

## DESARROLLOS RECIENTES EN ESTADÍSTICA: APORTES TEÓRICO-METODOLÓGICOS A LA INVESTIGACIÓN SOCIOLÓGICA

### *RECENT DEVELOPMENTS IN STATISTICS: THEORETICAL AND METHODOLOGICAL CONTRIBUTIONS TO SOCIOLOGICAL RESEARCH*

KARINA RDZ-NAVARRO Y RODRIGO A. ASÚN  
DEPARTAMENTO DE SOCIOLOGÍA, UNIVERSIDAD DE CHILE

Recibido: 04/04/2016  
Aceptado: 07/04/2016

**Resumen:** La presente investigación muestra y discute las ventajas empíricas de emplear herramientas estadísticas modernas (i.e., correlaciones policóricas, análisis factorial de ítems y modelos de ecuaciones estructurales) específicamente desarrollados para analizar variables latentes medidas con ítems ordinales categóricos, y compara estos procedimientos con dos tipos de análisis estadísticos clásicos desarrollados para trabajar con variables observadas (no latentes) medidas a nivel de intervalo (i.e., correlaciones de Pearson y análisis de regresión lineal). Los resultados de dos estudios Monte Carlo revelaron que los métodos estadísticos específicamente desarrollados para trabajar con variables observadas de tipo categórico y variables latentes permiten una mejor caracterización de los instrumentos de medida y los constructos teóricos, así como también una caracterización más precisa de la fuerza de la asociación entre las variables y la varianza explicada global del modelo. Se discute la relevancia de estas técnicas estadísticas para la investigación sociológica cuantitativa con la finalidad de iluminar un posible cambio en el enfoque metodológico y epistemológico de la investigación en ciencias sociales, desde el simple contraste de hipótesis nulas hacia el contraste de modelos teóricos alternativos.

**Keywords:** análisis de variable latente, análisis factorial de ítems, modelos de ecuaciones estructurales, variables ordinales.

**Abstract:** Current manuscript shows and discuss the empirical advantages of employing modern statistical tools (i.e., polychoric correlations, item factor analysis and structural equation modeling) specifically developed to analyze latent variables measured with ordinal categorical items, and compared these techniques to two classic statistical analyses developed for observed (non-latent) variables measured at an interval level of measurement (i.e., Pearson correlations and linear regression analysis). Results from two Monte Carlo studies revealed that statistical methods specifically developed to work with categorical observed variables and latent variables allowed for better characterization of measurement instruments and theoretical constructs, as well as more precise characterization of the strength of the association between variables and the overall variance explained by the model. The relevance of modern statistical techniques for sociological quantitative research is discussed in order to enlighten a possible shift in the methodological and epistemological scope in the social sciences from simple null-hypothesis testing to the testing of alternative theoretical models.

**Keywords:** latent variable analysis, item factor analysis, structural equation modeling, ordinal variables.

## Introducción

En su tarea de describir y explicar sus objetos de estudio, los investigadores sociales habitualmente se enfrentan ante las dificultades que encierra el que los constructos más interesantes desde un punto de vista teórico (e.g., clase social, ideología, identidad social, etc.) y potencialmente con mayor capacidad explicativa de los fenómenos de interés, son constructos complejos y multidimensionales que no son posibles de medir de manera directa pues constituyen lo que se denomina *variables latentes* (Borsboom, 2008) las cuales sólo pueden ser medidas indirectamente a través de múltiples variables observadas denominadas *indicadores* (i.e., variables observadas de tipo continuo) o *ítems* (variables observadas de tipo categórico ordinal o dicotómico).

Estas dificultades no siempre son recogidas en la práctica de investigación de las diversas disciplinas que componen las ciencias sociales. En el campo de la sociología cuantitativa, predomina la tradición de investigación social por encuestas (Ibáñez, 1986), que tiende a utilizar preguntas simples que suponen un continuo unidimensional, para medir conceptos notoriamente multidimensionales y latentes (e.g., posición política, religiosidad, etc.), lo cual implica negar por omisión la naturaleza compleja de los constructos de interés, mermando así el desarrollo teórico en la disciplina. Por otro lado, incluso cuando los investigadores sociales asumen el carácter complejo y latente de sus constructos y emplean múltiples ítems o indicadores para medir las variables latentes, en el análisis de datos tienden a emplear técnicas estadísticas diseñadas para variables observadas, y no técnicas específicas para variables latentes (Ramos-Zincke, 2005), lo que, como se verá en las próximas páginas, puede conducir a importantes distorsiones en los resultados obtenidos.

