas o superiores a 0,3. La matriz, mostró un determinante de

la matriz (0.028), la prueba de Bartlett ($p < 0,001$) y el estadístico KMO (0,79) y la solución final rotada con un total de 4 factores que explicaron el 45% de la varianza.

Tabla 3.

Resumen del análisis factorial exploratorio del test AWA SPSS

Ítem	F1	F2	F3	F4
3 Tengo derecho a pegarle a un animal si me molesta	0,634			
5* Le pegaría a mi mascota si me enfadara	0,785			
14* Si me canso de un animal, lo dejo en el campo	0,516		0,404	
17* Tengo como afición matar gorriones, palomas... pequeños animales en general	0,477			0,377
20 Me encantaría colaborar con un refugio de animales abandonados	0,428			
25* Los animales agresivos deben ser sacrificados de inmediato pues no pueden ser curados	0,427		0,417	
9 Nunca educaría a mi mascota a golpes		0,643	0,325	
10 Me preocupa que los toros sufran en el ruedo, aunque sean pocos minutos		0,601		
18 El abandono de animales me parece una práctica muy cobarde e irresponsable por parte de quien la realiza		0,549		
21 Yo no abandonaría a mi mascota		0,587		
24 Me da mucha pena ver al toro sufriendo en el ruedo, y que la gente se divierta		0,561		0,348
15* Las condiciones de vida de los animales de granja no les afectan porque son seres inferiores			0,710	
16* El abandono le produce al animal mucha sensación de libertad			0,630	
19* Los animales de granja, ni sufren ni padecen			0,652	

Tabla 3. (Continuación)

Ítem	F1	F2	F3	F4
1 Los animales sufren; si les pegas les duele		0,321		0,553
2 Todo animal doméstico debería estar bien cuidado				0,643
4* Me gusta que el coleo sea una seña de identidad colombiana			0,304	0,601
11* Me encantaría ir a cazar	0,387		0,311	0,526
22 Es necesario proteger a los animales con leyes		0,318		0,490

Nota: Método de extracción: análisis de componentes principales. Método de rotación: Varimax con normalización Kaiser.

a. La rotación ha convergido en 10 iteraciones.

En el programa Rstudio, tras la eliminación de los ítems antes relacionados, la matriz de correlación policórica presentó valores que variaron en el rango entre 0,08 y 0,69, todos los ítems presentaron correlaciones significativas y superiores a 0,3. Así mismo la matriz, mostró un determinante de la matriz (0.0007), la prueba de Bartlett ($p < 0,0001$) y el estadístico KMO (0,89) y la solución final rotada con un total de 2 factores que explicaron el 40% de la varianza.

Tabla 4.

Resumen del análisis factorial exploratorio del test AWA Rstudio

Ítem	Descripción	F1	F2
1	Los animales sufren; si les pegas les duele.		0,65
2	Todo animal doméstico debería estar bien cuidado.		0,54
10	Me preocupa que los toros sufran en el ruedo, aunque sean pocos minutos.		0,61
18	El abandono de animales me parece una práctica muy cobarde e irresponsable por parte de quien la realiza.		0,65

Tabla 4. (Continuación)

Ítem	Descripción	F1	F2
20	Me encantaría colaborar con un refugio de animales abandonados.		0,47
21	Yo no abandonaría a mi mascota.		0,56
22	Es necesario proteger a los animales con leyes.		0,64
24	Me da mucha pena ver al toro sufriendo en el ruedo, y que la gente se divierta.		0,77
11*	Me encantaría ir a cazar.	0,56	
14*	Si me canso de un animal, lo dejo en el campo.	0,68	
15*	Las condiciones de vida de los animales de granja no les afectan porque son seres inferiores.	0,67	
16*	El abandono le produce al animal mucha sensación de libertad.	0,68	
17*	Tengo como afición matar gorriones, palomas... pequeños animales en general.	0,47	
19*	Los animales de granja, ni sufren ni padecen.	0,50	
23*	Yo siempre compro mascotas en las tiendas de animales, porque en los refugios son feos y viejos.	0,47	
25*	Los animales agresivos deben ser sacrificados de inmediato pues no pueden ser curados.	0,56	
3*	Tengo derecho a pegarle a un animal si me molesta.	0,52	
5*	Le pegaría a mi mascota si me enfadara.	0,53	

Nota: Extracción ULS; rotación Oblimin; cargas inferiores a 0,3 suprimidas

En el programa FACTOR, tras la eliminación de los ítems antes relacionados, la matriz de correlación policórica presento, valores que variaron en el rango entre 0,06 y 0,71, todos los ítems presentaron correlaciones significativas y superiores a 0,3. Así mismo la matriz, mostró un determinante de la matriz (0.0002), la prueba de Bartlett ($p < 0,0001$) y el estadístico KMO (0,87) y la solución final rotada con un total de 2 factores que explicaron el 43% de la varianza.

Tabla 5.

Resumen del análisis factorial exploratorio del test AWA Factor.

Ítem	Descripción	F1	F2
1	Los animales sufren; si les pegas les duele.		0,73
2	Todo animal doméstico debería estar bien cuidado.		0,61
9	Nunca educaría a mi mascota a golpes.		0,49
10	Me preocupa que los toros sufran en el ruedo, aunque sean pocos minutos.		0,65
18	El abandono de animales me parece una práctica muy cobarde e irresponsable por parte de quien la realiza.		0,73
20	Me encantaría colaborar con un refugio de animales abandonados.		0,44
21	Yo no abandonaré a mi mascota.		0,60
22	Es necesario proteger a los animales con leyes.		0,67
24	Me da mucha pena ver al toro sufriendo en el ruedo, y que la gente se divierta.		0,79
11*	Me encantaría ir a cazar.	0,63	
12*	Yo tendría a los animales de granja encerrados, para organizarlos mejor.	0,61	
14*	Si me canso de un animal, lo dejo en el campo.	0,75	
15*	Las condiciones de vida de los animales de granja no les afectan porque son seres inferiores.	0,65	
16*	El abandono le produce al animal mucha sensación de libertad.	0,65	
17*	Tengo como afición matar gorriones, palomas... pequeños animales en general.	0,48	
19*	Los animales de granja, ni sufren ni padecen.	0,48	
23*	Yo siempre compro mascotas en las tiendas de animales, porque en los refugios son feos y viejos.	0,48	
25*	Los animales agresivos deben ser sacrificados de inmediato pues no pueden ser curados.	0,58	
3*	Tengo derecho a pegarle a un animal si me molesta.	0,56	
4*	Me gusta que el coleo sea una seña de identidad colombiana.	0,30	
5*	Le pegaría a mi mascota si me enfadara.	0,57	

Nota: Extracción RULS; rotación Promin; cargas inferiores a 0,3 suprimidas.

