



MODELOS DE ECUACIONES ESTRUCTURALES USOS EN INVESTIGACIÓN PSICOLÓGICA Y EDUCATIVA

Daiana Yamila Rigo¹

Danilo Donolo

*Consejo Nacional de Investigaciones Científicas y técnicas
Universidad Nacional de Río Cuarto, Argentina*

RESUMEN

En los últimos años, entre la comunidad científica, se puso de moda el uso de los Modelos de Ecuaciones Estructurales, un método clasificado como multivariado. Técnica que en la actualidad ha cautivado a los investigadores de las Ciencias Sociales, al permitir captar la complejidad de los fenómenos educativos y psicológicos que se pretenden estudiar. Este artículo, muestra un ejemplo, a modo didáctico, de cómo proceder cuando se está frente al desafío de analizar y llevar a cabo un Análisis Factorial Confirmatorio, las etapas y las decisiones a considerar lo largo del proceso.

Palabras clave:

Modelo de Ecuaciones Estructurales, Análisis Factorial Confirmatorio, Psicología, Educación.

ABSTRACT

In recent years, among the scientific community, the use of the Structural Equation Models, a method classified as multivariate, became fashionable. Technique that today has captivated the researchers of the Social Sciences, allowing to capture the complexity of the educational and psychological phenomena that are intended to study. This article shows an example, in a didactic way, of how to proceed when faced with the challenge of analyzing and carrying out a Confirmatory Factor Analysis, the stages and decisions to be considered throughout the process.

Keywords:

Model of Structural Equations, Confirmatory Factor Analysis, Psychology, Education.

¹ Correspondence about this article should be address to: Daiana-Yamila Rigo, Email: daianarigo@hotmail.com.

MODELS OF STRUCTURAL EQUATIONS
USES IN PSYCHOLOGICAL AND EDUCATIONAL RESEARCH

Es común encontrar entre los instrumentos que se usan en Psicología y en Educación escalas de tipo Likert para llevar a cabo las investigaciones, la práctica de adaptar y traducir y a la vez de validar localmente esos cuestionarios ha llevado a los profesionales a avanzar en el estudio, primero del análisis factorial exploratorio, y hoy en el análisis factorial confirmatorio. Asimismo, se ha incrementado el interés en conocer las relaciones entre variables que definen a los contextos instructivos y con esto la aplicación de modelos de ecuaciones estructurales con variables latentes, entre las técnicas multivariadas principales que se utilizan en investigaciones educativas.

Como en todo, en el modelo de ecuaciones estructurales -SEM- también hay reglas del juego, que invitamos a conocer. El artículo, no tiene el objetivo de realizar una exhaustiva revisión de la bibliografía sobre la temática, más bien pretende a través de un ejemplo, previa revisión de ideas centrales a tener presentes, mostrar cómo proceder para conseguir un adecuado análisis de los datos. Mostramos de manera didáctica qué decisiones como investigador se deberían tomar frente a los datos recolectados y disponibles.

El escrito se organiza en cuatro apartados. El primero muestra de manera resumida las principales contribuciones que la literatura actualiza nos acerca para comprender los principales supuestos que subyacen al SEM. El segundo presenta un ejemplo práctico, orientado a la validación de un instrumento, mostrando las decisiones apropiadas cuando se cuenta con datos categóricos. El tercero de los apartados comenta algunas sutilezas a considerar en la lectura de datos, y muestra que las reglas también pueden ser flexibles, y eso dependen de las decisiones y riesgos que el investigador se anime a tomar. Por último, en el cuarto, se realiza un cierre considerando los principales aportes que el SEM ha provisto al campo de la investigación educativa y psicológica.

Análisis multivariados en contextos educativos. Consideraciones generales

Los SEM se han puesto de moda dentro de las Ciencias Sociales por la complejidad creciente de los fenómenos en estudio, la necesidad de mirar varias variables al mismo tiempo ha exigido el uso de nuevas técnicas y entre ellas las multivariadas. Particularmente, Orgaz Baz (2008) entiende que este tipo de modelos de análisis reflejan el objetivo de las Ciencias del Comportamiento, en tanto y en cuanto posibilitan la construcción de esquemas explicativos, teorías que facilitan la comprensión de la realidad y de los fenómenos observados.

Khine (2013) refiriendo a la investigación educativa, menciona que el uso de esta técnica ha incrementado en los últimos años en tanto su análisis ayuda a identificar los factores y las interacciones entre las características de los estudiantes, las preferencias personales, los rasgos afectivos, el nivel motivacional, las habilidades de estudio, entre otros tantos factores que definen a las prácticas educativas. Un nuevo recurso, muy fructífero, da comienzo a una nueva generación de investigaciones educativas, para captar la complejidad de los fenómenos en estudio.

El incremento, se registra en varios estudios específicos, entre ellos, el análisis de Hershberger (2003), reporta cuatro razones principales sobre el desarrollo y el crecimiento del SEM, a saber: 1. el aumento notable del número de artículos en revistas que publican investigación que usan como técnica principal los SEM; 2. la extensión del número de revistas que publican estos artículos entre los años 1994-2001; 3. la hegemonía que los SEM han adquirido en comparación con otras técnicas multivariadas; y por último, 4. la gran cantidad de publicaciones dirigidas a introducir a los investigadores en aspectos básicos de la técnica.

Específicamente, en el campo de la Psicología Educativa, su expansión se relaciona directamente con la mayor disponibilidad de datos, lo que permite a los investigadores aplicar técnicas de análisis más complejas. Si lo representamos numéricamente, en el año 2000, 12 de 134 artículos presentaban resultados mediante técnicas de modelamiento, llegando al 2010 con una proporción de 57 sobre 141. Observándose, principalmente dos propósitos centrales en el uso de los SEM. Por un lado, para desarrollar y validar instrumentos sobre constructos centrales de

¹ Por su sigla en inglés *Structural Equation Modeling*.



interés. Y por otro lado, para estudiar y testear relaciones entre variables latentes en un modelo estructural que facilita comprender y valorar modelos teóricos (Nagengast y Trautwein, 2016).