Es posible que estas prácticas tengan su origen en el desconocimiento de las consecuencias que ellas implican para la capacidad que tendrán los investigadores sociales de describir y predecir su objeto de estudio. En el presente manuscrito trataremos de mostrar de manera simple, empleando ejemplos con datos simulados, las ventajas teórico-metodológicas que las técnicas estadísticas modernas, específicamente diseñadas para el análisis de variables latentes, ofrecen al investigador en ciencias sociales.

## La complejidad de la realidad social

La reflexión sobre la complejidad del objeto de estudio de las ciencias sociales siempre ha sido un lugar común en las distintas disciplinas del campo, no obstante, desde hace unas décadas la noción de *complejidad* ha ido adquiriendo un estatus epistemológico especial, llegando a defenderse la existencia de dos paradigmas científicos (Lewin, 2002; Vivanco, 2010): el tradicional paradigma clásico y el nuevo paradigma de la complejidad.

Más allá que pueda sorprender la ausencia de complejidad de esta propuesta dicotómica que pretende sintetizar la múltiple y diversa evolución de las disciplinas científicas, no cabe duda que a distintas velocidades y en distintos momentos, los diversos campos de la investigación social han ido asumiendo que trabajan con un objetos de estudio multidimensionales y latentes, que tienen una importante capacidad de agencia (Epstein, 2002), que deben ser comprendidos en su interrelación global y sistémica con otros sistemas (Luhmann, 1998; Meadows, Randers & Meadows, 2006), y que tienen relaciones lineales, no lineales y multidireccionales otros sistemas y variables externas (Cilliers, 2000).

Una manera de lidiar con este reconocimiento de la complejidad social, ha consistido en el desarrollo y uso de estrategias metodológicas que, al menos en principio, parecen epistemológica y técnicamente mejor situadas para estudiar situaciones complejas, como por ejemplo, la investigación cualitativa, estructural o participante (Martínez, 2006). No obstante, en esta misma línea, la investigación cuantitativa ha evolucionado desarrollando técnicas de

producción, análisis y procesamiento de información que permiten asumir y estudiar la complejidad social.

Dentro de esas nuevas tecnologías de investigación cuantitativa destacan, por ejemplo, los modelos de redes neuronales (cf, Abdi, Valentin & Edelman, 1999; Garson, 1998), las técnicas de simulación basadas en agentes (cf, Gilbert & Tema, 2000; Macy & Willer, 2002; Medina, 2011), la estadística para variables difusas (cf, Ragin & Pennings, 2005), el modelamiento de relaciones no lineales (cf, Kelava et al., 2011; Rdz-Navarro & Alvarado, 2015), entre otras. Evidentemente desarrollar y mostrar las aplicaciones y potencialidades de todas estas tecnologías excede las posibilidades de este artículo, por lo que la mirada se centrará en una de ellas que consideramos especialmente productiva para la investigación social actual: el uso de estadística para variables latentes medidas a través de ítems ordinales, como contrapartida al uso de la estadística para variables observadas continuas en la investigación social.

### La estadística para análisis de variables latentes medidas con ítems ordinales

Desde que a mediados del siglo XX Stevens (1946) propuso el concepto de niveles de medición (i.e., medición nominal, ordinal, intervalar y de razón), ha existido una notoria división en los estadísticos sociales. Algunos han argumentado que el uso de algunas operaciones estadísticas no es legítimo en presencia de datos ordinales como los que se obtienen al utilizar tests o escalas (DiStefano, 2002; Holgado-Tello, Chacón-Moscoso, Barbero-García, & Vila-Abad, 2010; Michell, 2009). Otros han argumentado que esa limitación es irrelevante en la práctica, por lo que es posible aplicar cualquier análisis estadístico a ese tipo de datos, homologándolos a información con nivel de medición de intervalo o de razón (Jamieson, 2004; Carifio & Perla, 2007; Norman, 2010). Y otros han argumentado que usar la estadística diseñada para variables de intervalo para analizar variables ordinales es una alternativa aceptable, pero sólo en determinadas situaciones (Bollen & Barb, 1981).