La parsimonia de la solución lograda por la matriz policórica, métodos robustos, rotaciones oblicuas es clara, puesto que se obtuvieron coeficientes más altos, mayor porcentaje de varianza explicada y teóricamente la acomodación de los ítems en cada factor, tiene un sustento teórico más apropiado desde la óptica de la psicología, en contraposición con la selección del criterio de uso común “Little Jiffy”.

6.3 Análisis factorial confirmatorio IECA

Esta segunda y tercera sección se enfoca en las diferencias y similitudes encontradas en el AFC y la TRI del instrumento IECA, se compara los resultados Del paquete AMOS para SPSS que es el paquete comúnmente utilizado por los psicólogos y Rstudio que es el paquete recomendado por la estadística para el análisis instrumento con variables categóricas.

Los resultados presentados corresponden a las respuestas a los 22 ítems de 310 participantes del cuestionario IECA, la base final no contenía valores perdidos.

En el software se ajustaron dos modelos, el modelo 1 en adelante F2 compuesto por dos factores que agrupa 12 de los 22 ítems y el modelo 2 en adelante F3 compuesto por tres factores, que agrupo el total de los 22 ítems.

Se usó la matriz de correlación Pearson para AMOS de SPSS y tetracórica para Rstudio, el método de estimación máxima verosimilitud y el de mínimos cuadrados Diagonalizados para Rstudio, lo anterior se realiza con el objeto de ilustrar lo que comúnmente realizan los psicólogos, inapropiado en la mayoría de los casos, en contraste con las recomendaciones derivadas de la estadística.

Los resultados obtenidos por AMOS, en relación con las correlaciones entre los factores y los ítems fluctuaron entre 0,05 y 0,73 para el modelo de dos factores y entre -0,03 y 0,86 para el modelo de tres factores.

Los resultados obtenidos por Rstudio, en relación con las correlaciones entre los factores y los ítems fluctuaron entre 0,18 y 0,98 para el modelo de dos factores y entre -0,30 y 0,99 para el modelo de tres factores.

Figura 8. Diagrama de flujo AFC paquete AMOS modelo dos factores (superior) modelo tres factores (inferior)

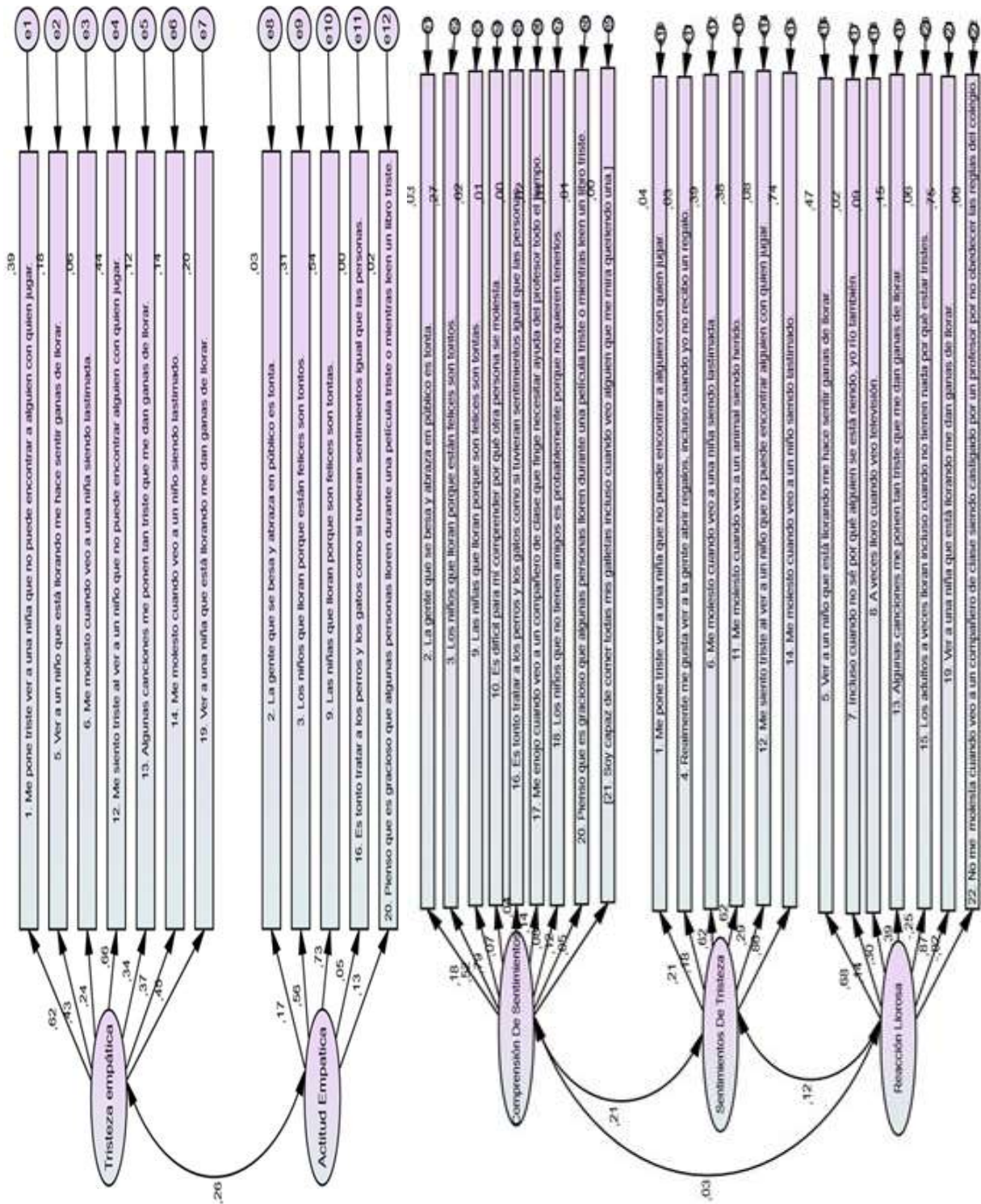


Figura 9. Diagrama de flujo AFC paquete Rstudio modelo dos factores (superior) modelo tres factores (inferior)