Los SEM son una gran familia de métodos estadísticos para modelar relaciones entre variables. Específicamente, es una técnica usada para especificar y estimar un modelo de relaciones lineales entre variables, ya sean éstas variables observadas (VO) o latentes (VL). Las VL son constructos hipotetizados que no pueden ser medidos directamente, representados por múltiples VO que sirven como indicadores de éstos. Además de estos dos elementos, en un modelo se encuentra el error de medida, que representan los errores de las variables observadas y todas las fuentes de varianza residual que no están explícitamente representadas en el modelo; es decir, otros aspectos que mide el indicador además del factor. Los modelos de ecuaciones estructurales, se definen como un diseño hipotetizado de relaciones lineales direccionales o no direccionales entre un conjunto de VO y VL. Las relaciones direccionales implican algún tipo de influencia direccional de una variable sobre otra, en tanto que, las relaciones no direccionales son de tipo correlacional y no implican ninguna influencia dirigida (MacCallum y Austin, 2000; Golob, 2003; Hoyle, 2012; Orgaz Baz, 2008).

Un SEM completo cuenta de dos partes fundamentales, el *modelo de medida*, el cual permite al investigador evaluar que tan bien las variables observadas combinan para identificar el constructo hipotetizado, y el *modelo estructural*, en el cual la teoría, permite al investigador distinguir qué variables independientes predicen cada variable dependiente, y describir la interrelación entre constructos (Fernández Pulido, 2008; Cupani, 2012). Algunos casos especiales de SEM son de particular interés y son de uso general en la práctica. El Análisis Factorial Confirmatorio (AFC), el cual sólo contiene el modelo de medida, ampliamente usado en Psicología y en Educación para una variedad de propósitos, entre ellos, evaluar las propiedades psicométricas de medidas y garantizar la validez de constructo de escalas de medición (Luo, 2011; Tomarken y Waler, 2005). El análisis de rutas, el cual solo contiene el modelo de relaciones estructurales, teniendo la particularidad de no contener variables latentes, y en su lugar las variables observadas son igualadas con las variables latentes. Y, un modelo full, que integra tanto el modelo estructural como el de medida, observándose tanto relaciones entre las variables latentes y sus mediciones observadas, como relaciones entre las variables latentes.

Dada la frecuencia de uso de la técnica en el campo disciplinar de las Ciencias Sociales, es importante conocer los supuestos asociados. Si consideramos que, en su mayoría, las variables que se usan son de tipo categóricas, más específicamente ordinales, provenientes de recolecciones de datos con cuestionarios de tipo Likert; y que además, no presentan una distribución normal. Esta realidad muestra la primera dificultad de no contar con datos continuos que presenten a su vez una distribución normal multivariada. Por lo general, las referencias indican que los investigadores omitiendo estos supuestos básicos, tienden a aplicar el método de Máxima Verosimilitud -MV-; no reparando en la naturaleza de los datos (Luo, 2011; West, Finch y Curran, 1995; Gao, Mokhtarian y Johnston, 2008).

Sin embargo, Finney y DiStefano (2006) basados en resultados empíricos concretos, sugieren que si los datos observados poseen muchas categorías, al menos 5 opciones de respuesta y tienen a una distribución próxima a la normal, entonces usar el método MV como técnica de estimación no resulta tan desacertado; los índices de bondad de ajuste, los parámetros estimados y los errores estándar no presentan grandes distorsiones, en tanto es un método robusto. El problema se presenta, cuando el número de opciones de respuesta decrece, o el tamaño de la muestra es pequeño, características que se asocian a una menor probabilidad de que los datos observados se distribuyan normalmente, perjudicándose la calidad de los resultados obtenidos a partir de la técnica por MV.

En los casos mencionados, se recomienda otros métodos, tales como: mínimos cuadrados ponderados (WLS, *Weighted Least Squares*), mínimos cuadrados ponderados diagonalizados (DWLS, *Diagonal Weighted Least Squares*), mínimos cuadrados ponderados robustos (WLSMV, *Weighted Least Squares Mean and Variance Adjusted*), distribución libre asintótica (ADF, *Asymptotic Distribution-Free*), mínimos cuadrados no ponderados (ULS, *Unweighted Least Squares*), entre otros. Su disponibilidad, dependerá del software que se use para llevar a cabo los

análisis, por ejemplo, dentro del AMOS se encuentran disponibles tanto el ULS como el ADF, mientras que dentro del MPLus se dispone del WLSMV.

Respecto de las características de uno u otro método, varios estudios tienden a analizar las ventajas y desventajas. Como ya se ha mencionado, el método de MV, es el más utilizado, ya que a partir de él se pueden obtener todos los índices de bondad de ajuste y no requiere de grandes muestras, sin embargo exige la asunción de normalidad multivariada. Por su parte, la estimación por GLD (*Generalized Least Squares*) muestra identificadas características que el método por MV, aunque es inadecuado cuando el modelo aumenta en complejidad. Con respecto a los métodos que no requieren, y como ventaja, la presencia de normalidad multivariada, el USL presenta la dificultad de no calcular toda la variedad de índices de bondad de ajuste posibles; no obstante, los estimadores WLS y ADF, requieren de grandes tamaños muestrales, con n mayor a 200 (Hoffmann, Stover, de la Iglesia y Fernández Liporace, 2013; Luo, 2011; Muthén y Kaplan, 1985; Valdivieso Taborga, 2013).