Pese a que el debate aún continúa abierto, la evidencia más reciente parece ser bastante concluyente en que existen condiciones y análisis estadísticos en los que no es inocuo tratar a datos ordinales como si fueran de intervalo. Por ejemplo, se ha demostrado que la correlación de Pearson –estadístico diseñado para variables de intervalo– aplicada a variables ordinales tiende a estimar de forma distorsionada la magnitud de la relación entre las variables y que la magnitud de esa distorsión no es constante, sino que varía en función de la asimetría de las variables analizadas (Holgado-Tello, Chacón-Moscoso, Barbero-García, & Vila-Abad, 2010), lo que tiende a generar importantes errores en los análisis estadísticos multivariados emplean como base el análisis de correlaciones (Berstein & Teng, 1989). Afortunadamente, hoy en día existen algunas soluciones que permiten lidiar con este problema. La más popular de ellas es estimar la magnitud de la relación entre dos variables ordinales con un estadístico alternativo: las correlaciones policóricas (denominadas también correlaciones tetracóricas cuando las variables analizadas son dicotómicas), desarrolladas a principios del siglo XX, pero recientemente diseminadas en el campo de las ciencias sociales gracias a los avances tecnológicos en informática.

La idea detrás de la estimación de las correlaciones policóricas es relativamente sencilla. El procedimiento asume que tras cada variable ordinal observada existe una variable continua subyacente con distribución normal que habría podido ser medida si dispusiésemos de un instrumento con infinitas categorías de respuesta, cuestión que en la práctica no es posible (Asún, Rdz-Navarro & Alvarado, 2016). Así, al utilizar preguntas o ítems con un número limitado de categorías de respuesta que se presentan al sujeto que responde (e.g., variables tipo Likert con 5 alternativas de respuesta que van desde el ‘muy de acuerdo’ al ‘muy en desacuerdo’) se obtiene una medición tosca de la variable continua subyacente, registrándola en un número de categorías ( $k$ ) definidas en función de  $k - 1$  parámetros *umbrales* ( $\tau_k$ ) que

constituyen los puntos de corte de la variable continua que dieron lugar a la distribución de frecuencias observada en la variable ordinal. De este modo, distintos valores de parámetros umbrales darán origen a distintas distribuciones de las variables ordinales efectivamente medidas, como se muestra en la Figura 1.

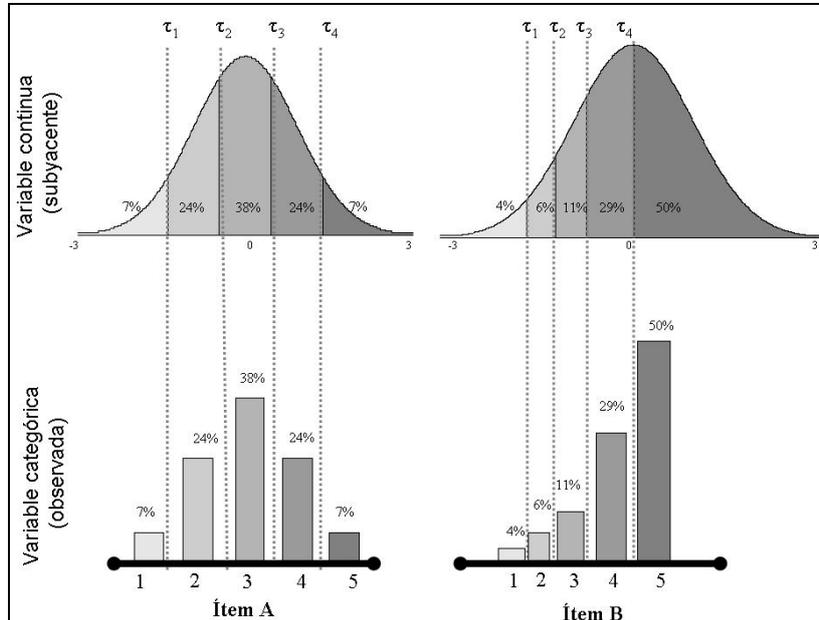


Figura 1. Relación entre la variable continua normal subyacente y las variables observadas ordinales.