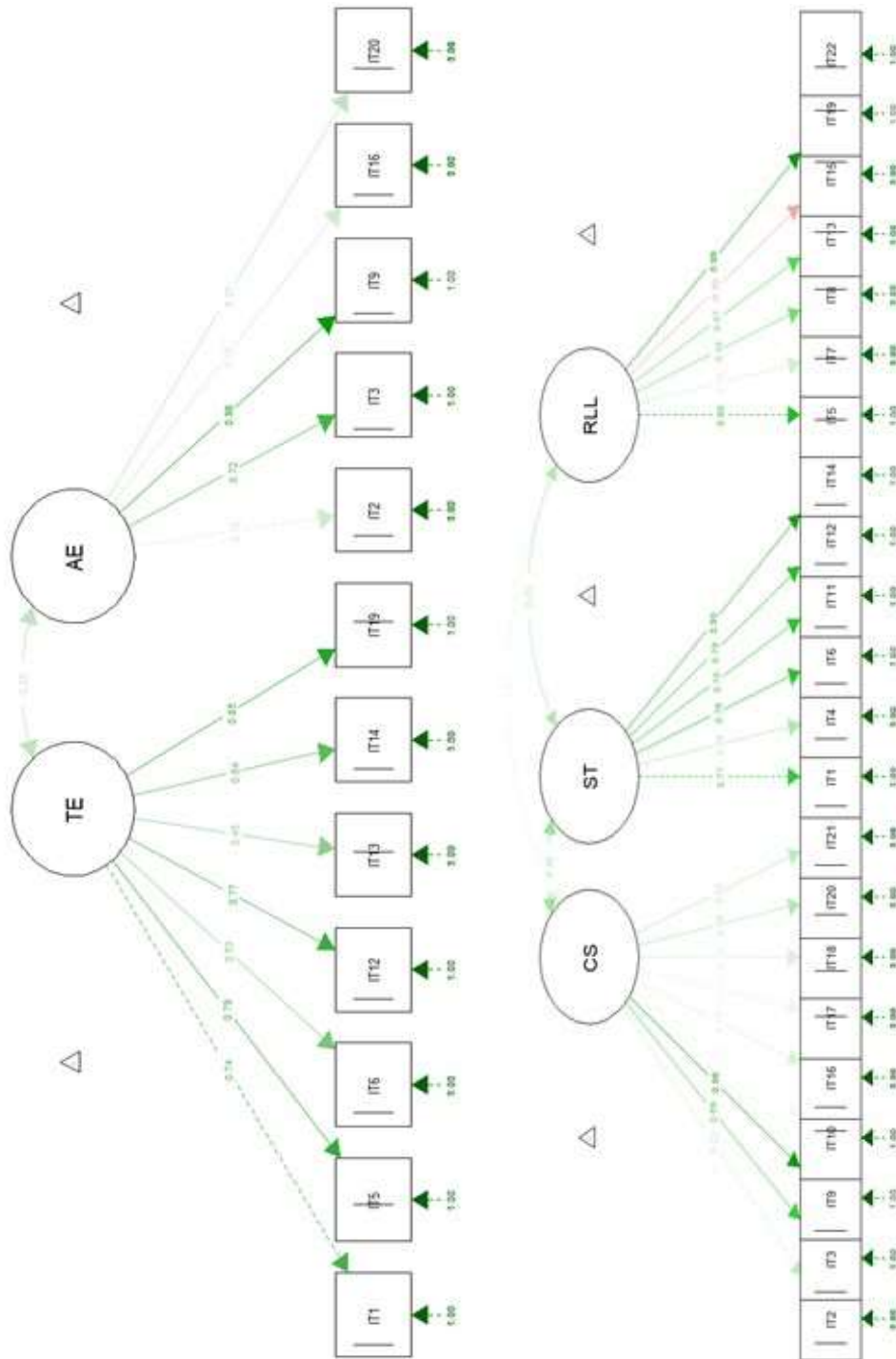


Tabla 6.

Estadísticas de bondad de ajuste AMOS

Modelo	Ajuste Absoluto		Ajuste Incremental		
	Chi cuadrado	RMSEA	CFI	TLI	NFI
Dos Factores	p<0,001	0,116	0,547	0,436	0,505
Tres Factores	p<0,001	0,061	0,689	0,652	0,555

Tabla 7.

Estadísticas de bondad de ajuste Rstudio

Modelo	Ajuste Absoluto		Ajuste Incremental		
	Chi cuadrado	RMSEA	CFI	TLI	NFI
<u>Dos Factores</u>	<u>p<0,001</u>	<u>0,099</u>	<u>0,843</u>	<u>0,805</u>	<u>0,756</u>
<u>Tres Factores</u>	<u>p<0,001</u>	<u>0,057</u>	<u>0,866</u>	<u>0,850</u>	<u>0,768</u>

Con relación a las medidas de bondad de ajuste, la ratio de verosimilitud de X2 muestra que existen diferencias significativas entre la matriz de datos observada y la estimada por el modelo ($p<0,05$) tanto para el modelo F2 como el F3, el error de aproximación cuadrático medio, expresa que el modelo F2 ($RMSEA> 0,05$) es inapropiado, caso contrario el modelo F3 ($RMSEA<0,05$) resulto apropiado.

Al contrastar las medidas de ajuste incremental obtenidas por AMOS, el índice de ajuste comparativo para los modelos F2 y F3 se encuentra por debajo del parámetro mínimo ($CFI<0,9$), el índice TOKER-LEWIS y la medida de ajuste normativo arrojaron valores menores del mínimo

esperado ($TLI < 0,9$, $NFI < 0,9$), por tal motivo existe evidencia para concluir que el modelo no se ajusta satisfactoriamente a los datos.

Al contrastar las medidas de ajuste incremental obtenidas por Rstudio, el índice de ajuste comparativo para los modelos F2 y F3 se encuentra por debajo, mas no muy alejado del parámetro mínimo ($CFI < 0,9$), el índice TOKER-LEWIS y la medida de ajuste normativo arrojaron valores menores, mas no alejados del mínimo esperado ($TLI < 0,9$, $NFI < 0,9$), por tal motivo existe evidencia para concluir que el modelo no se ajusta satisfactoriamente a los datos.

Finalmente, las medidas de parsimonia de AMOS muestran que, aunque los modelos no se ajustan, el F3, muestra un PRATIO, PCNI y PCFI mayor que el modelo F2, indicando un ajuste mejor. No obstante, el AIC es menor en el modelo F2. En conclusión, hay evidencia suficiente para rechazar los dos modelos.

La re-especificación que propone AMOS del modelo factorial propuesto llevaría a suprimir en el modelo F3 los ítems 10,16,18,21,15 y 22. Por otro lado, en el modelo F2 se aconseja suprimir los ítems 2,6,16,20 para observar los cambios que esto tendría en los estadísticos de bondad de ajuste.

