

Entre lo observable y lo latente: modelos de ecuaciones estructurales e investigación social

Between the observable and the latent: structural equations models and social research

Andrés Castillo Vargas

Universidad de Costa Rica, Costa Rica


andres.castillo@ucr.ac.cr

 <https://orcid.org/0009-0003-5794-3616>

Andrés Ruiz Sánchez

Universidad de Costa Rica, Costa Rica

andres.ruiz@ucr.ac.cr

 <https://orcid.org/0000-0003-2473-4626>

Recepción: 12 Enero 2024
Aprobación: 21 Agosto 2024



Acceso abierto diamante

Resumen

Introducción: El desarrollo de modelos y teorías que aborden los fenómenos socioculturales ha sido un reto para la investigación empírica en Ciencias Sociales, pues al trabajar con conceptos no experimentales y poco controlables, se ha vuelto necesaria la creación de metodologías más certeras para el estudio de conceptos no físicos y abstractos, entre las cuales se encuentran los Modelos de Ecuaciones Estructurales (MEE).

Objetivo: El objetivo es analizar el potencial de los MEE para mejorar la comprensión de fenómenos socioculturales en disciplinas pertenecientes a las Ciencias Sociales.

Método y técnica: En primera instancia, se realizó una revisión bibliográfica en torno a los orígenes de los MEE. Posteriormente, se detalló la estructura y etapas para su construcción, lo cual fue seguido de un análisis sobre las contribuciones que estos modelos ofrecen a la investigación social.

Resultados: Los MEE permiten comprender relaciones complejas entre variables propias de las Ciencias Sociales y considerar los errores de medición en los datos. Esto posibilita analizar simultáneamente varias relaciones de dependencia, a diferencia de otros enfoques como la regresión múltiple o análisis multivariantes.

Conclusiones: Los MEE favorecen la actividad investigativa y propician la generación de nuevos aprendizajes que contribuyen al entendimiento e interpretación de diversos fenómenos. Esto enfatiza la importancia de avanzar constantemente en el conocimiento de las complejidades inherentes a los procesos sociales, en aras de develar la relación existente entre lo observable y lo latente.

Palabras clave: Idealizaciones, Variables latentes, Relaciones causales, Teorías científicas, Modelos estadísticos.

Abstract

Introduction: The development of models and theories addressing sociocultural phenomena has been a challenge for empirical research in Social Sciences. Working with non-experimental and less controllable concepts has made it necessary to create more accurate methodologies for studying non-physical and abstract concepts, among which Structural Equation Models (SEM) are found.

Objective: The aim is to analyze the potential of SEM to enhance the understanding of sociocultural phenomena within disciplines belonging to the Social Sciences.

Method and technique: Initially, a bibliographic review was conducted regarding the origins of SEM. Subsequently, the structure and stages for their construction were detailed, followed by an analysis of the contributions these models offer to social research.

Results: SEM allows for the understanding of complex relationships among variables inherent to social sciences and accounts for measurement errors in data. This enables the simultaneous analysis of multiple dependency relationships, unlike other approaches such as multiple regression or multivariate analysis.

Conclusions: SEM is favorable for investigative activity and fosters the creation of new insights that contribute to the understanding and interpretation of various phenomena. This emphasizes the importance of constantly advancing knowledge related to the inherent complexities of social processes, to unveil the relationship between the observable and the latent.

Keywords: Idealizations, Latent variables, Causal relationships, Scientific theories, Statistical models.

Introducción

Tradicionalmente, uno de los elementos que impulsan el avance de la ciencia es el desarrollo de metodologías que establecen criterios mínimos de excelencia y rigurosidad en la búsqueda del conocimiento. No obstante, existe un componente esencial dentro de este proceso, al cual se le ha otorgado poca atención y, sin embargo, es indispensable en el perfeccionamiento de teorías y modelos científicos; este elemento es conocido como las idealizaciones.

Las idealizaciones consisten en asumir un hecho como real con el objetivo de explicar una teoría científica. Esta noción de gran utilidad cuando se carece de determinados datos o de teorías auxiliares que fortalezcan o validen un modelo a estimar (Laymon 1989; Cuevas y Martínez 2011). Las idealizaciones no solo son necesarias porque ayudan a la producción de conocimiento científico y a construir teorías, sino también porque facilitan el acercamiento a un objeto de estudio en particular. Permiten obtener predicciones o explicaciones con base en circunstancias o modelos idealizados.

En cierta manera no habría progresividad científica sin el uso de las idealizaciones, pues las mismas evidencian como gracias a la identificación de un mundo, contexto o característica “ideal” es posible comprender el mundo real y acceder al conocimiento científico, es decir, son una manera “poco compleja o simplificada” de comprender las teorías y modelos científicos (Laymon 1989). Ahora bien, aunque las idealizaciones son importantes en el desarrollo de la ciencia, su uso debe ser prudente, pues como expone Laymon (1989), puede que no se conozca con certeza si las idealizaciones propuestas se encuentran muy cerca o muy lejos de la verdad. De allí, surge la importancia de realizar idealizaciones cada vez más realistas, a pesar de lo contradictorio del adjetivo, que contribuyan al establecimiento de predicciones más exactas en relación con una teoría científica determinada.

Considerando la relación existente entre las idealizaciones, la ciencia y los modelos científicos, las teorías científicas pueden ser entendidas de dos maneras; como «un conjunto (generalmente infinito) de enunciados –hipótesis y leyes- sobre un aspecto de la realidad, ordenados sistemáticamente mediante relaciones de deductibilidad que pueden someterse a contrastación» (Cuevas y Martínez 2011, 26); o como «sistemas ideales constituidos, mediante definición, por [un] conjunto de enunciados o leyes teóricas que afirman o niegan que un sistema real empírico está de acuerdo o no con el modelo [teórico de referencia]» (Cuevas y Martínez 2011, 27).

De acuerdo con Bunge (1981), las teorías científicas son sistemas lógicos y deductivos que permiten hacer predicciones comprobables por medio de la observación y la experimentación. Para dicho investigador, las teorías son afines a los modelos científicos, los cuales son definidos como una posible representación conceptual, esquemática e hipotética de una cosa o de una sustancia real o, así, presumida. Al respecto, Guevara y Valdez (2004) argumentan que la característica principal de un modelo es ser una construcción imaginaria que reemplaza a un elemento de la realidad con el fin de representar de una manera simplificada un aspecto de un sistema teórico.

Los modelos tienen una función central en el desarrollo de teorías pues, permiten establecer relaciones entre el lenguaje y los objetos por medio de enunciados, indicadores, ecuaciones o diagramas, representados de manera conceptual o gráfica (Cuevas y Martínez 2011). Es así, que un modelo debe facilitar la visualización o comprensión conceptual del objeto o cosa modelada (Guevara y Valdez 2004). Por ejemplo, en las teorías relacionadas con la lógica existen modelos de árboles deductivos; en la teoría atómica existen diagramas de densidad de probabilidades; y en biología matemática se encuentran grafos dirigidos que enlazan distintas funciones biológicas.

El desarrollo de modelos y teorías que aborden los fenómenos socioculturales ha sido un reto para la investigación empírica en Ciencias Sociales, pues al trabajar con conceptos no experimentales y poco controlables, se ha vuelto necesaria la creación de distintas metodologías y análisis de datos que permitan un

acercamiento cada vez más certero a todos aquellos conceptos no físicos y abstractos, conocidos comúnmente como constructos (Casas Guillén 2002). Precisamente, una de estas metodologías son los Modelos de Ecuaciones Estructurales (MEE), los cuales son un buen ejemplo de cómo las idealizaciones, las teorías y los modelos científicos convergen empíricamente dentro de la investigación social; justo en las siguientes líneas se caracterizarán algunos aportes que este tipo de modelos han realizado al estudio del comportamiento humano. Tomando en consideración el papel que los MEE emplean actualmente en las Ciencias Sociales, este artículo analiza el potencial de los MEE para mejorar la comprensión de fenómenos socioculturales en disciplinas pertenecientes a las Ciencias Sociales, con particular énfasis en los modelos CB-SEM (modelos de covarianza).

Con la finalidad de organizar la presentación de los fundamentos, reflexiones y conclusiones esbozadas en este artículo, se inicia con un breve recorrido en torno a los orígenes de los MEE. Posteriormente, se detalla la estructura y etapas para su construcción, seguido de un análisis sobre las contribuciones ofrecidas de estos modelos a la investigación social. Finalmente, se concluye resaltando su utilidad para la indagación empírica en Ciencias Sociales.