De este modo, en lugar de centrarse en el cálculo de la correlación entre las variables observadas, las correlaciones policóricas permiten estimar la relación existente entre las variables continuas subyacentes, de modo que la estimación de la magnitud de la relación entre las variables no se vea afectada por el grado de asimetría introducido por la recodificación arbitraria de las variables que generan los parámetros umbrales. En consecuencia, actualmente se recomienda que toda vez que se esté en presencia de ítems ordinales, se empleen matrices de correlaciones policóricas para estimar la magnitud de la relación entre las variables.

El interés de los científicos sociales no sólo se centra en el análisis de la relación entre variables observadas sino principalmente en el análisis de variables latentes. Como se mencionó previamente, las variables latentes son constructos que no tienen un correlato empírico directo, por lo que deben ser medidos a partir de múltiples indicadores o ítems que son manifestaciones empíricas de la variable latente. Así por ejemplo, en investigación por encuestas, se suelen medir variables latentes a través de escalas tipo Likert, es decir, de un conjunto preguntas simples con respuesta cerrada que se suponen miden el mismo constructo complejo, ante las cuales las personas manifiestan su grado de acuerdo seleccionando una respuesta dentro de una serie de categorías generalmente ordenadas en función de su intensidad (Likert, 1932).

Se ha demostrado que el uso múltiples preguntas o ítems para la medición de constructos latentes contiene menos error de medida que intentar medir una variable latente con una sola pregunta, pues los errores de medida introducidos en el fraseo o la respuesta a cada uno de ellos tiende a compensarse con los errores en direcciones opuestas de los otros ítems (Likert, Roslow & Murphy, 1934), pero ello sólo es cierto si es que se puede obtener evidencia de que realmente se está midiendo el mismo constructo latente con el conjunto de

ítems empleado, aunque dicho constructo sea multidimensional y distintos grupos de ítems hayan sido diseñados para medir sus diferentes dimensiones. Una de las herramientas para aportar ese tipo de evidencia es el análisis factorial confirmatorio, pues permite determinar si tras un conjunto de datos es posible suponer la presencia de una variable latente común, compuesta por aquella porción de la varianza compartida por todos los ítems de la escala (cf, Ferrando & Anguiano-Carrasco, 2010; Mulaik, 2009).

El análisis factorial confirmatorio (AFC) no sólo permite determinar si un conjunto de ítem miden un mismo constructo, sino que al mismo tiempo, permite determinar si el constructo medido es unidimensional o multidimensional, es decir, permite saber si una variable latente está compuesta por distintos dominios teóricos o temáticos que tienden a tener cierta autonomía entre si, de manera que, pese a conformar un constructo común, las personas pueden tener grados diferenciales de magnitud en cada dominio teórico. Si ese fuese el caso, para caracterizar correctamente a los sujetos no bastaría con obtener una sola puntuación, sino que se deben obtener tantas puntuaciones como dimensiones tenga el constructo, agregando las respuestas que los sujetos dieron a cada uno de los ítems que, según el diseño del instrumento y la teoría confirmada por el AFC, pertenecen a cada dimensión.

En términos generales, existen dos versiones de AFC: El denominado *AFC clásico*, diseñado para analizar un conjunto de indicadores continuos que opera mediante el análisis de covarianzas o correlaciones de Pearson, y el *AFC de ítems* (Wirth & Edwards, 2007), también conocido como AFI (i.e., análisis factorial de ítems), que permite reconocer y abordar explícitamente el carácter ordinal de las variables observadas en la investigación social, pues opera sobre la base de matrices de correlación policórica –o tetracóricas-. Por ello, se tiende a afirmar que el AFI es una técnica que permite estimar los “modelos de medida” de variables latentes complejas medidas con ítems (Cupani, 2012).