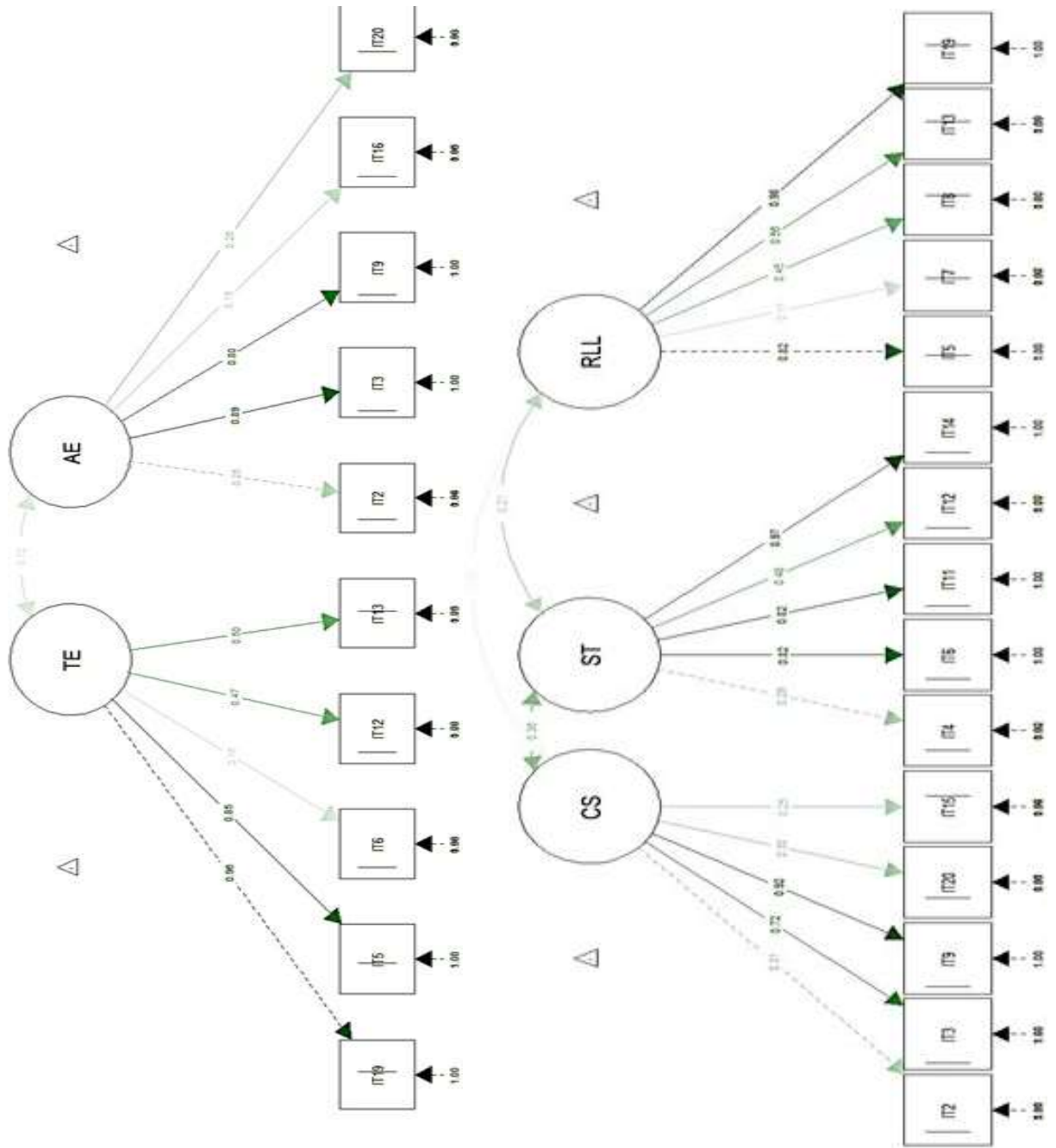
La re-especificación que propone Rstudio del modelo factorial propuesto llevaría a suprimir en el modelo F3 los ítems 1, 10, 16,17,18, 21, 22 y mover el ítem 15 al primer factor . Por otro lado, en el modelo F2 se aconseja suprimir los ítems 14 y 1 para observar los cambios que esto tendría en los estadísticos de bondad de ajuste.

Una vez efectuados los cambios en AMOS, el modelo de dos factores aun no cumple los criterios mínimos el ajuste absoluto, el mismo sigue mostrando un valor inapropiado ($RSMEA = 0,121$) por su parte, los criterios de ajuste incremental, aunque aumentaron ($CFI = 0,696$, $TLI = 0,579$ y $NFI = 0,660$) no alcanzan el mínimo valor aceptado 0,9. El modelo de tres factores

aun no cumple los criterios mínimos el ajuste absoluto, el mismo sigue mostrando un valor inapropiado ($RSMEA=0,072$) por su parte, los criterios de ajuste incremental, aunque aumentaron ($CFI= 0,761$, $TLI= 0,716$ y $NFI = 0,670$) no alcanzan el mínimo valor aceptado 0,9, A fin de ajustar el modelo la opción para ambos casos sería la eliminación de ítems adicionales, lo que tendría como consecuencia un detrimento en la interpretación teórica del constructo medido.

Una vez efectuados los cambios en Rstudio, el modelo de dos factores cumplen los criterios mínimos el ajuste absoluto ($RSMEA=0,027$) por su parte, los criterios de ajuste incremental, aunque aumentaron ($CFI= 0,987$, $TLI= 0,983$ y $NFI = 0,915$) alcanzan el mínimo valor aceptado 0,9. El modelo de tres factores cumple los criterios mínimos el ajuste absoluto ($RSMEA=0,047$) por su parte, los criterios de ajuste incremental aumentaron ($CFI= 0,9521$, $TLI= 0,942$ y $NFI = 0,90$) alcanzando el mínimo valor aceptado 0,9. Los cambios elaborados en la versión de tres factores, tiene consecuencias positivas en la acomodación del ítem 15 en el primer factor, pues el ítem se ubica teóricamente de forma más apropiada, Es importante aclarar que hay indicadores como el AIC que no se puede calcular en mínimos cuadrados Diagonalizados, razón por la cual no se menciona dicho criterio.