Metodología

La presente investigación se desarrolló a partir de un enfoque cualitativo. Se llevó a cabo un análisis descriptivo basado en la revisión, evaluación y síntesis crítica de fuentes bibliográficas vinculadas con los MEE. Esta revisión de antecedentes posibilitó analizar el potencial de estos modelos para mejorar la comprensión de fenómenos socioculturales en disciplinas pertenecientes a las Ciencias Sociales. La ejecución de este estudio se llevó a cabo en tres etapas investigativas. En primer lugar, se volcó la atención a los orígenes de los MEE, con particular interés en su utilización dentro de investigaciones sociales. En la segunda etapa, se consultó una diversidad de fuentes bibliográficas que versaban en torno a la estructura y pasos a seguir para la construcción de un MEE, contrastándolas con el objetivo de plantear una secuencia típica en la elaboración de estos modelos. En la tercera y última etapa, se generaron reflexiones que resaltaron los beneficios, limitaciones y aportes de los MEE dentro de la investigación social.

La columna vertebral de la metodología consistió en la búsqueda y selección de material bibliográfico en línea, vinculado con los temas centrales del estudio escritos tanto en español como inglés. Debido a que en esta investigación se contempla una breve revisión histórica sobre la creación y utilización de los MEE, las fuentes bibliográficas datan de 1921 (uno de los primeros textos en los que se plantea este tipo de modelos), hasta el año 2024; en este sentido, la fecha de publicación no constituyó un criterio de exclusión. Para localizar las fuentes de interés para el estudio, se utilizaron motores de búsqueda en línea (como por ejemplo repositorios de artículos científicos y bibliotecas digitales) en los cuales se introdujeron palabras clave vinculadas con los temas centrales de la investigación, permitiendo así la identificación, clasificación y análisis de los textos que sirvieron como materia prima para la elaboración de este artículo.

El análisis de estas fuentes consistió en la síntesis de las ideas y argumentos centrales de cada uno de los textos consultados, los cuales fueron agrupados en categorías preestablecidas de acuerdo con su pertinencia para cada uno de los apartados de interés en el estudio. Una vez categorizada esta información, se procedió a la construcción del cuerpo de cada uno de estos apartados, entrelazando y contrastando las ideas clave de las autorías consultadas.

Origen y características de los modelos de ecuaciones estructurales

El término genérico “modelos de ecuaciones estructurales” (MEE) [SEM por sus siglas en inglés] ha sido utilizado en los procesos de investigación empírica desde hace muchos años, datándose desde la década de 1920 (Tarka, 2018). Hace referencia al procedimiento analítico que «designa aproximaciones varias al análisis de la

causalidad entre una serie de variables independientes y dependientes, con la particularidad de incluir distintas interrelaciones entre ellas y diferenciar entre variables observadas y latentes» (Cea D'Ancona 2002, 515).

Dicho concepto se introdujo hace más de 80 años por el genetista estadounidense Sewall Wright (1921), quien:

trabajando sobre patrones de covariación entre varias características de cobayas, desarrolló una forma de romper las correlaciones observadas en un sistema de ecuaciones que matemáticamente describían su hipótesis respecto a las relaciones causales. Las relaciones entre variables fueron representadas [posteriormente por dicho investigador] en un pathdiagram (diagrama de ruta, senderos o trayectoria), por lo que su método llegó conocerse como pathanalysis, [constituyéndose en uno de los primeros métodos de investigación que utilizaban como recurso un modelo gráfico] (García Veiga 2011, 17).

Justamente, Maruyama (1997) comenta que en un inicio los MEE fueron llamados *pathanalysis*, pues tomaban como referencia el nombre utilizado por Wright (1921) en sus investigaciones. No obstante, con la inclusión del término en las Ciencias Sociales y el desarrollo de las nuevas tecnologías de la información y la comunicación, este paso a denominarse “modelado causal”. Este término luego fue reemplazado por uno menos controversial y más descriptivo que no produjera tanto debate y discusión en la comunidad científica respecto al tipo de causalidad referida; este fue “ecuación estructural”. Dicho concepto es muy apropiado, según Ruiz, Pardo y San Martín (2010) pues «el nombre que reciben los modelos de ecuaciones estructurales es debido a que es necesario utilizar un conjunto de ecuaciones para representar las relaciones propuestas por la teoría» (35).

Por otra parte, los MEE son como una familia de modelos estadísticos que buscan explicar las relaciones existentes entre múltiples variables, por medio de la estructura de las interrelaciones expresadas mediante un conjunto de ecuaciones similares a un grupo de regresiones (Hair et al. 2006; Tarka 2018). Esta característica ha hecho argumentar a distintos científicos sociales, que uno de los principales alcances de estos modelos es estimar la fuerza y direccionalidad de todas las relaciones hipotetizadas entre las variables previstas en un modelo teórico (Maruyama 1997). En palabras de Ruiz, Pardo y San Martín:

la gran ventaja de este tipo de modelos es que permiten proponer el tipo y dirección de las relaciones que se espera encontrar entre las diversas variables contenidas en él, para pasar posteriormente a estimar los parámetros[1] que vienen especificados por las relaciones propuestas a nivel teórico. Por este motivo se denominan también modelos confirmatorios, ya que el interés fundamental es “confirmar” mediante el análisis de la muestra, las relaciones propuestas a partir de la teoría explicativa que se haya decidido utilizar como referencia (2010, 34).

Sobre esta misma línea discursiva, Gutiérrez Doña (2008, 11) opina que los MEE ofrecen la posibilidad de realizar análisis complejos al unir un análisis factorial confirmatorio con una regresión lineal, permitiendo de este modo el estudio de estructuras latentes. Es decir, «los constructos hipotéticos se definen como variables latentes, que son operacionalizados por varios indicadores, de manera que es posible estimar el error de medición de los indicadores individualmente y analizar las regresiones *libres de error* entre los constructos hipotéticos». Ciertamente por esta cualidad, Ruiz, Pardo y San Martín (2010) defienden que los MEE son un grupo de modelos estadísticos multivariantes que permiten estimar el efecto y las relaciones entre distintas variables de una manera relativamente sencilla y eficaz.

Para Kesteren (2021), los MEE constituyen una herramienta útil para el estudio de variables latentes. Casas Guillén comparte esta idea, argumentando que «estos modelos no prueban la causalidad, pero ayudan al [personal] investigador en la toma de decisiones, rechazando las hipótesis (...) cuando se contradicen con los datos, esto es, con la estructura de covarianzas o correlaciones subyacente entre las variables» (2002, 1).

Específicamente, sobre el término “causal”:

en las ciencias sociales hace referencia al conjunto de estrategias y técnicas de elaboración de modelos (...) que explican los fenómenos con objeto de contrastarlos empíricamente. Sus orígenes se encuentran en el path-analysis, literalmente traducido como análisis de senderos [de ruta o trayectorias], cuyo objeto es el estudio de los efectos de unas variables

consideradas como causas sobre otras tomadas como efectos (...) El análisis path es una técnica similar a la regresión pero con poder explicativo, que estudia los efectos directos e indirectos en el conjunto de las variables observables, asumiendo la existencia de relaciones lineales entre ellas, [así como] la incorrelación de los errores de regresión y la ausencia de errores de medición de las variables (Casas Guillén 2002, 2).

Puede afirmarse que el término “causalidad” es utilizado en estos modelos como una forma de explicar una relación empírica sin necesidad de poseer un diseño experimental, es decir, alude al contraste que se logra establecer entre la teoría y las técnicas utilizadas sin presentarse una manipulación directa de las variables. En términos más sencillos, hipotetiza una relación causa-efecto. Ruiz, Pardo y San Martín explican esta característica, justificando que «la especificación teórica del modelo permite proponer estructuras causales entre las variables, de manera que unas variables causen un efecto sobre otras variables que, a su vez, pueden trasladar estos efectos a otras variables, creando [así diversas] concatenaciones de variables» (2010, 34).