Finalmente, dado que el objetivo del investigador social o sociológico al medir variables latentes está habitualmente centrado en determinar su relación con otros constructos también latentes, es importante disponer de formas de evaluar dichas relaciones. La evidencia ha mostrado que cualquier intento de obtener puntuaciones observadas globales de un constructo medido a través de múltiples ítems –ya sea empleando la simple sumatoria de respuestas, puntuaciones factoriales o estimaciones theta derivadas de la teoría de respuesta al ítem- conduce a introducir cierto grado de error de medida en la puntuación global. El error de medida sesgará la estimación de la verdadera relación entre los constructos, afectando así la explicación de los fenómenos. Es por ello que una técnica que permita estimar la asociación entre variables latentes sin tener que estimar sus puntuaciones observadas globales, permitirá una mejor apreciación de la relación entre ambas. Una de esas técnicas son los modelos de ecuaciones estructurales (o modelos SEM, por sus siglas en inglés), que tienen la ventaja añadida de permitir incorporar también la estimación de efectos directos e indirectos de las variables independientes sobre las dependientes (Batista-Foguet & Coenders, 2000; Muthén & Asparouhov, 2015), además de relaciones lineales y no lineales (cf, e.g., Rdz-Navarro & Alvarado, 2015). Dado que se puede argumentar que los modelos de ecuaciones estructurales son una generalización del análisis factorial confirmatorio, que permite extenderlo desde los modelos de medida de variables latentes hacia los modelos de estructura o de relación entre variables latentes, ambas técnicas conforman una unidad coherente denominada *análisis de variables latentes* (Cupani, 2012; Muthén, 2002).

A continuación intentaremos, utilizando datos simulados y reales, mostrar algunas de las ventajas que se pueden obtener de emplear estas técnicas de variable latente: a) la capacidad del AFI para evidenciar la verdadera estructura multidimensional de una variable latente y, b) la capacidad de los modelos SEM para estimar la verdadera magnitud de relación entre un conjunto de variables independientes, intervinientes y dependientes.

## Estudios Monte Carlo

### Estudio 1: Correlaciones de Pearson versus correlaciones policóricas

Para mostrar las ventajas del uso de las correlaciones policóricas para la investigación social imaginamos un ejemplo ficticio concreto, suponiendo la existencia de un investigador que busca medir un constructo teórico unidimensional continuo (e.g., actitudes hacia los inmigrantes), y para ello elabora 12 preguntas a ser aplicadas a una muestra a través de una encuesta. Como al investigador no le es posible formular cada pregunta de modo que los sujetos respondan con un número infinito de categorías de respuesta, decide utilizar un formato de respuesta cerrado tipo Likert de cinco alternativas que van del muy en desacuerdo al muy de acuerdo y se puntúan de 1 a 5. Una vez aplicado su instrumento, el investigador se ve en la necesidad de evaluar cuán bien (o mal) mide cada pregunta el constructo teórico, para lo cual decide utilizar el análisis factorial confirmatorio.

Con esta situación hipotética en mente, se simuló una población en la cual existía un factor continuo (F) con distribución normal estándar (i.e., media igual a 0 y desviación típica igual a 1) que se relacionaba con 12 indicadores continuos (i.e.,  $V_i$ ) que simbolizaban las respuestas que el investigador habría obtenido si su instrumento o dispositivo de medición le hubiese permitido un formato de infinitas respuestas. La relación del factor con los indicadores continuos está dada por parámetros lambda (i.e.,  $\lambda_i$ ), también llamados *cargas factoriales*, que expresan la magnitud de la relación entre el factor y el indicador de manera similar a una correlación. Los parámetros lambda se simularon de acuerdo a dos valores (i.e., .3 y .7) de modo tal que indicadores fuesen de regular calidad y otros de buena calidad.

A continuación, se seleccionó una serie de valores para los parámetros umbrales (i.e.,  $\tau_{k,i}$ ) que sirvieron para transformar las variables continuas subyacentes en ítems de tipo ordinal con cinco categorías de respuesta (i.e., variables  $X_i$ ) y definen la relación entre las variables continuas  $V_i$  (no medidas) y las variables ordinales  $X_i$  (efectivamente medidas) como se muestra en la Figura 2.