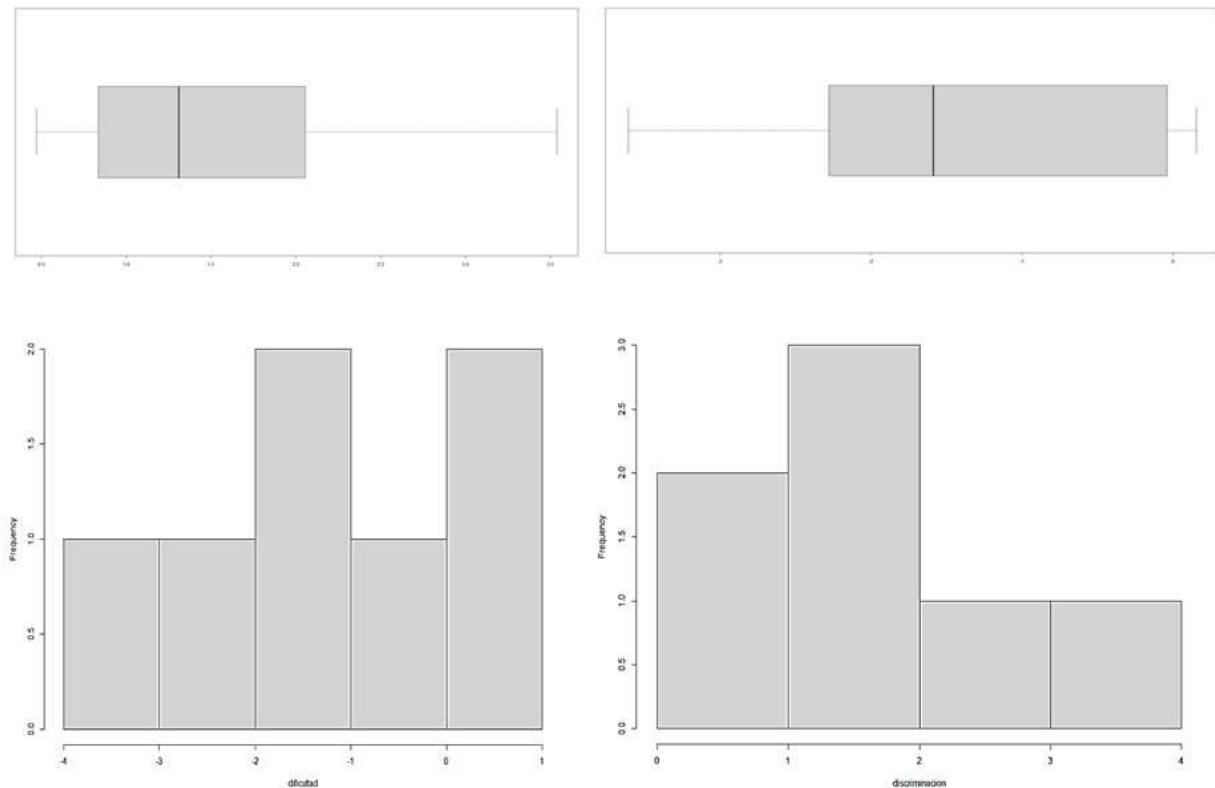
Figura 10. Diagrama AFC paquete Rstudio modelo dos factores (superior) modelo tres factores (inferior)



6.4 Análisis teoría de respuesta al ítem IECA

Se estimó los parámetros de cada curva característica con el modelo de dos parámetros, la decisión de trabajar con el modelo 2PL se basa en que el mismo presenta un mejor ajuste a los datos de la muestra, pues el modelo 3PL presentó problemas de convergencia en el paquete *f* de R studio tras 500 interacciones. No se evidenciaron datos atípicos, los valores de discriminación no evidencian la presencia de valores negativos o cero. Se resalta que se probará la estructura de dos factores para el IECA conformado por 12 ítems, por ser aquella que presentó los valores más altos de ajuste el AFC.

Figura 11. Boxplot y grafico de barras dificultad y discriminación modelo 2PL



Las curvas características y los estimadores de los parámetros de los ítems para el modelo 2PL y bajo la estructura IECA dos factores, se observa en gris los ítems 6 (factor 1) y 16 (factor2) que se alejan de la expectativa y debe considerarse la posibilidad de eliminarlos.

Tabla 8.

Estimadores de los parámetros F1

Ítem	D_i	α_i
1	-1,74	1,31
5	-0,18	2,70
6	-3,61	0,47
12	-1,59	1,43
13	0,16	1,02
14	-2,82	0,66
19	0,10	3,54

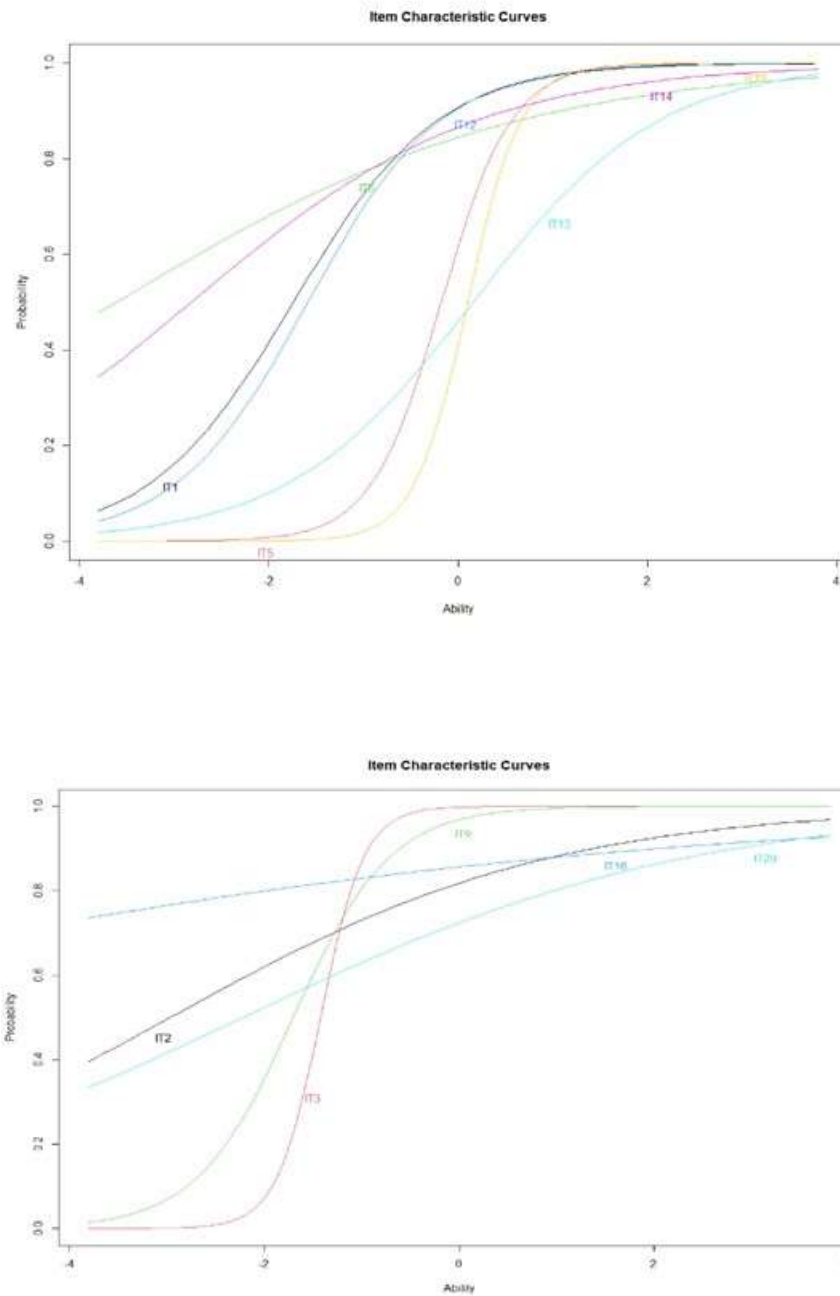
Nota: D_i =Dificultad α_i =Discriminación

Tabla 9.