Una vez conocidos a grandes rasgos los orígenes de esta metodología, es indispensable especificar el tipo de variables que se pueden distinguir dentro de estos modelos. Ruiz, Pardo y San Martín (2010, 36) realizan una muy buena descripción de estas, la cual vale la pena transcribir:

- Variable observada o indicador: mide a los sujetos [de investigación]. Por ejemplo, las preguntas de un cuestionario referidas a conductas realizadas.
- Variable latente: característica que se desearía medir, pero que no se puede observar y que está libre de error de medición. Por ejemplo, una dimensión de un cuestionario o un factor en un análisis factorial exploratorio.
- Variable error: representa tanto los errores asociados a la medición de una variable como el conjunto de variables que no han sido contempladas en el modelo y que pueden afectar a la medición de una variable observada. Se considera que son variables de tipo latente por no ser observables directamente. El error asociado a la variable dependiente representa el error de predicción.
- Variable de agrupación: variable categórica que representa la pertenencia a las distintas subpoblaciones que se desea comparar. Cada código representa una subpoblación.
- Variable exógena [indicadora]: afecta a otra variable [a la vez que] no recibe efecto de ninguna variable. Las variables independientes de un modelo de regresión son exógenas, [en un diagrama de ruta estas variables son las que no reciben ninguna flecha].
- Variable endógena [criterio]: recibe efecto de otra variable. La variable dependiente de un modelo de regresión es endógena. Toda variable endógena debe ir acompañada de un error, [en un diagrama de trayectoria estas variables son las que reciben las flechas].

Ahondando un poco en la descripción anterior, García Veiga (2011, 18) expone que «cualquier efecto estructural se representa como una flecha recta y unidireccional, cuyo origen es la variable indicadora y cuyo final, donde se encuentra la punta de la flecha, es la variable dependiente».

Asimismo, para efectos de la diagramación de dichas variables y el establecimiento de las posibles relaciones entre ellas, es importante considerar la siguiente simbología presente en la Tabla 1, con el objetivo interpretar adecuadamente los gráficos resultantes de un MEE (Cea D’Ancona 2002; García Veiga 2011):

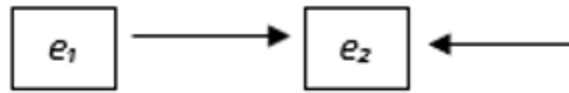
Tabla 1
Simbología utilizada en el diagrama de un Modelo de Ecuaciones Estructurales

Criterio	Símbolo
Las variables observadas (indicadores) se representan en un cuadrado.	
Las variables latentes (constructos) figuran en un círculo o elipse.	
Las flechas rectas indican relaciones causales entre las variables conectadas (asociación mediante regresión).	
Las flechas curvas y de doble dirección denotan correlaciones (covariación).	
La varianza de una variable puede representarse utilizando una flecha de doble cabecera desde una variable a sí misma.	
Las líneas discontinuas indican trayectorias o senderos no significativos.	
A cada flecha se le asocia un coeficiente.	0.5
La inexistencia de flecha significa que dichas variables no están directamente relacionadas (independencia condicional).	

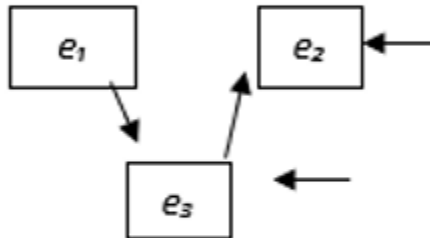
elaboración propia con base en Cea D'Ancona (2002) y García Veiga (2011).

Por otra parte, como complemento a la Tabla 1, la Figura 1 ilustra algunas de las relaciones que pueden presentarse en un diagrama de senderos:

a) Relación directa: e_1 causa e_2



b) Relación causal indirecta: e_1 causa e_2 a través del efecto de e_3



c) Relación espuria o no causa entre e_1 y e_2 : e_3 provoca efecto sobre e_1 y e_2

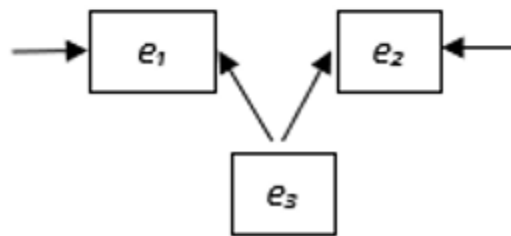


Figura 1

Relaciones entre variables de un Modelo de Ecuaciones Estructurales
Elaborado con base en Casas Guillén (2002, 2).

Estructura de un MEE

Los MEE están conformados por dos partes fundamentales o submodelos: un modelo de medida y un modelo de relaciones estructurales. De acuerdo con Gutiérrez Doña (2008) y Tarka (2018), en la estructuración de estos modelos se procede de lo general a lo particular, es decir, primero se hace un planteamiento teórico basado en un ejemplo gráfico que exponga el modelo de medición para especificar posteriormente el modelo estructural. En dicho contexto, el modelo de regresión es especificado como el modelo estructural, mientras que el análisis de factores es especificado como el modelo de medición. Para Cea D'Ancona (2002), las características principales de dichos submodelos son las siguientes:

Modelo de medida o de medición (análisis de factores)

Incluye las relaciones de las variables latentes con las variables empíricas que los representan. Por medición se entenderá el proceso por el que un concepto se ve representado en una variable latente y esta se vincula a variables observadas. Su ecuación se expresa de la siguiente forma en el caso de variables endógenas:

$$Y = \Lambda \eta + \epsilon$$

- Y = variables observadas dependientes.
- Λ = matriz de coeficientes factoriales de Y . Muestra la relación de η a Y .

- η = es un vector de variables no observadas o latentes.
- ε = error de medición de los indicadores endógenos.

Para variables exógenas:

$$X = \Lambda \cdot \zeta(x_i) + \delta$$

- X = sería un vector de variables observadas independientes o exógenas.
- Λ (lambda) = es una matriz de pesos factoriales que relaciona las variables manifiestas con las latentes.
- $\zeta(.i)$ = es un vector de variables exógenas latentes.
- δ (delta) = es un vector de errores de medición.

En este sentido, vale agregar que Cea D'Ancona destaca que «el modelo de medición incluye las relaciones de las variables latentes (o constructos) con sus indicadores (o variables empíricas)» (2002, 529); y que la finalidad última de este modelo es poder comprobar la adecuación de los indicadores elegidos en la medición de los términos conceptuales de interés.

Modelo estructural (análisis de regresión)

Es el modelo que describe relaciones causales entre variables latentes. Su ecuación se expresa de la siguiente manera:

$$\eta = \Gamma \zeta(x_i) + B\eta + \zeta$$

Donde cada elemento de la ecuación representa:

- η (*eta*) = vector de variables endógenas latentes.
- Γ (gamma) es una matriz de coeficientes [γ] que relaciona las variables latentes exógenas (ξ) con las endógenas (η).
- ξ (x_i) = vector de variables exógenas latentes (indicadoras).
- β = matriz con relaciones de las variables latentes endógenas entre sí.
- ζ (zeta) = errores o perturbación

Dada la complejidad que puede representar el entendimiento de los signos y la nomenclatura empleada en las ecuaciones de estos modelos; la Tabla 2 esquematiza los principales símbolos utilizados en esta labor, con el objetivo de brindar una mejor comprensión de estos:

Tabla 2

Significado de la simbología en las ecuaciones de un Modelo de Ecuaciones Estructurales

Símbolo	Significado
ξ (xi)	Variables latentes exógenas (o independientes) representadas por las X.
η (eta)	Variables latentes endógenas (o dependientes) representadas con las Y.
γ (coeficientes gama)	Coefficientes que miden el efecto <i>directo</i> de las ξ (variables latentes exógenas) sobre las η (variables latentes endógenas).
β (coeficientes beta)	Son los efectos de las η (variables latentes endógenas) sobre las η (variables latentes endógenas).
ϕ (Coeficientes phi)	Son las correlaciones entre las variables exógenas latentes.
λ (coeficientes lambda)	Son coeficientes de regresión que relacionan cada variable observable con su parte no observable o latente.
ε (coeficientes epsilon)	Son coeficientes sobre errores de medición en las variables endógenas observables; es decir, son los errores de medición de las Y (variables dependientes).
δ (coeficientes delta)	Son coeficientes sobre errores de medición en variables exógenas observadas; es decir, son los errores de medición de las X (variables independientes).
ζ (zeta)	Es el término de error de medición para cada ecuación que relaciona un grupo de variables endógenas y exógenas explicativas, con una variable endógena criterio. En otras palabras, son los errores de medición de cada ecuación lineal estructural.