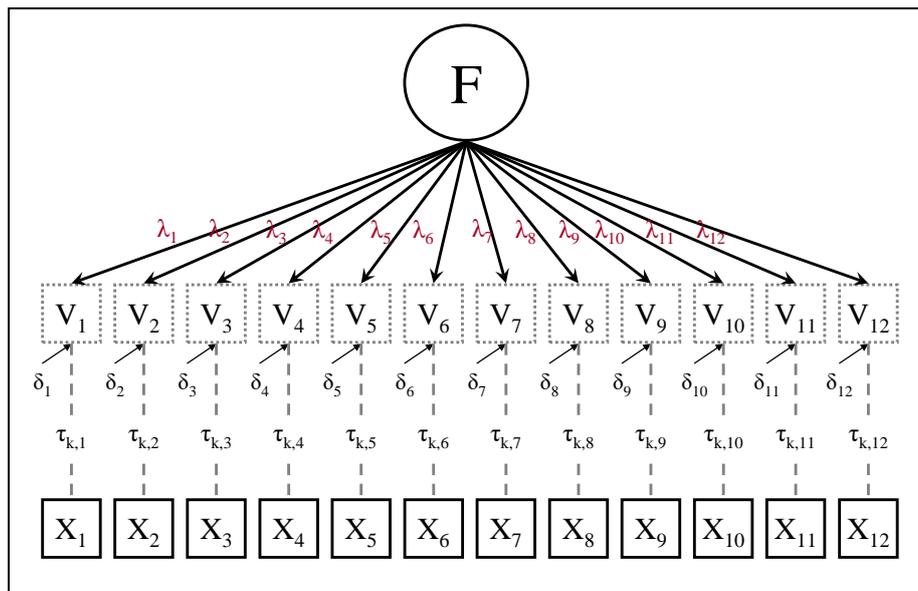


Figura 2. Modelo factorial simulado

Debido a que en la práctica es habitual encontrar situaciones en las cuales las respuestas de los sujetos a los ítems presentan distintos niveles de asimetría, se seleccionaron distintos tipos de valores de parámetros umbrales que dieron lugar a cinco tipos de distribuciones de ítems: simétricos, con asimetría moderada (positiva y negativa) y asimetría alta (positiva y negativa), como se muestra en la Figura 3. Para simplificar la simulación y el análisis de datos, los 4 primeros ítems fueron generados con una distribución simétrica, los cuatro siguientes con una distribución con asimetría moderada y los siguientes con asimetría alta. En las situaciones de asimetría, se generaron dos ítems con asimetría positiva y dos con asimetría negativa.

Una vez generados los datos de la población en el software R versión 3.1.0, se obtuvieron 500 muestras de 500 sujetos cada una, para evitar que los resultados fueran afectados por las características aleatorias de cada muestra particular de respuestas. Cada una de las muestras fue analizada con el software Mplus 7.4 (Muthén & Muthén, 1998-2015), utilizando dos versiones del AFC: su versión clásica en base a correlaciones de Pearson y su versión para ítems en base a correlaciones policóricas. Los resultados fueron analizados globalmente, examinando el promedio del parámetro lambda estimado para cada ítem en ambos tipos de análisis y el porcentaje sesgo relativo de dicho parámetro obtenido con cada estimación. Se consideró aceptable obtener un sesgo relativo de cada parámetro igual o inferior a 5% en valores absolutos.