La elección adecuada del método de estimación, es requisito previo para apreciar la bondad de ajuste entre los modelos propuestos y los datos empíricos. Específicamente, en lo que a ello concierne, la literatura recomienda el uso de múltiples indicadores para evaluar la calidad del ajuste del modelo. Entre los más destacados están la prueba de significación estadística de X^2 , el cual reporta un buen ajuste cuando se acepta la hipótesis nula ($p > 0,05$); el GFI (*Goodness of Fit Index*) y el CFI (*Comparative Fit Index*), cuyos valores oscilan entre 0 y 1, siendo aceptable valores cercanos o superiores a 0,90; y el RMSEA (*Root Mean Square Residual*), considerado óptimo cuando sus valores se encuentran entre 0,06 y 0,08, y del cual se recomienda informar el intervalo de confianza -IC- (Fernández Pulido, 2008; Orgaz Baz, 2008; Ruiz, Pardo y San Martín, 2010).

En el marco de la propuesta y de uso cada vez más frecuente de los SEM en áreas de educación y psicología, el artículo busca mostrar qué decisiones y caminos el investigador puede tomar frente a los datos, evaluando fortalezas y limitaciones. Buscamos hacerlo en un lenguaje claro, no matemático, sin formulas, ni cálculos. Nos inclinamos por un artículo de tipo didáctico que muestre los procedimientos para realizar un AFC, cuando se pertenece a un área disciplinar ajena al campo de la estadística; pero su uso para hacer análisis de datos y avanzar en desarrollos teóricos son un pilar clave del área de estudio. No obstante, el artículo no pretende profundizar en cada paso sino introducir a noveles científicos.

Antesala del AFC, decisiones sobre la marcha e interpretación de resultados

A continuación mostramos los pasos y las decisiones que se toman frente a los datos, los resultados y las posibilidades que se ofrecen desde el material y la bibliografía específica consultada. Específicamente, pondremos a prueba un modelo de medida a través de un AFC a modo de mostrar una herramienta útil para validar instrumentos provenientes de áreas educativas y psicológicas. Recordemos que el AFC busca explicar la correlación entre un grupo de variables observadas en términos de un pequeño grupo de factores latentes (Luo, 2011). Se consideran diversas etapas para aplicar esta técnica, que a continuación nombramos, definimos y ejemplificamos.

Usaremos como modelo la estructura del MIDAS -*Multiple Intelligences Developmental Assessment Scales*-, un cuestionario tipo Likert que mide las Inteligencias Múltiples en el marco de la teoría propuesta por Gardner (1983), desarrollado por Shearer (1987). El cuestionario está compuesto por 119 ítems distribuidos en 8 escalas iniciales y 26 sub-escalas en su versión para adultos y adolescentes (ver cuadro 1), ofreciendo 6 alternativas de respuesta para describir el grado en el que la habilidad consultada se presenta en el desarrollo personal de cada sujeto. Instrumento que fue aplicado y contestado a una muestra de 638 estudiantes universitarios, de diversas carreras y años de cursado.

¹ Para mayor detalle, puede consultar el manual del autor de MIDAS, Shearer (1987).

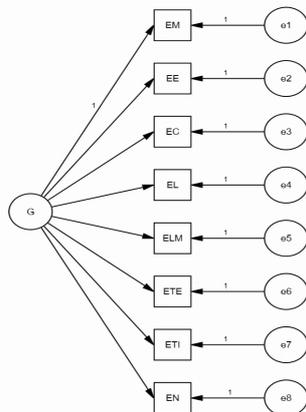


Cuadro 1. Resumen de escalas y sub-escalas del MIDAS (1987)

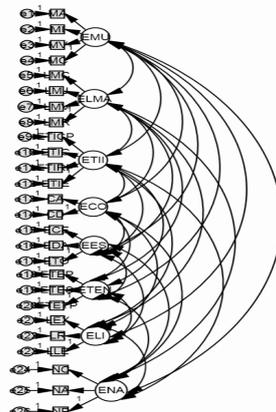
Escalas	Sub-escalas
Musical	Apreciación, Habilidad Vocal, Habilidad Instrumental, Composición.
Cinestésica-Corporal	Atletismo, Movimientos Expresivos.
Lógica-Matemática	Juegos Estratégicos; Habilidad Diaria para las Matemáticas; Resolución de Problemas Diarios, Matemáticas en la Escuela.
Espacial	Conocimiento Espacial, Diseño Artístico, Modelación de Objetos.
Lingüística	Habilidad Retórica, Sensibilidad Expresiva, Escrita-Lectura Académica.
Interpersonal	Sensibilidad Social, Persuasión Social y Relaciones Interpersonales.
Intrapersonal	Conocimiento Personal, Auto-Eficacia, Meta-Cognición, Resolución de Problemas.
Naturalista	Cuidado de Animales, de Plantas, Ciencia.

En la primera etapa de un AFC, el investigador especifica el modelo tratando de comprobar una hipótesis. En nuestro ejemplo, para analizar la validez de constructo del MIDAS se formulan hipótesis acerca de la estructura que subyace a los datos y que pueden ser compatibles con distintas alternativas teóricas sobre inteligencia. Una de ellas, podría postular que las ocho habilidades que mide el MIDAS se definen bajo una estructura unidimensional correspondiente a la teórica del factor general de inteligencia formulada por Spearman (1904), a la cual llamaremos Modelo A. Otra, podría proponer una estructura que defina ocho factores independientes, coherentes con los desarrollos teóricos de la propuesta inicial de Gardner (1983; 2001), a la cual denominaremos Modelo B, o bien proponer una estructura cuatridimensional, agrupando las inteligencias en áreas principales -académicas, artísticas, emocional, naturalista- afín a la descripción que se ofrece en el libro *La Mente Reformulada* (Gardner, 2001), a la cual nombraremos como Modelo C. Cabe destacar que, en el primero se utilizan como variables observadas las 8 escalas del MIDAS, mientras que en el segundo y último modelo se utilizan las sub-escalas como tales, a continuación presentamos los distintos diagramas de flujo (*path diagram*) para cada modelo propuesto:

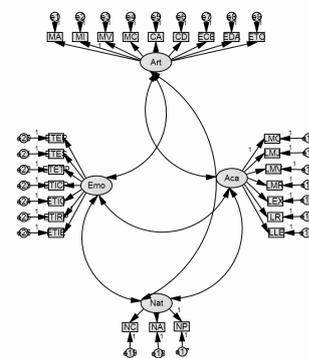
Modelo A



Modelo B



Modelo C



Como segundo paso, se procede a la identificación del modelo, etapa fundamental, en tanto permite conocer y asegurar que pueden ser estimados los parámetros del modelo propuesto. Esta fase está asociada al concepto de grados de libertad (gl), y actualmente los softwares brindan dicha información, aunque los cálculos se podrían realizar a mano -gl disponible: $VO (VO+1)/2$;

donde VO es el número de variables observadoras del modelo propuesto-. Concretamente, la identificación del modelo viene dada por la diferencia entre los gl disponibles y los consumidos (o número de parámetros a estimar). Un modelo estará sobreidentificado, cuando los gl sean >0 , suponiendo que este modelo es uno entre otras posibles alternativas, mientras que cuando los gl sean negativos, se asumirá que el modelo no está identificado, es decir, excede los grados de libertad disponibles, y si los gl fuesen igual a 0 entonces el modelo está saturado; en otras palabras, no cumple el criterio de simplificación de los datos observados, y aunque el modelo tuviera un ajuste adecuado, no tendría sentido, siendo un modelo único, sin opciones de ningún tipo.

En nuestro ejemplo, para cada hipótesis la identificación del modelo es:

Modelo	gl disponibles	GL consumidos	Diferencia
A - Un factor	36	16	20
B - Ocho Factores	351	80	271
C- Cuatro factores	351	58	293

Observamos que, en los tres modelos la diferencia es mayor a 0, con lo cual es posible pasar a la siguiente etapa de modelización.

La tercera fase es quizás la más compleja y la que mayor consideración puede llevar, consiste en la estimación de los parámetros en el modelo de medida propuesto.

Como primer paso se recomienda evaluar la calidad de la base de datos con la que se cuenta: el tamaño de la muestra, los datos perdidos, la multicolineidad de la muestra, los casos atípicos multivariados y la distribución normal multivariada; que pasaremos a analizar sobre la muestra de 638 sujetos.

Respeto del tamaño de la muestra, se considera adecuado tener entre 10 o 20 unidades de análisis por variable observada, sin embargo varios autores recomiendan un tamaño mínimo de 200 sujetos para cualquier tipo de SEM (Ferrando y Anguiano Carrasco, 2010). En uno u otro caso, con n de 638 sujetos se cumple con lo requerido. Cabe destacar que el número de sujetos se redujo en 12, al detectar la presencia de datos perdidos en algunas unidades de análisis correspondiente a 5 sub-escalas del MIDAS.

Para el diagnóstico de la multicolinealidad de las variables observadas que ingresaremos a los modelos se utiliza el test de esfericidad de Bartlett y la prueba de Kaiser-Meyer-Olkin (KMO). En el test de Bartlett, se busca rechazar la H_0 , la cual plantea que la matriz de correlación poblacional es idéntica; es decir, que las variables no están correlacionadas en la población, y en la prueba de KMO, se busca informar sobre el grado de relación conjunta entre las variables; resultados cercanos a 1, indican que las relaciones entre las variables son más sustanciales o que no hay tantos factores como ítems tenga el instrumento, en general se acepta un valor igual o superior a 0,80, indicando que la matriz de correlación es apropiada para factorizar.

Particularmente, ambas pruebas fueron aplicadas a todas aquellas variables observadas que forman parte de los tres modelos propuestos, por un lado a las 8 escalas y por otra parte a las 26 sub-escalas. La medida de adecuación muestral de Kaiser-Meyer-Olkin fue de 0,804 y 0,823 respectivamente y la prueba de esfericidad de Bartlett arrojó un p-valor $<0,0001$ en ambos casos, estos resultados confirman que la comunalidad entre los ítems, escalas y sub-escalas es baja.

En lo que respecta a la presencia de datos extremos (*outliers*) se analiza a través de la distancia de Mahalanobis al cuadrado (D^2), una medida estadística de la distancia multidimensional de un individuo respecto al centroide o media de las observaciones, contemplada en la mayoría de los paquetes estadísticos más usados en la actualidad. Un dato extremo multivariado son casos con puntuaciones extremas en dos o más variables. La distancia encontrada se informa junto a un valor la probabilidad, el cual indica las chances de que un individuo cualquiera esté a una distancia mayor o igual a D^2 . Cuando esa probabilidad es cercana a 0,001, indica presencia de atipicidad; es decir, una combinación muy rara entre las variables en estudio para un caso o varios casos en particular (González, Abad y Lévy, 2006). Al respecto, la experiencia nos indica que se puede ser aún más exigente y eliminar aquellos casos con una probabilidad igual o menor 0,10, dado que los casos atípicos multivariados no solo afectan a la media y al desvío sino también a las correlaciones y la distribución normal multivariada (Arias,



2008; Gao, Mokhtarian y Johnston, 2008; Muñoz García y Amón Uribe, 2013; West, Finch y Curran, 1995).