Elaboración propia con base en Gutiérrez Doña (2008, 14).

Etapas en la construcción de un MEE

De acuerdo con Casas Guillén los MEE analizan las relaciones causales y no causales entre distintas variables por medio de indicadores de medida de los constructos (o factores) identificados teóricamente, en donde la persona investigadora diseña un modelo que intenta representar de forma sencilla las relaciones existentes entre las variables latentes y observadas. «La hipótesis de partida de todos estos modelos es que reproducen exactamente la estructura de varianzas y covarianzas de las variables objeto de estudio» (2002, 3), siguiendo una metodología que pasa por diferentes etapas: especificación, identificación, estimación de parámetros, evaluación del ajuste, reespecificación del modelo e interpretación de resultados (Tabla 3).

Tabla 3

Etapas para elaborar un Modelo de Ecuaciones Estructurales

Etapas
1. Especificación del modelo teórico a utilizar.
2. Identificación de parámetros.
3. Medición de variables observadas (fase de estimación).
4. Valoración de ajuste del modelo (fase de ajuste).
5. Evaluación del modelo (fase de evaluación e interpretación).

Elaboración propia con base en Casas Guillén (2002) y Cea D'Ancona (2002).

A continuación, se presenta una breve descripción de los pasos que deben seguirse para elaborar un MEE (Casas Guillén 2002; Ruiz, Pardo y San Martín 2010, Stoffels et al. 2023). Si bien, dicha descripción busca ser detallada, es importante aclarar que no pretende ser exhaustiva.

Estimación del modelo y formulación de la teoría que lo sustenta

Alude a la identificación de variables de acuerdo con una teoría y tema en particular. De acuerdo con Ruiz, Pardo y San Martín:

dicha teoría debe estar formulada de manera que se pueda poner a prueba con datos reales. En concreto, debe contener las variables que se consideran importantes y que deben medirse a los sujetos. El modelo teórico debe especificar las relaciones que se espera encontrar entre las variables (correlaciones, efectos directos, efectos indirectos, bucles). Si una variable no es directamente observable, deben mencionarse los indicadores que permitan medirla. Lo normal es formular el modelo en formato gráfico; a partir de ahí es fácil identificar las ecuaciones y los parámetros (2010, 39).

En palabras de Casas Guillén:

En esta fase el investigador aplica sus conocimientos teóricos del fenómeno estudiado al planteamiento de las ecuaciones matemáticas relativas a los efectos causales de las variables latentes y a las expresiones que las relacionan con los indicadores o variables observables. Del mismo modo, la claridad del modelo se encontrará determinada por el grado de conocimiento teórico que posea el investigador sobre el tema de estudio, si la información es poco exhaustiva o detallada, la asignación de los parámetros será confusa a priori, por lo que el investigador debe realizar diversos análisis exploratorios de los datos hasta configurar el modelo, y efectuar [posteriormente] el análisis confirmatorio del mismo (2002, 3).

Identificación de parámetros

Una vez formulado el modelo teórico, cada parámetro debe ser identificado y definido correctamente (Stoffels et al. 2023). Igualmente, debe ser derivable de la información contenida en la matriz de varianzas-covarianzas. Este es un proceso complejo que debe realizarse cuidadosamente para no incurrir en errores que afecten posteriormente el análisis del modelo. En este sentido, una de las reglas o estrategias más utilizadas para saber si un modelo está identificado «es la regla de los grados de libertad, obtenidos como la diferencia entre el número de varianzas y covarianzas (ecuaciones) y el número de parámetros a estimar» (Casas Guillén 2002, 5). Byrne afirma que:

Los modelos estructurales pueden ser identificables (just-identified model), sobre identificados (over-identified model) o sub identificados (under identified model). Un modelo identificable es aquel en el cual existe una correspondencia de uno-a-uno entre los datos y los parámetros estructurales, [en donde los grados de libertad son positivos - $df = 0$ -], (...) un modelo sobre identificado es aquel en el cual el número de parámetros a ser estimados es menor que el número de varianzas y covarianzas de las variables observadas, (...) [y] un modelo sub identificado es aquel en el cual el número de parámetros a ser estimados es mayor que el número de varianzas y covarianzas, [lo cual equivaldría a decir que la cantidad de parámetros estimados libremente en el modelo es mayor que el número de datos conocidos] (1998, 9).

Fase de estimación

Una vez seleccionadas las variables que formarán parte del modelo, se decide cómo se medirán las variables observables (Thakkar 2020). La importancia de estas mediciones radica en que gracias a ellas se obtienen las varianzas y las covarianzas sobre las cuales se basa la estimación de los parámetros. García Veiga opina que los MEE «se basan en las correlaciones existentes entre las variables medidas en una muestra de sujetos de manera transversal. Por tanto, para poder realizar las estimaciones, basta con medir a un conjunto de sujetos en un momento dado (...). [Del mismo modo], hay que tener en cuenta que las variables deben permitir el cálculo de

las correlaciones [a realizar], por ello deben ser variables cuantitativas [preferiblemente] continuas» (2011, 21).

Por otra parte, sobre la estimación del modelo, vale rescatar los siguientes métodos que además tienden a ser los más utilizados en esta etapa:

- Mínimos cuadrados no ponderados (ULS por sus siglas en inglés): busca minimizar una mitad de la suma de los cuadrados de cada elemento de la matriz residual. Así, según Cea D'Ancona (2002), se da un peso empíricamente determinado a cada residuo cuadrado antes de producirse la suma. Suele ser un método adecuado cuando otros métodos han fallado.

- Mínimos cuadrados generalizados (GLS por sus siglas en inglés): es una variante del método de mínimos cuadrados no ponderados. Al igual que el anterior, busca la minimización de los residuos. En este proceso de estimación, se otorgan pesos desiguales a los diversos residuos, lo cual conlleva la ponderación distinta de las observaciones para corregir las varianzas-covarianzas desiguales de los términos de error. Funciona de manera más óptima cuando hay normalidad en los datos (Cea D'Ancona 2002).

- Máxima verosimilitud (ML por sus siglas en inglés): es el método de estimación de parámetros más utilizado (Cea D'Ancona 2002; Tabachnick y Fidell 2007). Se trata de un método que maximiza la probabilidad de los parámetros a partir de los datos empíricos. Mediante este procedimiento se buscan los valores de los parámetros que con mayor probabilidad han ocasionado las covarianzas o correlaciones observadas, de manera que exige el cumplimiento del supuesto de normalidad, homocedasticidad[2] e independencia de los residuos. De acuerdo con Cea D'Ancona (2002), este método junto al GLS pueden obtener estimaciones carentes de sesgo a partir de las 100 unidades de análisis.

Fase de ajuste

Una vez estimados los parámetros del modelo se procede inicialmente a valorar su ajuste (Thakkar 2020). «La etapa de diagnóstico de la bondad del ajuste se refiere a la exactitud de los supuestos del modelo especificado para determinar si el modelo es correcto y sirve como aproximación al fenómeno real, precisando así su poder de predicción» (Casas Guillén 2002, 2). En otras palabras, es la comprobación del ajuste global de los datos observados, lo cual consiste en contrastar las varianzas y covarianzas predichas con las inicialmente observadas (Stoffels et al. 2023).

Según expone Cea D'Ancona (2002), dicho procedimiento consiste en restar a cada celda de la matriz observada la celda correspondiente de la matriz esperada (o residual), si el resultado es igual a cero [0] significará que existe un ajuste perfecto, de forma que entre más ceros posea la matriz residual mejor será el ajuste.

Para estimar y valorar la bondad del modelo, en esta etapa se han identificado tres tipos de índices:

- Índice de ajuste global: es utilizado para comprobar el ajuste global del modelo considerando las diferencias existentes entre la matriz de varianzas-covarianzas observadas y la predicha (matriz residual). El indicador más utilizado es la razón de verosimilitud χ^2 (Cea D'Ancona 2002; Tabachnick y Fidell 2007).

- Índice de ajuste incremental: compara el modelo trazado con el modelo nulo, es decir, compara el modelo propuesto con el modelo más sencillo posible, en el cual no se especifica ninguna relación entre las variables y todos los senderos son fijados a cero. Los indicadores que se incluyen en el ajuste incremental son varios, no obstante, todos comparten la característica de oscilar en un rango de 0 a 1 [entre más cercano este a uno se dice que hay mejor ajuste] (Cea D'Ancona 2002; Tabachnick y Fidell 2007).