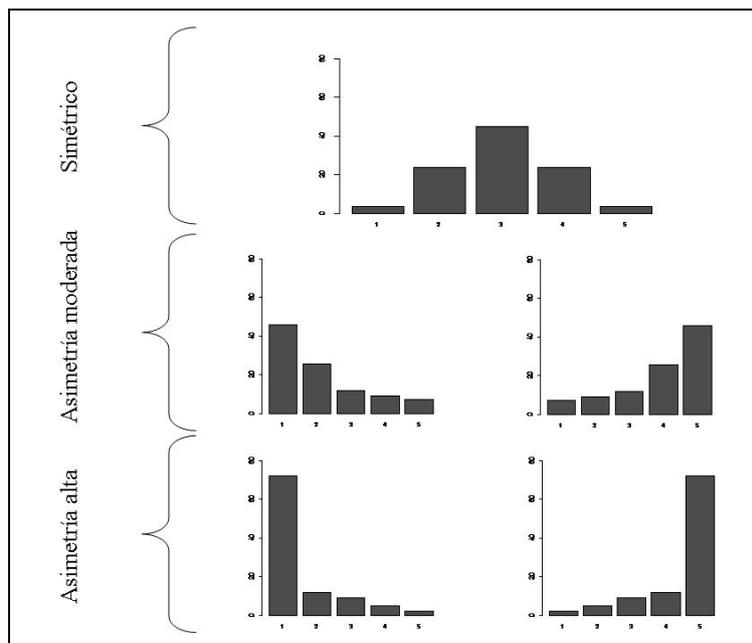


Figura 3. Tipos de ítems simulados

## Resultados

En la Tabla 1 se puede apreciar que los parámetros lambdas (es decir, la magnitud de la relación entre la variable latente y la respuesta dada a los indicadores observados) tiende a recuperarse de manera ligeramente atenuada, pero este sesgo está dentro del rango aceptable (i.e., es inferior a un 5%) cuando se emplea AFC con correlaciones policóricas (i.e., AFI), mientras que alcanza valores inaceptablemente altos (cerca de un 25% de subestimación) cuando el AFC se realiza con correlaciones de Pearson y existe asimetría en las respuestas observadas.

Es interesante notar que la magnitud de la atenuación de los parámetros lambda obtenidos al utilizar correlaciones de Pearson depende del grado de asimetría de los ítems, con lo que si se hubiese utilizado un análisis factorial exploratorio con correlaciones de Pearson, es probable que los ítems con asimetrías similares hubiesen agrupado entre si, formando factores espurios o falsos (Berstein & Teng, 1989; Green, 1983), que habrían confundido aún más la interpretación teórica del constructo.

Tabla 1. Resultados de la simulación modelo factorial

	Análisis		AFC		AFC	
	Ti	$\lambda_r$	Pearson		Policóricas	
			Media	%Sesgo	Media	%Sesgo
X <sub>1</sub>	S	.70	.679	-3.0	.671	-4.1
X <sub>2</sub>	S	.30	.288	-4.0	.287	-4.3
X <sub>3</sub>	S	.70	.675	-3.6	.669	-4.4
X <sub>4</sub>	S	.30	.291	-3.0	.289	-3.7
X <sub>5</sub>	M <sub>P</sub>	.70	.616	<b>-12.0</b>	.670	-4.3
X <sub>6</sub>	M <sub>P</sub>	.30	.268	<b>-10.7</b>	.289	-3.7
X <sub>7</sub>	M <sub>N</sub>	.70	.618	<b>-11.7</b>	.671	-4.1
X <sub>8</sub>	M <sub>N</sub>	.30	.266	<b>-11.3</b>	.287	-4.3
X <sub>9</sub>	A <sub>P</sub>	.70	.526	<b>-24.9</b>	.670	-4.3
X <sub>10</sub>	A <sub>P</sub>	.30	.233	<b>-22.3</b>	.289	-3.7
X <sub>11</sub>	A <sub>N</sub>	.70	.527	<b>-24.7</b>	.670	-4.3
X <sub>12</sub>	A <sub>N</sub>	.30	.233	<b>-22.3</b>	.286	-4.7
Fiabilidad		.80	.73	<b>-8.4</b>	.78	-2.8

Nota: AFC: análisis factorial confirmatorio. Ti: tipo de ítem.  $\lambda_r$ : parámetro lambda poblacional (valor real). S: ítem simétrico. M<sub>P</sub>: ítem con asimetría positiva moderada. M<sub>N</sub>: ítem con asimetría negativa moderada. A<sub>P</sub>: ítem con asimetría alta positiva. A<sub>N</sub>: ítem con asimetría alta negativa. Negrilla: sesgo inaceptable.

## Estudio 2: Detección de la relación entre variables latentes

Con el fin de poner en evidencia la pérdida de poder explicativo sobre una variable dependiente al trabajar con estadística para variables observadas en vez de estadística para variables latentes, se simuló el modelo descrito en la Figura 4, donde dos variables independientes cualesquiera (F1 y F2) moderadamente correlacionadas entre sí ( $\phi_{21} = .3$ ), se emplean para explicar una tercera variable F3.