Se procedió a calcular la D^2 con nuestra base de datos, sobre todo sabiendo que el instrumento es sensible a la identificación de casos atípicos (Rigo y Donolo, 2012), comprobando la presencia de 24 casos con puntuaciones marginales multivariadas (como se muestra en parte en la tabla 1), los cuales fueron quitados de la muestra, quedando un $n=602$ sujetos.

Tabla 1.
Distancia de Mahalanobis para la muestra de 626 sujetos

Caso n°	D^2	p
590	233,403	,000
58	83,949	,000
10	81,114	,000
51	56,316	,001
38	55,154	,001
50	54,030	,001
117(...)	53,934	,001
(...)327	53,578	,001

Cabe destacar, que la eliminación de los casos atípicos multivariados o su tratamiento, es condición necesaria y previa a comprobar la normalidad multivariada.

Al respecto, determinar si la muestra presenta o no una distribución normal multivariada es el paso previo a decidir el método de estimación de parámetros por el cual se optará. West, Finch y Curran (1995), menciona que la normalidad univariada es una condición necesaria pero no suficiente, es más puede que las distintas variables no presenten una distribución normal univariada, pero todas juntas muestren una multivariada. La prueba usada para asegurar la presencia de normalidad multivariada en los datos recolectados es el coeficiente de curtosis multivariada de Mardia. En lo que alude a su interpretación, los valores críticos (c.r.) asociados al coeficiente deberían de caer dentro del intervalo $[-1,96 - 1,96]$ para comprobar normalidad multivariada. Mientras que, el contraste de normalidad univariada, se valora con las pruebas de Kolmogorov-Smirnov-Lilliefors y de Shapiro-Wilks.

Cuando el procedimiento se llevó con la muestra que aún conservaba los 24 casos extremos multivariados identificados, el valor crítico obtenido fue mayor a 2 (c.r.=2.354), revelando falta de normalidad multivariada en los datos. No obstante, cuando se eliminan los casos atípicos, el c.r.<1.96, presentó un valor de 1,008, comprobándose el supuesto de normalidad multivariada (sobre el modelo que mejor ajusta, ver Tabla 2).

Finalmente, habiendo aumentado la calidad de la base de datos, se está frente a la tarea de seleccionar el método para la estimación de parámetros a usar, en este caso, dada la revisión realiza y la presencia de una distribución normal multivariada de los datos se puede aplicar el de Máxima Verosimilitud (MV) para estimar los valores de los parámetros desconocidos, y sus respectivos errores de medición. Y luego poder atender a la lectura de los parámetros estandarizados, con respecto a la bondad de los modelos propuestos, a fin de determinar cuál es el que mejor se ajusta a los datos muestrales obtenidos. Ambas etapas, estimación de parámetros y evaluación de ajuste del modelo se dan de manera simultánea.

Se aplica el método de MV para realizar el AFC, para cada uno de los modelos propuestos: A, B y C, en la Tabla 2 se aprecian los valores obtenidos para cada índice de bondad de ajuste considerado.

Tabla 2
 Índices de bondad de ajuste para cada modelo. AFC. Instrumento MIDAS, n= 602

Modelo	X^2	gl	p	GFI	CFI	RMSEA (IC al 90% de confianza)
A	448,139	20	0,0001	0,835	0,803	0,189 (0,174 - 0,204)
B	1047,605	271	0,0001	0,888	0,850	0,069 (0,065 - 0,073)
C	2007,991	293	0,0001	0,759	0,669	0,098 (0,094 - 0,102)

Como se aprecia, el Modelo B es el que mejor ajuste presenta a los datos empíricos. No obstante, cabe hacer algunas aclaraciones al respecto. Observamos que en ninguno de los tres modelos la prueba de significación estadística X^2 , ofrece un nivel de significación superior a 0,05, lo cual indica que la estructura teórica propuesta en el modelo es significativamente distinta a la indicada en la matriz de covarianza de los datos, no obstante, se recomienda usar otras pruebas, dado que raramente es usado como única prueba o concluyente para interpretar el ajuste del modelo. La decisión es tal, en tanto es una prueba muy sensible al tamaño de la muestra, y tiende a rechazarse la H_0 cuando el $n > 200$ (Fernández Pulido, 2008). Por este motivo, otro criterio adoptado para evaluar el ajuste global del modelo es mirar no tanto el nivel de significación obtenido, sino la razón X^2/gl , donde valores menores a 5 corresponderían a un ajuste razonable (Marsch y Hocevar, 1985). Claramente, el Modelo B satisface este último criterio. En función de los diferentes valores obtenidos en los diversos estadísticos de bondad de ajuste, el Modelo B presenta el mejor ajuste aceptable en función de los criterios de referencia que se manejan en la literatura. Proponiendo una estructura de 8 factores subyacentes y como indicadores observados cada una de las 26 sub-escalas que integran al MIDAS. A esta interpretación se le debe sumar la lectura e interpretación de los parámetros estimados estandarizados que se presentan en la Figura 1. En ella se muestra la solución estandarizada hallada para el modelo de ocho factores.