- Índice de ajuste de parsimonia: se busca obtener un modelo sencillo, en el cual haya pocos parámetros y al mismo tiempo muchos grados de libertad. En este sentido, entre más parámetros libres haya es más probable que el modelo tenga un buen ajuste (Cea D'Ancona 2002; Tabachnick y Fidell 2007).

En la siguiente Tabla 4 se pueden apreciar los indicadores más utilizados de acuerdo con cada índice de bondad de ajuste:

Tabla 4
Indicadores según índices de bondad de ajuste y cortes aceptables

Ajuste global o absoluto	Valor de ajuste esperado (puntos de cortes aceptables)
Índice de razón de verosimilitud χ^2	Estadísticamente no significativo ($p > 0,05$), que coincida con un valor pequeño ($\chi^2 = 0,0$ ajuste perfecto)
Bondad de ajuste (GFI)	Grande $\geq 0,90$
Raíz cuadrada de la media de los residuos (RMSR)	Pequeño (0,0 ajuste perfecto)
Índice de centralidad (CI)	Grande $\geq 0,90$
Ajuste incremental	Valor de ajuste esperado
Tucker-Lewis (TLI)	
Ajuste relativo (RFI)	
Ajuste Normado (NFI)	Elevados $\geq 0,90$
Ajuste incremental (IFI)	
Ajuste comparativo (CFI)	
Índice de no centralidad relativa (RNI)	
Ajuste de parsimonia	Valor de ajuste esperado
Ajuste parsimonioso (PFI)	No posee valor referencial, no obstante, se sugieren valores de 0,06 y 0,90 como indicador de diferencias sustanciales entre los modelos comparados
χ^2 normado	$< 2,0$
Bondad de ajuste ajustado (AGFI)	$\geq 0,90$
De bondad de ajuste de parsimonia (PGFI)	$\geq 0,90$
Error de la raíz cuadrada media de aproximación (RMSEA)	$\leq 0,05$
Información de Akaike (AIC)	Pequeño
Estadístico N crítico (CN)	≥ 200

Elaboración propia con base en Cea D'Ancona (2002).

Fase de evaluación e interpretación

Una vez realizados los procedimientos de ajuste necesarios, se procede a hacer una valoración técnica de los valores estimados para los parámetros con el objetivo de interpretar el modelo en todas sus partes y conocer su capacidad predictiva (Cea D'Ancona 2002). Este análisis permitirá la identificación de conclusiones y recomendaciones importantes vinculadas al tema de investigación elegido, o en su defecto, —si el modelo no fuera válido— contribuiría a sentar las bases para su rediseño (Thakkar 2020).

De igual manera, una vez identificadas las distintas etapas en la elaboración de un MEE, se vuelve significativo señalar los aportes de Cea D'Ancona (2002) respecto a los supuestos básicos que deben considerarse cuando se construye un modelo de este tipo, los cuales destacan que todas las variables deben encontrarse estandarizadas (según la “métrica que le demos al modelo”), es decir, medidas como desviaciones de su media; los errores de medición deben tener un valor esperado igual a cero y no estar correlacionados.

Además, los errores de medición deben cumplir los supuestos de homocedasticidad (varianza constante) y las correlaciones teóricas deben ser viables tanto entre indicadores exógenos (X) como endógenos (Y).

Con respecto a los tipos de análisis que se utilizan en los MEE, es importante resaltar los supuestos básicos del análisis de regresión lineal, pues son estas premisas las que guían, en gran parte, las etapas de valoración establecidas en dichos modelos, por ello, Cea D'Ancona (2002) recomienda tener en consideración los siguientes supuestos:

1. Tamaño muestral: la importancia de estimar de manera correcta el tamaño de la muestra a utilizar es un dato relevante, pues «la fiabilidad de la mayoría de los índices de ajuste aumenta, cuando se incrementa el tamaño de la muestra» (Cea D'Ancona 2001, 519) de ahí que se recomienda que debe de haber 10 veces más casos que variables observadas. Aunque este requisito actualmente se ha flexibilizado bastante.

2. Normalidad multivariable: «su existencia es imprescindible para garantizar que los valores críticos que determinan la significatividad de cada coeficiente del modelo, y la del modelo en su conjunto, sean correctos» (Cea D'Ancona 2002, 520).

3. Las variables observadas y latentes han de ser continuas: este supuesto es de gran relevancia pues en ciencias sociales es muy común encontrar variables nominales y ordinales, estas últimas por lo general se trabajan como continuas, lo que hace que se incumpla este supuesto y el supuesto anterior [aunque en la actualidad existe la tecnología necesaria para soportar análisis con datos categóricos u ordinales].

4. Ausencia de colinealidad entre las variables: la colinealidad hace referencia a la correlación que hay entre las variables independientes, por lo tanto, es esperado que esta sea la mínima posible, de no ser así se incurre en la dificultad de aumentar la magnitud del error típico.

Todos los supuestos anteriores son relevantes no solo para el uso de programas estadísticos que analicen los datos recolectados, sino también para que él o la investigadora puedan discernir la información más relevante que deben reportar.

En síntesis, la utilidad de los MEE para la investigación empírica en Ciencias Sociales reside en el carácter global que muestran respecto a la relación existente entre las variables de estudio o interés investigativo (Tarka 2018), Esto facilita el agrupamiento de grandes cantidades de información en algunos pocos datos, los cuales pueden representarse de una manera visualmente sencilla y agradable. Asimismo, al carecer del error de medición son ideales en el estudio de constructos no medibles directamente, como lo pueden ser las actitudes y la intención conductual (Cea D'Ancona 2002; Gutiérrez Doña 2008; Ruiz, Pardo y San Martín 2010).

En esa dirección, Ruiz, Pardo y San Martín (2010) consideran que las fortalezas principales de estos modelos son:

- permiten su representación gráfica,
- facilitan la posibilidad de hipotetizar efectos causales entre las variables,
- admiten la concatenación de efectos entre variables y
- contribuyen a identificar relaciones recíprocas entre variables.

En lo que respecta al estudio de la causalidad o predictibilidad, los MEE no sólo facilitan los análisis estadísticos, sino también contribuyen a la toma de decisiones por parte del personal investigador, al lograr identificar modelos e indicadores que explican diversos fenómenos de la realidad humana (Shaheen et al. 2017). Finalmente, recapitulando las principales características de los MEE, la Figura 2 presenta a continuación un mapa conceptual que integra los aspectos fundamentales desarrollados con anterioridad.