Estimadores de los parámetros F2

Ítem	D_i	α_i
2	-2,96	0,51
3	-1,42	4,45
9	-1,70	2,02
16	-8,86	0,20
20	-2,21	0,44

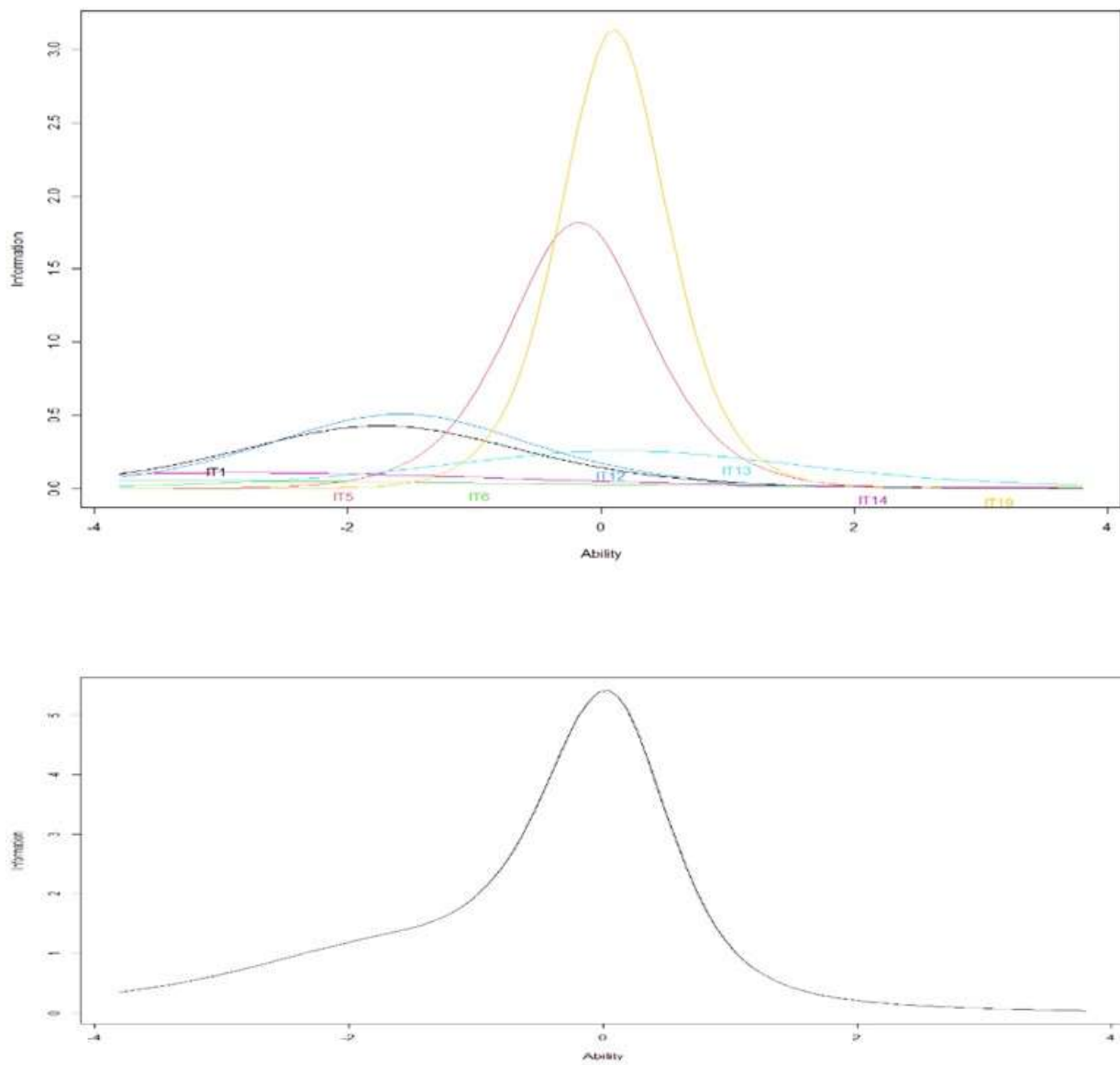
Nota: D_i =Dificultad α_i =Discriminación

Figura 12. ICC modelo 2PL F1 (superior) F2 (inferior)

La dificultad de los ítems que se ajustan al criterio fluctuó entre -2.96 y 0,16 lo que sugiere que el test mide el rasgo latente colocando el énfasis en los sujetos cuya empatía presenta valores en el promedio y por debajo del mismo.

En relación con la función de información de los ítems y del test, se puede apreciar que el rango de habilidad medido por el test IECA corresponde a la habilidad promedio y por debajo, se observa que los ítems 15 y 19 miden a los participantes con empatía promedio, los demás ítems aportan la mayor cantidad de información sobre participantes con valores por debajo del promedio.

Figura 13. Función de información de los ítems (superior) y test (inferior) modelo 2PL



7. Conclusiones

Este trabajo propuso comparar al interior de las Técnicas Estadísticas AFE, AFC y TRI los resultados obtenidos tras seleccionar determinados criterios de análisis a fin de establecer pautas que faciliten el manejo adecuado dentro de los procesos de adaptación o validación de pruebas en psicología, recurriendo a una muestra de 310 participantes para el IECA y 357 para el AWA.

En el marco del AFE, se estudió la estructura factorial del AWA, se observaron diferencias importantes para discutir. Al estudiar la estructura factorial del AWA se observó que la correlación de Pearson infraestimó los coeficientes en comparación con la correlación tetracórica, en relación con la composición y organización de los factores la extracción por componentes principales y Mínimos Cuadrados no Ponderados Robustos, finalizó en estructuras diferentes, la segunda estructura, estimada por Mínimos Cuadrados no Ponderados Robustos tuvo cargas factoriales más altas y ajustada mejor a la teoría que pretende medir el instrumento.