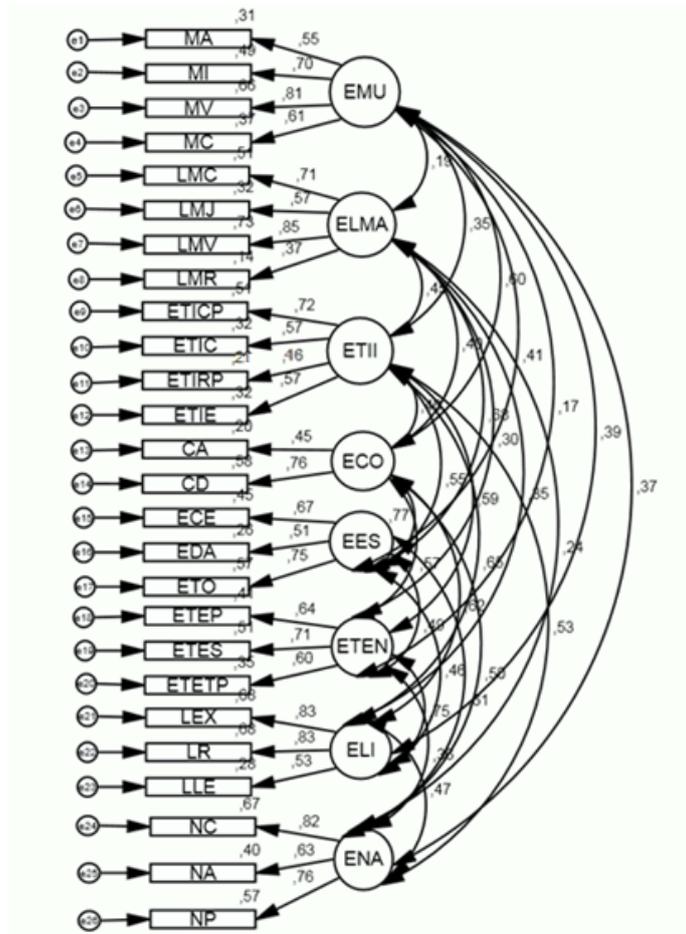


Figura 1. Diagrama de flujos (path diagram) del Modelo B - Análisis Factorial Confirmatorio del MIDAS

La lectura se realiza sobre las estimaciones estandarizadas, en tanto la covarianza, los pesos factores y la correlación múltiple al cuadrado no estandarizados son difíciles de explicar e interpretar en tanto sus valores están sujetos a la unidad de medida de las variables (Arbuckle, 2007).

En el diagrama de flujo, las flechas que unen las variables latentes con los indicadores muestran las relaciones entre las variables observadas o indicadores y las variables latentes, especificadas a partir de unos coeficientes que en este caso se denominan cargas o pesos factoriales. Mientras que, la relación entre dos variables latentes se dibuja con flechas de doble punta y se lee como 'r' o correlación. En lo que respecta a los valores que se ubican por encima de los rectángulos, representan la proporción de la varianza del ítem explicada por el factor de pertenencia, y refiere a la correlación múltiple al cuadrado (R^2).

De este modelo es interesante resaltar que presenta pesos factoriales por encima de 0,45 mostrando la importancia que tiene el factor en la determinación de la variable observada, al igual que lo muestran las correlaciones múltiples al cuadrado, las cuales expresan el porcentaje de varianza que explica el factor respecto de las puntuaciones en una sub-escala. Las correlaciones entre los factores son todas significativas y positivas, de mayor magnitud entre todos los factores respecto a los ETTII (inteligencia intrapersonal) y ETEN (inteligencia interpersonal), las de menos magnitud refieren a los factores que representan a las inteligencias artísticas, a saber EMU (inteligencia musical), ECO (inteligencia corporal) y EES (inteligencia espacial). Asimismo, se observan una

fuerte correlación entre los factores que representan a la inteligencia lingüística (ELI) y lógica-matemática (ELMA), resultados que coinciden con investigaciones previas con grupos semejantes (Rigo & Donolo, 2014).

Con lo cual estamos en condiciones de apreciar que el modelo confirma la propuesta que tanto de Gardner (1983; 2001), como Shearer (1987) comparten en sus formulaciones sobre la existencia de ocho factores definidos por cada una de las inteligencias múltiples propuestas en la teoría de base. Ocho inteligencias correlacionadas que se definen en función de una muestra particular de estudiantes universitarios de cinco facultades -Ciencias Humanas, Exactas, Agrarias, Económicas e Ingeniería-, un modelo de los tantos posibles de encontrar.

Asimismo, otro aspecto a mirar es la significación de todos los estimadores para concluir con esta etapa. Particularmente encontramos que todos son significativos ($p < .01$), lectura que se realiza en la salida de datos no estandarizadas de los pesos factoriales, covarianza y varianza (ver Tabla 3).

Tabla 3
Solución no estandarizada de la covarianza entre variables exógenas

			Estimación	Sd	c.r.	p
ELI	<-->	ENA	1,469	,208	7,062	***
ETII	<-->	ECO	1,110	,178	6,237	***
EES	<-->	ETEN	1,378	,195	7,062	***
ETEN	<-->	ELI	1,590	,196	8,116	***

La primera estimación que se muestra en la Tabla 3 refleja el valor de la covarianza entre la escala musical y naturalista, seguida del desvío estándar, el valor crítico -C.R.- y el p-valor (para dos colas). Cuando el c.r. excede el valor de 1,96 podría concluirse que es significativo, en este caso es igual a 7,062 $>$ 1,96, lo que es igual a decir que la covarianza entre ELI y ENA es significativamente diferente de cero a un nivel de significación de 0,001 (Arbuckle, 2007).

Lejos de concluir en la interpretación del modelo, el investigador que lleva a cabo el AFC puede, guiado por índices de modificación reajustar el modelo; sin embargo, esta última etapa no es muy recomendada, tal es así que Ruiz, Pardo y San Martín (2010), explican que el reajuste de las posibles relaciones teóricas, incluyendo y excluyendo factores y variables de manera tentativa deriva en modelos inestables y que pierden sus buenas propiedades cuando se replican con otras muestras. En parte como consecuencia de la dificultad de justificar teóricamente la incorporación de tales modificaciones al modelo.

Anotaciones adicionales

A continuación se comentan algunas interpretaciones comunes en el mundo de los SEM y sobre todo referidas a la técnica de AFC, tomar notas ayuda a comprender la sutileza, limitaciones y recaudos a considerar cuando hacemos uso de ellas.