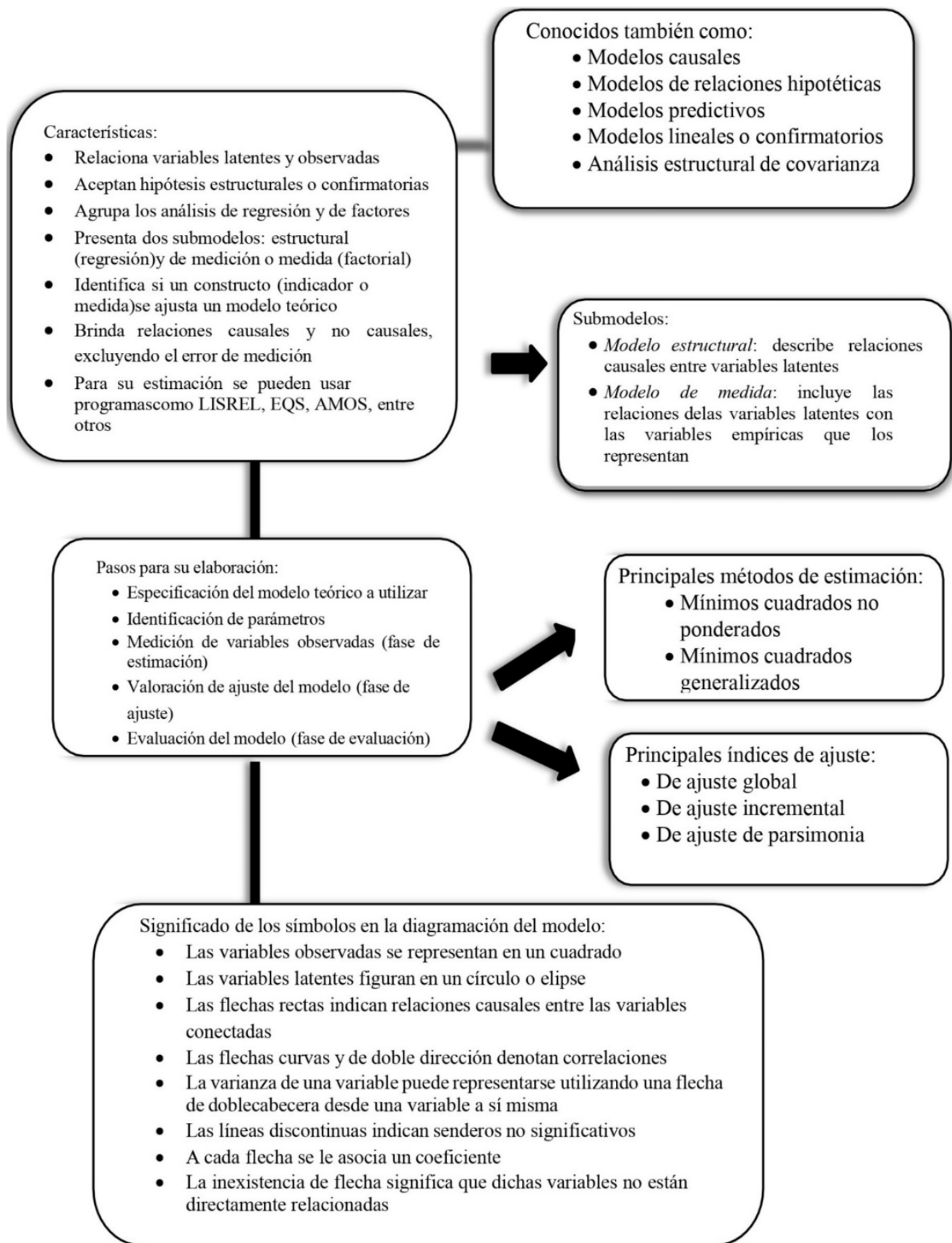


Figura 2

Mapa conceptual sobre modelos de ecuaciones estructurales

Elaboración propia con base en Cea D'Ancona (2002), Gutiérrez Doña (2008) y Ruiz, Pardo y San Martín (2010).

Aportes de los MEE en la investigación social: alcances y limitaciones

Como parte de la sociedad, resulta fundamental poder identificar y comprender diversos procesos, estructuras y dimensiones, especialmente, cuando existen variables latentes que no cuentan con una denominación precisa y estas, al igual que otras variables, ayudan a explicar y describir la realidad social circundante. Por tanto, es necesario contar con sistemas y métodos que puedan darles sentido.

Alaminos et al (2015) exponen que los MEE son como una narración para intentar explicar un fenómeno o proceso social; donde se seleccionan un conjunto de personajes, llamados variables que son considerados relevantes para una historia. Estos interactúan entre sí, y sus acciones están conectadas y explicadas dentro de la trama. Algunas acciones de personajes se explican por las de otros, estableciendo relaciones causales entre ellos. Asimismo, puede haber personajes cuyas acciones están explicadas con detalle dentro de la historia, aunque su impacto en los demás personajes puede no ser relevante en ese momento específico de la narrativa. Esta es una forma creativa de comprender los aportes de los modelos de ecuaciones estructurales para la investigación en Ciencias Sociales.

Dentro de la investigación social, la versatilidad de estos modelos en la exploración de conexiones complejas entre variables, tanto observables como no observables, los convierte en una herramienta esencial para la comprensión de fenómenos en entornos no experimentales (Ghaleb y Murat 2024). A diferencia de enfoques tradicionales, como la regresión múltiple, los MEE son capaces de abordar diversas relaciones de dependencia al mismo tiempo, haciéndose útiles para investigar situaciones donde múltiples variables interactúan entre sí.

Sin embargo, a pesar de estas fortalezas, los MEE también tienen ciertas áreas de mejora que deben ser consideradas a la hora de ser aplicados. En las siguientes líneas se explorarán con detenimiento estas limitaciones, para ello se brindarán ejemplos de aplicación y señalamientos relevantes de distintas personas autoras, con el fin de potenciar aún más el impacto de los MEE y sus aportes en el ámbito de la ciencia.

Los MEE son una herramienta importante para la investigación social, utilizada en diversas Ciencias Sociales como la Psicología, la Educación y la Sociología. «Una de las ventajas de los SEM es que permiten proponer el tipo y dirección de las relaciones que se esperan encontrar entre las variables continuas en estudio para realizar la verificación de relaciones propuestas a nivel teórico» (González Martínez 2018, 3). Estos modelos permiten comprender relaciones complejas entre variables y considerar los errores de medición en los datos; lo cual posibilita analizar simultáneamente varias relaciones de dependencia, a diferencia de otros enfoques. Algunas de estas otras técnicas incluyen la regresión múltiple, análisis de factores, análisis multivariantes y el análisis discriminante, las cuales solamente pueden examinar una relación a la vez. Esto puede representar un desafío a la hora de investigar fenómenos complejos a nivel social que involucran múltiples variables interconectadas (Cupani 2012; Ghaleb y Murat 2024).

De acuerdo con González Martínez (2018), cuando las variables están interconectadas, los MEE muestran cómo una puede tener influencia en otras, y como está varía en fuerza y dirección. Esto puede ser especialmente útil cuando los efectos no son uniformes y las interacciones entre variables son necesarias para entender el panorama completo. Además, los MEE permiten evaluar qué tan bien se ajusta el modelo teórico a los datos observados utilizando diferentes índices de ajuste, lo cual es fundamental para determinar si el modelo teórico refleja la realidad con precisión, lo que contribuye a la construcción de modelos más sólidos (González Martínez 2018; Shaheen et al. 2017). En este sentido, la construcción de modelos sólidos es fundamental para la progresividad de la ciencia y el conocimiento científico en el estudio de variables humanas o sociales.

Los MEE se destacan por estimar y evaluar relaciones entre variables latentes, es decir, aquellas que representan constructos teóricos que no son observables directamente, como la inteligencia o la satisfacción (Kesteren 2021). A diferencia de otras técnicas que utilizan una única medida para representar constructos y

no modelan el error de medición, los MEE permiten el uso de múltiples medidas para representar un constructo, controlando así el error de medición específico de cada variable. Todo esto ayuda a evaluar la validez de los constructos medidos (Cupani 2012).

A la hora de interpretar resultados dentro de los MEE, es necesaria una evaluación de diversas pruebas con el fin de determinar si la estructura teórica propuesta se ajusta adecuadamente a los datos empíricos, lo cual implica comparar los valores de los parámetros estimados con la matriz observada de covarianzas. A pesar de su potencial para evaluar modelos teóricos, los MEE no establecen causalidades en sí mismas, sino que ayudan a seleccionar hipótesis causales basadas en evidencia empírica, permitiendo que se rechacen modelos causales que contradigan los datos recopilados. Su valor está en la especificación previa de relaciones complejas entre variables y en la evaluación de cuántas de estas relaciones concuerdan con los datos recopilados empíricamente (Cupani 2012).

En áreas como la Psicología Educativa (como en otras Ciencias Sociales), por ejemplo, los MEE tienen dos principales propósitos: por un lado, desarrollar y validar instrumentos de medición relacionados con constructos de interés, y por otro, estudiar y probar relaciones entre variables latentes en modelos estructurales. Esta capacidad de comprender y evaluar modelos teóricos a través de MEE ha permitido tener un enfoque más profundo y sofisticado en la investigación, lo que resulta en una comprensión más precisa de los fenómenos estudiados (Rigo y Donolo 2018).

Un ejemplo de aplicación corresponde al estudio de Vargas Hernández y Montero Rojas (2016) en el cual utilizaron modelos de ecuaciones estructurales para examinar los posibles factores causales que influían en el rendimiento académico en cursos de matemáticas en la Universidad Nacional de Ingeniería de Nicaragua. Su variable dependiente fue el rendimiento académico en un curso de Matemáticas, el cual midieron en una escala de 0 a 100. Utilizaron distintas variables independientes para analizar su relación con el rendimiento académico, incluyendo la inteligencia fluida de las personas estudiantes, sus hábitos de estudio y sus actitudes negativas hacia las matemáticas. Asimismo, se consideraron factores relacionados con las personas docentes, como la edad, su nivel académico y la participación en cursos pedagógicos. También se tomaron en cuenta variables como la estabilidad emocional, las técnicas de enseñanza, el grado de trabajo, la asistencia a clases y medidas de apoyo académico.