En el marco del AFC se estudió la estructura de dos y tres factores del IECA, se observaron diferencias importantes, la matriz de correlación de Pearson estimada por el programa AMOS para SPSS incluso con la eliminación de los ítems 10,16,18,21,15 y 22 de la estructura de 3 factores y los ítems 2,6,16 Y 20 en el modelo de dos factores, los datos no ajustaron al modelo ni lograron los valores mínimos esperados. Por otro lado, cuando se utilizó el programa Rstudio utilizando la matriz de correlación tetracórica y luego de la eliminación de los ítems 1, 10, 16,17,18, 21, 22 y el cambio el ítem 15 al primer factor en la estructura de tres factores y la eliminación de los ítems 14 y 1 en la estructura de dos factores, mostró que dichos cambios tuvieron buenos efectos en los

estadísticos de bondad de ajuste, se resalta que mover el ítem 15 mejoró sustancialmente la interpretación de la estructura del instrumento IECA.

En relación con la TRI y el AFC se analizó la estructura de dos factores del instrumento IECA, se obtuvo un ajuste global del modelo para los ítems, la decisión de llevar a cabo el análisis de la estructura de dos parámetros radica en que la misma fue la que presentó óptimos valores de bondad de ajuste, el análisis de los criterios aconseja eliminar de la estructura el ítem 6 y 16, lo que tiene por consecuencia una pérdida menor de información, una recomendación que se extrae de los resultados serían incluir preguntas que discriminen mejor a las personas con mayor trazo latente, aunque los resultados están en consonancia con lo expuesto la teoría de las pruebas de empatía, cuyo propósito principal es medir aquellos individuos que presentan bajos niveles de empatía es aconsejable que un instrumento logre identificar los valores a lo largo de todo el constructo latente.

A manera de síntesis los resultados y conclusiones indicarían la necesidad de seguir las recomendaciones en relación con el análisis de datos categóricos en torno a la validación de test psicológicos, las diferencias en los resultados producto de seleccionar entre uno u otro criterio, lleva a concluir que seguir las recomendaciones brindadas por la estadística no solo son una obligación metodológica, sino que hacerlo permite descubrir con mayor precisión la estructura teórica subyacente de los test psicológicos.

Se observa, por ejemplo, que la matriz factorial resultante del AFE muestra mayor relevancia teórica del atributo medido, por otro lado, la reacomodación del ítem 15 del AFC tiene lógica desde el punto de vista teórico y responde a una de las principales críticas a la estructura analizada en el IECA, finalmente la óptica de análisis y los resultados obtenidos por la TRI señala limitaciones del AFC, si bien se logró responder a los objetivos propuestos en el presente trabajo,

vale destacar que el presente estudio necesita de socialización en la comunidad de psicólogos que analizan datos categóricos para asegurar que el uso de las técnicas sea el adecuado y pueda ser empleada como aconseja la estadística.

Referencias

- Aldás, J., & Uriel, E. (2005). *Análisis multivariante aplicado: aplicaciones al marketing, investigación de mercados, economía, dirección de empresas y turismo*. España: Thomson.
- Barajas Marcos, I, Prieto Alaiz, M. (2015). *Análisis factorial con variables categóricas*.
- Bryant, Brenda. (1982). *An index of empathy for children and adolescents*. *Child Development*, 53, 413-425. . Disponible en: <http://www.jstor.org/stable/1128984>
- Calderón, K. A. (2011). *Adaptación de la escala de empatía de Bryant, B. (1982) Para niños, niñas y adolescentes en Costa Rica/Adaptation of the Bryant, B. index of empathy for children and adolescents (1982) in Costa Rica*. *Actualidades Investigativas en Educación*, 11(2). Disponible en: <https://revistas.ucr.ac.cr/index.php/aie/article/view/10193/18051>
- Carretero-Dios, H., & Pérez, C. (2005). *Normas para el desarrollo y revisión de estudios instrumentales*. *International Journal of clinical and health psychology*, 5(3), 521-551. Disponible en: <https://www.redalyc.org/pdf/337/33705307.pdf>
- Castrillón Moreno, D. A., Luna Montaña, I., Avendaño Prieto, G., & Pérez-Acosta, A. M. (2007). *Validación del Body Shape Questionnaire (Cuestionario de la Figura Corporal) BSQ para la población colombiana*. *Acta Colombiana de Psicología*, Vol. 10, no. 1 (ene.-jun. 2007); p. 15-23. Disponible en: <https://repository.ucatolica.edu.co/bitstream/10983/543/1/v10n1a03.pdf>
- Chahín-Pinzón, N., & Briñez, B. L. (2015). *Propiedades psicométricas de la Escala de Ideología de Género en adolescentes colombianos*. *Universitas Psychologica*, 14(1), 15-23. Disponible en: <https://www.redalyc.org/pdf/647/64739086019.pdf>

- Del Barrio, Victoria, Aluja, Anton y García, Luis. (2004). *Bryant's empathy index for children and adolescents: psychometric properties in the spanish language*. *Psychological Reports*, (95), 257-262. <https://doi.org/10.2466/pr0.95.1.257-262>
- De Wied, M., Maas, C., Van Goozen, S., Vermande, M., Engels, R., Meeus, W., & Goudena, P. (2007). *Bryant's empathy index*. *European Journal of Psychological Assessment*, 23(2), 99-104. Disponible en: DOI 10.1027/1015-5759.23.2.99
- Dominguez-Lara, S. A., Fernández-Arata, M., Merino-Soto, C., Navarro-Loli, J. S., & Calderón De la Cruz, G. (2018). Inventario de Violencia y Acoso Psicológico en el Trabajo (IVAPT) in Colombia: the dangerous Little Jiffy. *Revista Salud Uninorte*, 34(2), 536-537.
- Ferrando, P. J., & Lorenzo-Seva, U. (2014). *Exploratory Item Factor Analysis: Some additional considerations*. *Anales De Psicología / Annals of Psychology*, 30(3), 1170-1175. Disponible en: <https://doi.org/10.6018/analesps.30.3.199991>
<https://revistas.um.es/analesps/article/view/analesps.30.3.199991/165451>
- Fernández-Pinto, I., López-Pérez, B., & Márquez, M. (2008). *Empatía: Medidas, teorías y aplicaciones en revisión*. *Anales de Psicología/Annals of Psychology*, 24(2), 284-298. <https://revistas.um.es/analesps/article/view/42831>
- Gil, B. M. (2015). *La actitud hacia el bienestar animal en el ámbito educativo*. *Enseñanza de las ciencias: revista de investigación y experiencias didácticas*, 33(1), 271-272. Disponible en: <http://invenio2.unizar.es/record/15510/files/TESIS-2014-056.pdf>
- González, M. J. (2008). *El análisis de reactivos con el modelo Rasch*. *Manual técnico A*. Serie: medición y metodología. Instituto Nacional para la evaluación en educación.