Generalización Vs. Limitación. La bibliografía consultada muestra que aunque el modelo propuesto presente un ajuste aceptable, el investigador no ha probado el modelo en sí, solo ha confirmado que es uno de los varios modelos posibles. En el proceso de validación, la muestra y su conformación tienen un rol importante. Al respecto, MacCallum y Austin (2000) y Tomarken y Waler (2005), resaltan la importancia de considerar que, el uso de las conclusiones derivadas del SEM está limitado a la muestra, las variables y la ocasión en la cual se tomaron las medidas. Tales limitaciones rara vez se reconocen. Por el contrario, los resultados suelen interpretarse como generalizables.

Varios Vs. Uno. La importancia de que los investigadores generen y evalúen modelos alternativos, ofrece protección contra un sesgo de confirmación y ayuda a no reforzar el apoyo de un modelo único y favorito (MacCallum & Austin, 2000; Tomarken y Waler, 2005), por tanto, es honesto formular y contrastar varios modelos soportados por la teoría o por corrientes teóricas opuestas.



Interpretación Solista Vs. Integral. Para decidir que un modelo es bueno, no solo hay que comprender las diversas familias de índices de ajuste, también hay que mostrar y analizar los parámetros estimados. De hecho, es perfectamente posible encontrar que las relaciones entre variables siendo débiles, o incluso cercanas a cero, junto a un modelo con excelente ajuste. Por lo tanto, es importante prestar atención a las estimaciones de los parámetros, incluso cuando los distintos índices reporten un buen ajuste.

Grande Vs. Pequeña. Los estudios con muestras pequeñas no son frecuentes, la revisión realizada por MacCallum y Austin (2000) muestra que alrededor del 18% de los estudios usan muestras con menos de 100 individuos. En general, las publicaciones rara vez incluyen una consideración explícita de si la muestra disponible es lo suficientemente grandes. Al respecto, Ruiz, Pardo y San Martín (2010), aluden a la importancia del tamaño de la muestra sobre la bondad de ajuste, cuanto más pequeño es el n , se está frente a mayores posibilidades de que los valores de ajuste no se deterioren. Una muestra adecuada supera los 200 casos, mientras que aquellas configuradas con menos de 100 unidades de análisis no son una buena garantía del modelo estimado. Asimismo, hay que tener presente que cuanto mayor es el número de variables (mayores a 20) es más difícil reproducir las varianzas observadas y mayor debe ser el tamaño muestral (una razón de 10 sujetos por cada variable observada es lo aconsejado).

A modo de consideraciones finales

En la actualidad, las técnicas de análisis multivariante tiene un gran impacto en diversos campos disciplinares, incluida la psicología y la educación, en correspondencia con el avance de uso de paquetes estadísticos más sofisticados que permiten con un solo clic acceder a los diversos análisis incluidos en los SEM, como el AFC. No obstante, ese clic debe estar orientado por conocimientos que guíen los pasos a seguir y las interpretaciones a realizar.

En la línea de guiar el proceso de hacer escribimos este artículo, a modo de portafolio en el cual se pueda encontrar aquellas herramientas, caminos y atajos que sirvan a los investigadores interesados en aplicar la técnica, con las sutilizas que el caso amerite, sobre todo insertos en las Ciencias Sociales.

El desafío no termina acá, los análisis multivariados cada vez son más justificados en el ámbito psicológico y educativo, tal es así que no sorprende que la tercera edición del *Handbook of Educational Psychology* (Corno y Anderman, 2016), incluya un capítulo sobre Modelos de Variables Latentes, sus limitaciones y ventajas, queda claro en el camino realizado que las relaciones bivariadas exceden la complejidad de las dimensiones que explican y ayudan a transformar no solo la teoría sino las prácticas educativas. Al respecto, Rigo (2016), expresa que los diseños de investigación y las técnicas en Psicología Educativa van evolucionando para atender a la necesidad de crear puentes entre las variables que definen a los docentes, las tareas académicas y los alumnos, como modo de estudiar las interrelaciones entre los aspectos que definen a la clase.

Agradecemos a aquellos que en la tarea de aprender se mostraron abiertos a consultas y generosos en la tarea de compartir conocimientos y experiencias, ya sea en un contexto presencial, contestando un e-mail o enviando un artículo, en esto las nuevas tecnologías de la comunicación y las redes sociales, tales como *ResearchGate*, tiene un rol fundamental en la construcción de una comunidad científica más solidaria y con responsabilidad social.