Los resultados mostraron que tres variables del estudiantado, la inteligencia fluida, los hábitos de estudio y las actitudes negativas hacia las matemáticas, influyen de forma significativa en su desempeño. Además, las primeras dos, no solo impactan directamente en el rendimiento académico en matemáticas, sino que también tienen incidencia en las actitudes negativas hacia la asignatura; por lo que tanto, los hábitos de estudio como la inteligencia fluida influyen en el comportamiento de las actitudes negativas hacia las matemáticas (Vargas Hernández y Montero Rojas 2016).

Por otra parte, las variables de la persona docente, tales como la edad, el nivel académico y la asistencia a cursos pedagógicos, también mostraron efectos directos de relevancia en el rendimiento académico de las y los estudiantes (Vargas Hernández y Montero Rojas 2016). Estudios como este destacan la utilidad de los MEE para analizar relaciones complejas en ámbitos naturales como lo es el educativo y su capacidad para identificar factores determinantes dentro de estos.

Manzano Patiño (2017) señala que si bien los MEE tienen distintas ventajas, existen también algunas limitaciones relevantes. Una de ellas es la necesidad de contar con “muestras grandes” para su uso, por ejemplo, $n > 200$; lo cual puede ser un desafío en investigaciones con recursos y tiempo limitados. Además, menciona que, para cada variable observada, se recomienda tener al menos diez casos en la muestra; por lo que, si se desean evaluar modelos complejos con diversas variables y parámetros, se requeriría una muestra aún más grande, lo que puede ser costoso y difícil de obtener en la práctica.

Otra limitación importante está relacionada con la formulación del modelo, debido a que plantear un modelo de MEE requiere un profundo conocimiento teórico del fenómeno estudiado. Sin una base teórica sólida, pueden desarrollarse modelos que se ajusten bien a los datos, pero que no tengan suficiente sustento

teórico; lo cual puede dar paso a conclusiones erróneas o a la construcción de modelos sin suficiente relevancia en el contexto real (Manzano Patiño 2017).

La confiabilidad de las variables latentes es otro aspecto que destacar; ya que, si se emplean variables latentes poco confiables, los resultados del MEE pueden no tener mucha confiabilidad y solidez. Finalmente, el conocimiento sobre estadística es fundamental para aplicar el MEE, incluyendo la comprensión profunda de conceptos y técnicas de modelado; sin esto, es posible que la metodología sea utilizada de manera errónea o los resultados no se interpreten correctamente, lo que a su vez representa un obstáculo para la persona investigadora (Manzano Patiño 2017).

Medrano y Muñoz Navarro (2017) por su parte, expresan que si bien los MEE se corresponden a una metodología sólida para trabajar con datos no experimentales, su eficacia y utilidad están ligadas directamente a la existencia de un modelo teórico sólido que los respalde. Destacan que es común observar un uso inadecuado de esta metodología, donde las personas investigadoras incluyen o excluyen parámetros en función de la mejora del ajuste del modelo sin una base teórica sólida. Esto genera limitaciones significativas, ya que la solidez de un modelo no depende solo del ajuste estadístico, sino también de su coherencia interna con la teoría.

Por otro lado, una limitación común que estos destacan también es la insistencia por el ajuste estadístico en detrimento de la fundamentación teórica; resaltando que el fin del MEE no es solamente ajustar bien los datos, sino contrastar teorías y conceptualizaciones, y la pérdida del razonamiento teórico en favor del ajuste estadístico puede ser perjudicial. En esta línea, Medrano y Muñoz Navarro señalan que:

ausencia de un criterio externo puede llevar a que muchos modelos SEM sean simplemente un modelo estadístico que explique la covariación entre diferentes constructos teóricos. En efecto, muchos investigadores construyen modelos “ciegos” o “mecánicos” donde se “crean” una o varias variables latentes para explicar las variaciones causales, sin un criterio externo que permita contrastar dicho modelo (2017, 237).

Resulta pertinente mencionar que los modelos MEE deben estar anclados en la realidad, vinculando las variables observables con conceptos o fenómenos de relevancia práctica (Tarka 2018). La falta de este anclaje puede convertir a estos modelos en construcciones estadísticas desvinculadas de la realidad, explicando la covariación entre constructos teóricos sin un claro vínculo con el contexto real (Barrett 2007, como se citó en Medrano y Muñoz Navarro 2017).

Dentro de las recomendaciones para abordar las limitaciones identificadas en el uso de los MEE se encuentran establecer conexiones claras entre las variables observables y fenómenos con relevancia práctica. Esto no solo fortalecerá la aplicabilidad de los resultados, sino que también proporcionará una base más sólida para la toma de decisiones. Además, se recomienda un enfoque más riguroso en la identificación y discusión de modelos que presenten deficiencias; la transparencia en la presentación de resultados, incluso cuando estos no cumplen con las expectativas, es de carácter esencial para el progreso de la producción de conocimiento científico. Esta comunidad puede verse beneficiada al abordar con apertura las limitaciones y reforzar un diálogo que conduzca a mejoras en la metodología y la comprensión de los fenómenos sociales.

Igualmente, otra recomendación podría ser explorar nuevos enfoques y desarrollar modelos más flexibles que respondan a necesidades contextualizadas y pertinentes que se adapten a los fenómenos que se buscan estudiar. También, la aplicación de los MEE en contextos culturales con mayor diversidad y adaptar esta metodología a realidades más amplias podría diversificar su utilidad y relevancia global, así como, generar nuevas preguntas y enfoques de investigación; enriqueciendo con ello la comprensión de determinados objetos de estudio bajo distintas perspectivas.

Conclusiones

Los MEE representan herramientas sumamente valiosas en la investigación social, el explorar y trabajar con conexiones complejas entre variables, los convierte en estrategias fundamentales para entender cómo estas se relacionan en distintos contextos. Más allá de confirmar supuestos teóricos, los MEE juegan un rol esencial al refinar y ajustar teorías científicas ya establecidas, robusteciendo de manera sustancial la investigación realizada por las Ciencias Sociales.

Estos modelos no solo permiten representar de forma gráfica las relaciones entre variables y brindar una narrativa visual que facilite la interpretación de los fenómenos, sino que además establecen hipótesis causales que enriquecen la comprensión de distintas dinámicas socioculturales.

La flexibilidad que otorgan los MEE al poder estudiar relaciones simultáneas de dependencia resulta especialmente valiosa al analizar procesos complejos que involucran múltiples variables interrelacionadas; tal y como sucede con frecuencia en la investigación social.

Indudablemente, los MEE favorecen la actividad investigativa y propician la generación de nuevos aprendizajes que contribuyen al entendimiento e interpretación de diversos fenómenos. Subrayan con especial énfasis la importancia de avanzar constantemente en el conocimiento de las complejidades inherentes a los procesos sociales, en aras de develar la intrincada y cautivadora relación existente entre lo observable y lo latente.

Agradecimientos

Agradecemos el significativo apoyo brindado por las personas asistentes de investigación Nathalia Vásquez Navarro y Sebastián “Teo” Ramírez Estrada.