- Herzog Jr, H. A., Betchart, N. S., & Pittman, R. B. (1991). *Gender, sex role orientation, and attitudes toward animals*. *Anthrozoös*, 4(3), 184-191. Disponible en: <https://doi.org/10.2752/089279391787057170>
- Herzog, H. A., & Mathews, S. (1997). *Personality and attitudes toward the treatment of animals*. *Society & animals*, 5(2), 169-175. Disponible en: <https://doi.org/10.1163/156853097X00060>
- Herzog, H., Grayson, S., & McCord, D. (2015). *Brief measures of the animal attitude scale*. *Anthrozoös*, 28(1), 145-152. Disponible en: <https://doi.org/10.2752/089279315X14129350721894>
- Holgado-Tello, F. P., Morata-Ramírez, M. Á., & Barbero-García, M. I. (2018). *Confirmatory Factor Analysis of Ordinal Variables: H Simulation Study Comparing the Main Estimation Methods*. *Avances en Psicología Latinoamericana*, 36(3), 601-617.
- Jiménez Ardila, L. O., Avendaño-Prieto, B. L., Giraldo, A., & Montañez, G. (2018). *Análisis psicométrico de la prueba EDEMO-32 con una muestra de menores no infractores*. *Investigación en psicología: aplicaciones e intervenciones*. Bogotá: Editorial Universidad Católica de Colombia. Disponible en: http://repository.ucatolica.edu.co:8080/bitstream/10983/23563/1/investigacion-en-psicologia_Cap07.pdf
- Kline, P. (1997). *An easy guide to factor analysis*. Routledge.
- Korkmaz, S., Goksuluk, D., & Zararsiz, G. (2014). *MVN: An R package for assessing multivariate normality*. *The R Journal*, 6(2), 151-162. Disponible en: <https://www.mclibre.org/descargar/docs/revistas/the-r-journal/the-r-journal-12-en-201412.pdf#page=151>

- Li, C. H. (2016). *Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares*. *Behavior research methods*, 48(3), 936-949.
- Linacre, L. M. (2012). *Winsteps Tutorial 1. June*. Disponible en: <http://www.winsteps.com/a/winsteps-tutorial-4.pdf>
<http://dx.doi.org/10.12804/revistas.urosario.edu.co/apl/a.4932>
- Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., & Tomás-Marco, I. (2014). *El análisis factorial exploratorio de los ítems: una guía práctica, revisada y actualizada*. *Anales de Psicología/Annals of Psychology*, 30(3), 1151-1169. Disponible en: <https://www.redalyc.org/pdf/167/16731690031.pdf>
http://e-spacio.uned.es/fez/eserv/tesisuned:ED-Pg-MCyS-Mamorata/MORATA_RAMIREZ_MAngeles_Tesis.pdf
- McDonald, S. E., Vidacovich, C., Ascione, F. R., Williams, J. H., & Green, K. E. (2015). *The Children's Treatment of Animals Questionnaire: A Rasch Analysis*. *Anthrozoös*, 28(1), 131-144. Disponible en: <https://doi.org/10.2752/089279315X14129350722172>
- Martínez, C. M., & Rondon, M. A. R. (2012). Introducción al análisis factorial exploratorio. *Revista colombiana de psiquiatría*, 41(1), 197-207. Recuperado de <https://www.redalyc.org/pdf/806/80624093014.pdf>
- Mazas Gil, B. (2015). *La actitud hacia el bienestar animal en el ámbito educativo (No. ART-2015-90384)*. Disponible en: <https://ensciencias.uab.es/article/view/v33-n1-mazas>
- Orozco, L. C. (2010). *Medición en salud, diagnóstico, evaluación de resultados. Un manual crítico más allá de lo básico*. Bucaramanga, Colombia: Publicaciones UIS.

- Palmer, A., Jiménez, R., & Montaña, J. (2000). Tutorial sobre coeficientes de correlación con una o dos variables categóricas. *Revista Electrónica de Psicología*, 4(2). Recuperado de: https://www.researchgate.net/publication/331640990_Tutorial_sobre_coeficientes_de_correlacion_con_una_o_dos_variables_categoricas#fullTextFileContent
- Rodríguez-Jiménez, O. R., Rosero-Burbano, R. F., Sanabria, M. L. B., & Mateus, L. H. D. (2011). Producción de Conocimiento en Psicometría en Instituciones de Educación Superior de Bogotá y Chía. *Revista Colombiana de Psicología*, 20(1), 9-25. Recuperado de <http://www.scielo.org.co/pdf/rcps/v20n1/v20n1a02.pdf>
- Spearman, C. (2010). *The proof and measurement of association between two things*. *International journal of epidemiology*, 39(5), 1137-1150. Recuperado de <https://doi.org/10.1093/ije/dyq191>
- Tobar, R. A. T., & Rubio, C. V. V. (2014). *Escala de Sociotropía-Autonomía (SAS): propiedades psicométricas de la adaptación a Colombia*. *Psicogente*, 17(32). Disponible en: <http://revistas.unisimon.edu.co/index.php/psicogente/article/view/1447>
- Tobar, R. A. T., Prieto, B. L. A., López, H. A. A., & Sarmiento-López, J. C. (2017). *Propiedades psicométricas de un stroop emocional computarizado para evaluar vulnerabilidad cognitiva (Psychometric properties of a computerized emotional Stroop for assessing cognitive vulnerability)*. *CES Psicología*, 10(2), 103-115. Disponible en: <http://revistas.ces.edu.co/index.php/psicologia/article/view/3953>
- Velicer, W. F., & Jackson, D. N. (1990). *Component analysis versus common factor analysis: Some issues in selecting an appropriate procedure*. *Multivariate behavioral research*, 25(1), 1-28. Disponible en: https://www.researchgate.net/profile/Wayne_Velicer/publication/232934380_Component

[_Analysis_Versus_Common_Factor_Analysis_Some_Issues_in_Selecting_an_Appropriate_Procedure/links/09e4150b683c289198000000/Component-Analysis-Versus-Common-Factor-Analysis-Some-Issues-in-Selecting-an-Appropriate-Procedure.pdf](#)

Wright, B. D., y Mok, M. M. (2004). *An overview of the family of Rasch measurement models. Introduction to Rasch measurement*, 1-24.

Wright, B. D., y Stone, M. H. (2004). *Making measures*. Phaneron Press.