Referencias

- Arbuckle, J. (2007). *Amos™ 18 User's Guide*. Chicago: Amos Development Corporation.
- Arias, B. (2008). Desarrollo de un ejemplo de análisis factorial confirmatorio con LISREL, AMOS y SAS. En M., Verdugo, M. Crespo, M. Badía y B. Arias (Coord). *Metodología en la investigación sobre discapacidad. Introducción al uso de las ecuaciones estructurales*, (pp. 75-120). Salamanca: Publicaciones del INICO.
- Corno, L. & Anderman, E. (2016). *Handbook of Educational Psychology*. New York: Routledge.
- Cupani, M. (2012). Análisis de ecuaciones estructurales: conceptos, etapas de desarrollo y un ejemplo de aplicación. *Revista Tesis, 1*, 186-199.
- Fernández Pulido, R. (2008). Modelos de medida y análisis factorial confirmatorio. En M., Verdugo, M. Crespo, M. Badía y B. Arias (Coord). *Metodología en la investigación sobre discapacidad. Introducción al uso de las ecuaciones estructurales*, (pp. 29-41). Salamanca: Publicaciones del INICO.
- Ferrando, P. & Anguiano Carrasco, C. (2010). El análisis factorial como técnica de investigación en psicología. *Papeles del Psicólogo, 31*(1), 18-33.
- Finney, S. & DiStefano, C. (2006). Non-normal and Categorical data in structural equation modeling. En G. Hancock y R. Mueller (Edit.). *Structural equation modeling: a second course* (pp. 269–314). Greenwich, Connecticut: Information Age Publishing
- Gao, S., Mokhtarian, P. & Johnston, R. (2008). Non-normality of data in structural equation models. *Transportation Research Record Journal of the Transportation Research Board, 2082*, 116-124.
- Gardner, H. (1983). *Frames of mind: The theory of multiples intelligences*. New York: Basic.
- Gardner, H. (2001). *La inteligencia reformulada. Las inteligencias múltiples en el siglo XXI*. Barcelona: Paidós.
- Golob, T. (2003). Structural equation modeling for travel behavior research. *Transportation Research Part B: Methodological, 37*, 1-25.
- González, N., Abad, J. & Lèvy, J.P. (2006). Normalidad y otros supuestos en análisis de covarianzas. En J. Lévy. (Edit.): *Modelización con estructuras de covarianzas* (pp. 31-57). La Coruña: Netbiblo.
- Hershberger, S. (2003). The growth of structural equation modeling: 1994-2001. *Structural Equation Modeling, 10*(1), 35-46.
- Hoffmann, A., Stover, J., de la Iglesia, G. & Fernández Liporace, M. (2013). Correlaciones policóricas y tetracóricas en estudios factoriales exploratorios y confirmatorios. *Ciencias Psicológicas, VII*(2), 151-164
- Hoyle, R. (2012). *Handbook of structural equation modeling*. New York: The Guilford Press.
- Khine, M. (2013). *Application of structural equation modeling in educational research and practice*. Rotterdam: The Netherland.
- Luo, H. (2011). Some aspects on confirmatory factor analysis of ordinal variables and generating non-normal data. Uppsala, Suecia: Universidad de Uppsala. Recuperado de: <http://www.diva-portal.org/smash/get/diva2:405108/FULLTEXT01.pdf>
- MacCallum, R. & Austin, J. (2000). Applications of structural equation modeling in psychological research. *Annu. Rev. Psychol, 51*, 201-226.
- Marsch, H. & Hocevar, D. (1985). Application of Confirmatory Factor Analysis to the Study of Self-Concept. First- and Higher Order Factor Models and Their Invariance Across Groups. *Psychological Bulletin, 97*(3), 562-582.
- Muñoz García, J. & Amón Uribe, I. (2013). Técnicas para detección de outliers multivariantes. *Revista en Telecomunicaciones e Informática, 3* (5), 11-25.
- Muthén, B. & Kaplan, D. (1985). A comparison of some methodologies for the factor analysis of non-normal Likert variables: a note on the size of the model. *British Journal of Mathematical and Statistical Psychology, 38*, 171-189.
- Nagengast, B. & Trautwein, U. (2016). The prospects and limitations of latent variable models in Educational Psychology. En L. Corno y E. Anderman (Edit.). *Handbook of Educational Psychology* (pp. 41-58). New York: Routledge.



- Orgaz Baz, B. (2008). Introducción a la metodología SEM: conceptos y propósitos fundamentales. En M., Verdugo, M. Crespo, M. Badía y B. Arias (Coord). *Metodología en la investigación sobre discapacidad. Introducción al uso de las ecuaciones estructurales*, (pp. 14-28). Salamanca: Publicaciones del INICO.
- Rigo, D. (2016). El compromiso en el aula de educación primaria. Entre rasgos personales y factores contextuales. En G. Maldonado, A. Bono y D. Sigal (Comp.). *Qué investigamos en la Facultad de Ciencias Humanas. Actas de las Jornadas de Investigación de la Facultad de Ciencias Humanas 2015* (pp. 294-307). Río Cuarto, Universidad Nacional de Río Cuarto: UNIRIO. Recuperado de: <https://www.unrc.edu.ar/unrc/comunicacion/editorial/repositorio/978-987-688-191-3.pdf>
- Rigo, D. & Donolo, D. (2012). ¿De qué modo somos inteligentes? Resultados para pensar la educación! *Revista Cultura y Educación*, 24(1): 5-15.
- Rigo, D. & Donolo, D. (2014). *Inteligencias. Teorías recientes, Creencias arraigadas y desempeños sociales*. Colección Cuadernos de Educación N° 2. La Laguna, Tenerife: Sociedad Latina de Comunicación Social. Recuperado de: <http://issuu.com/revistalatinadecomunicacion/docs/cde02#/signin>
- Ruiz, M., Pardo, A. & San Martín, R. (2010). Modelos de educaciones estructurales. *Papeles del Psicólogo*, 31(1), 34-45.
- Shearer, B. (2007). *The MIDAS: A professional manual*. Kent, Ohio: Research and consulting Inc.
- Spearman, Ch. (1904). General intelligence, objectively determined and measured. *American Journal of Psicología*, 15, 201-293.
- Tomarken, A. & Waler, N. (2005). Structural equation modeling: strengths, limitations, and misconceptions. *Annual Review of Clinical Psychology*, 1, 31-65.
- Valdivieso Tabora, C. (2013). Efecto de los métodos de estimación en las modelaciones de estructuras de covarianzas sobre un modelo estructural de evaluación del servicio de clases. *Comunicaciones en Estadística*, 6(1), 21-44.
- West, S., Finch, J. & Curran, P. (1995). Structural equation models with nonnormal variables. Problems and Remedies. En R. Hoyle (Edit.). *Structural equation modeling: Concepts, issues and applications*, (pp. 56-75). Thousand Oaks: Sage.

Received: 03/14/2017
Accepted: 02/26/2019