Referencias

- Alaminos, Antonio et al. 2015. *Introducción a los modelos estructurales en investigación social*. Cuenca, Ecuador: Pydlos Ediciones.
- Bunge, Mario. 1981. *Teoría y realidad*. Barcelona, España: Editorial Ariel.
- Byrne, Barbara. 1998. *Structural equation modeling with LISREL, PRELIS and SIMPLIS*. Londres, Inglaterra: Psychology Press.
- Casas Guillén, Mercedes. 2002. «Los modelos de ecuaciones estructurales y su aplicación en el Índice Europeo de Satisfacción del Cliente». Ponencia presentada en las X Jornadas Madrid 2002. Asociación Española de Profesores Universitarios de Matemáticas para la Economía y la Empresa, Madrid, España, setiembre.
- Cea D’Ancona, María Ángeles. 2002. *Análisis multivariable. Teoría y práctica en la investigación social*. Madrid, España: Editorial Síntesis.
- Cuevas Badallo, Ana, y Oscar Martínez Ortubay. 2011. «Tema 1. Introducción a la noción de ciencia». Manuscrito inédito. Máster Oficial en Estudios Sociales de la Ciencia, Escuela de Ciencia, Centro de Altos Estudios Universitarios (OEI), Madrid, España.
- Cupani, Marcos. 2012. «Análisis de Ecuaciones Estructurales: conceptos, etapas de desarrollo y un ejemplo de aplicación». *Revista Tesis* 1: 186-199. https://www.researchgate.net/publication/274716879_Analisis_de_Ecuaciones_Estructurales_conceptos_etapas_de_desarrollo_y_un_ejemplo_de_aplicacion
- Ghaleb Marwan y Murat Yaslioglu. 2024. «Structural Equation Modeling (SEM) for Social and Behavioral Sciences Studies: Steps Sequence and Explanation». *Journal of Organizational Behavior Review* 6(1): 69-108. https://www.researchgate.net/publication/377776329_Structural_Equation_Modeling_SEM_for_Social_and_Behavioral_Sciences_Studies_Steps_Sequence_and_Explanation
- García Veiga, María Ángeles. 2011. «Análisis causal con ecuaciones estructurales de la satisfacción ciudadana con los servicios municipales». Tesis de maestría inédita, Universidad de Santiago de Compostela, Santiago de Compostela, España. http://eio.usc.es/pub/mte/descargas/proyectosfinmaster/proyecto_610.pdf
- González Martínez, Edgar Santiago. 2018. «Aplicación de un modelo de ecuaciones estructurales en un estudio observacional». Reporte práctico técnico para optar por diplomado, Universidad Veracruzana. <https://www.uv.mx/personal/jorggomez/files/2021/11/Gonzalez-Martinez-2018-APLICACION-DE-UN-MODELO-DE-ECUACIONES-ESTRUCTURALES-EN-UN-ESTUDIO-OBSERVACIONAL.pdf>
- Guevara, Minerva, y Ricardo Valdéz. 2004. «Los modelos en la enseñanza de la química: algunas de las dificultades asociadas a su enseñanza y a su aprendizaje». *Educación química* 15(3): 243-247. <https://doi.org/10.22201/fq.18708404e.2004.3.66181>
- Gutiérrez Doña, Benicio. 2008. *Modelos lineales estructurales: conceptos básicos, ecuaciones y programación con LISREL*. San José, Costa Rica: Instituto de Investigaciones Psicológicas de la Universidad de Costa Rica. <https://iip.ucr.ac.cr/sites/default/files/contenido/cuamet2.pdf>

- Hair, Joseph et al. 2006. *Multivariate Data Analysis*. 6.. ed. New Jersey, Estados Unidos: Prentice Hall.
- Kesteren, Erik-Jan van. 2021. *Structural equations with latent variables: computational solutions for modern data problems*. Utrecht, Países Bajos: Utrecht University. <https://doi.org/10.33540/130>
- Laymon, Ronald. 1989. «Applying Idealized Scientific Theories to Engineering». *Synthese* 81(3): 353-371. <https://doi.org/10.1007/BF00869321>
- Manzano Patiño, Abigail Patricia. 2017. «Introducción a los modelos de ecuaciones estructurales». *Investigación en Educación Médica* 7(25): 67-72. <https://doi.org/10.1016/j.riem.2017.11.002>
- Maruyama, Geoffrey. 1997. *Basics of structural equation modeling*. Londres, Inglaterra: Sage Publications.
- Medrano, Leonardo Adrián, y Roger Muñoz Navarro. 2017. «Aproximación Conceptual y Práctica a los Modelos de Ecuaciones Estructurales». *Revista Digital de Investigación en Docencia Universitaria* 11(1): 219-239. <http://dx.doi.org/10.19083/ridu.11.486>
- Rigo, Daiana Yamila, y Danilo Donolo. 2018. «Modelos de ecuaciones estructurales: usos en investigación psicológica y educativa». *Revista Interamericana de Psicología* 3(52): 345-357. <https://doi.org/10.30849/rip%20ijp.v52i3.388>
- Ruiz, Miguel, Antonio Pardo y Rafael San Martín. 2010. «Modelos de ecuaciones estructurales». *Papeles del Psicólogo* 31(1): 34-45. <https://www.redalyc.org/articulo.oa?id=77812441004>
- Shaheen Farhat et al. 2017. «Structural equation modeling (SEM) in social sciences & medical research: a guide for improved analysis. ». *International Journal of Academic Research in Business and Social Sciences* 7(5): 132-143. <http://dx.doi.org/10.6007/IJARBS/v7-i5/2882>
- Stoffels Malou et al. 2023. «Steps and decisions involved when conducting structural equation modeling (SEM) analysis» *Medical teacher* 45(12):1343-1345. <https://pubmed.ncbi.nlm.nih.gov/37813119/#:~:text=DOI%3A-,10.1080/0142159X.2023.2263233,-No%20abstract%20available>
- Tabachnick, Barbara, y Linda Fidell. 2007. *Using Multivariate Statistics*. 5.. ed New York, Estados Unidos: Pearson. https://hispania.ch/pluginfile.php/77114/mod_resource/content/0/Using%20Multivariate%20Statistics%20%28Tabachnick%20and%20Fidell%29.pdf
- Thakkar Jitesh. 2020. «Procedural Steps in Structural Equation Modelling». En: *Structural Equation Modelling. Studies in Systems, Decision and Control*, vol 285. Singapore: Springer. https://doi.org/10.1007/978-981-15-3793-6_3
- Tarka, Piotr. 2018. «An overview of structural equation modeling: its beginnings, historical development, usefulness and controversies in the social sciences». *Qual Quant* 52: 313-354. <https://doi.org/10.1007/s11135-017-0469-8>
- Vargas Hernández, Marcia Mercedes, y Eiliana Montero Rojas. 2016. «Factores que determinan el rendimiento académico en Matemáticas en el contexto de una universidad tecnológica: aplicación de un modelo de ecuaciones estructurales». *Universitas Psychologica* 15(4). <https://doi.org/10.11144/Javeriana.upsy15-4.fdra>
- Wright, Sewell. 1921. «Systems of mating. I. The biometric relations between parent and offspring». *Genetics* 6(2): 111-123. <https://doi.org/10.1093/genetics/6.2.111>

Notas

- 1 Un parámetro estadístico es una medida poblacional, es decir, son las características de la población que serán calculadas y probadas dentro del MEE.
- 2 La homocedasticidad refiere a que la varianza de error de los residuos es constante (Cea D'Ancona 2002)

Información adicional

Contribución de las personas autoras: El Dr. Andrés Castillo Vargas y el M.Sc. Andrés Ruiz Sánchez contribuyeron por igual a la redacción de este artículo. Ambas personas aportaron en el proceso de investigación, así como en la revisión y aprobación de la versión final del artículo. Este artículo se desprende del proyecto de investigación titulado: “Actitudes hacia la difusión y la divulgación de la ciencia en la investigación académica: un modelo estructural desde la teoría de la acción razonada”, tesis para optar por el grado de Doctor en Ciencias Sociales y de la Comunicación por la Universidad de Salamanca, España, realizado por el Dr. Andrés Castillo Vargas, autor de este artículo. Y del proyecto de investigación N° 723-C1-304: “Actitudes y prácticas de personas investigadoras hacia la divulgación de la ciencia” desarrollado por los autores en el Instituto de Investigaciones Psicológicas de la Universidad de Costa Rica.

Apoyo financiero: Universidad de Costa Rica, la Universidad de Salamanca y la Fundación Carolina

Agradecimientos: Agradecemos el significativo apoyo brindado por las personas asistentes de investigación Nathalia Vásquez Navarro y Sebastián “Teo” Ramírez Estrada.



Disponible en:

<https://www.redalyc.org/articulo.oa?id=72978397003>

Cómo citar el artículo

Número completo

Más información del artículo

Página de la revista en redalyc.org

Sistema de Información Científica Redalyc
Red de revistas científicas de Acceso Abierto diamante
Infraestructura abierta no comercial propiedad de la
academia

Andrés Castillo Vargas, Andrés Ruiz Sánchez
**Entre lo observable y lo latente: modelos de ecuaciones
estructurales e investigación social**
**Between the observable and the latent: structural
equations models and social research**

Reflexiones

vol. 104, núm. 2, 2025

Universidad de Costa Rica, Costa Rica

reflexiones.fcs@ucr.ac.cr

ISSN: 1021-1209

ISSN-E: 1659-2859

DOI: <https://doi.org/10.15517/rr.v104i2.58268>