Cada uno de los factores o variables latentes fue medido usando 4 ítems de 5 alternativas de respuesta puntuados del 1 al 5, según tres tipos de asimetría. En la primera condición, todos los ítems tenían distribuciones simétricas. En la segunda condición, todos los ítems tenían distribuciones moderadamente asimétricas en la misma dirección (i.e., asimetría unilateral). Mientras que en la tercera condición, todos los ítems tenían distribuciones moderadamente asimétricas, pero en distintas direcciones (i.e., asimetría alterna): la mitad tenían asimetría negativa, mientras que la otra mitad tenían asimetría positiva.

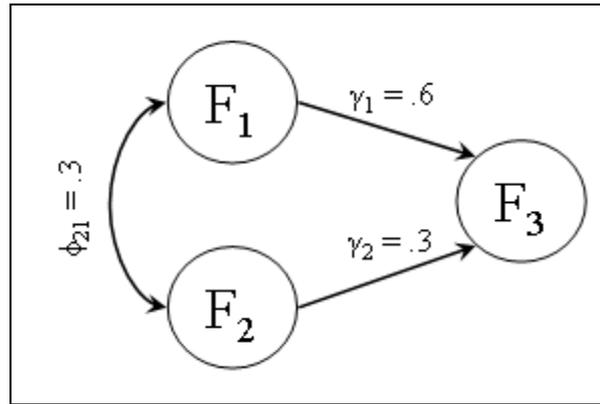


Figura 4. Modelo causal poblacional simulado. F<sub>1</sub> y F<sub>2</sub> son factores latentes exógenos (variables independientes). F<sub>3</sub> es un factor latente endógeno (variable dependiente).  $\phi_{21}$  representa la correlación entre los factores exógenos.  $\gamma_1$  es la magnitud del efecto de F<sub>1</sub> sobre F<sub>3</sub> y  $\gamma_2$  es la magnitud del efecto de F<sub>2</sub> sobre F<sub>3</sub>.

Una vez generados los datos de la población en R versión 3.1.0, se obtuvieron al azar 500 muestras de 1000 sujetos cada una que habrían respondido el cuestionario de 12 ítems (i.e., 4 ítems por cada factor). Cada una de las muestras fue analizada con el software Mplus 7.4 (Muthén & Muthén, 1998-2015) de dos maneras: (a) realizando una regresión lineal tradicional -que es un análisis estadístico para variables observadas-, empleando la estimación ML (que en este caso es matemáticamente equivalente a la estimación por mínimos cuadrados -OLS- que proponen por defecto los más difundidos software de análisis estadístico como SPSS), sobre la suma de puntos obtenidos en los 4 ítems que miden cada factor; (b) un análisis con un modelo SEM basado en las correlaciones policóricas con estimación ULSMV, que es un procedimiento estadístico para variables latentes diseñado para trabajar con variables observadas ordinales.

Los resultados fueron analizados globalmente, examinando los parámetros  $\gamma$  (también conocidos como parámetros Beta en el contexto del análisis de regresión) estimados en ambos análisis y el porcentaje sesgo relativo de dicho parámetro obtenido con cada estimación. Al igual que en la simulación anterior, se consideró aceptable obtener un sesgo relativo de cada parámetro igual o inferior a 5%.

## Resultados

Si se comparan en la Tabla 2 los parámetros poblacionales simulados con los obtenidos por el análisis SEM, se observará que ellos siempre fueron estimados dentro de los límites de lo aceptable, aunque usualmente con leves sobreestimaciones. Por el contrario, los análisis de regresión obtuvieron en todos los casos subestimaciones inaceptables del valor real del parámetro simulado.

Si estos fuesen los resultados de un estudio con datos reales y la estrategia de análisis utilizada hubiese sido aplicar una regresión lineal sobre la sumatoria de puntuaciones de los ítems individuales tratados como una sola puntuación compuesta de cada factor, se habrían obtenido tres conclusiones erróneas, crecientemente más graves:

1. La correlación entre las variables independientes hubiera parecido menor que la real, llevándonos a creer que se trataba de dos variables mucho menos relacionadas entre sí;

2. Los efectos de las variables independientes (F1 y F2) sobre la variable dependiente (F3) se juzgarían como mucho menores que lo que son en la realidad, incluso pudiendo alcanzar valores no significativos si dispusiéramos de muestras más pequeñas;
3. La capacidad global que tendría nuestro modelo de explicar la variable dependiente habría aparecido mucho menor de lo real –un 50% menor– lo que nos habría llevado a creer que las variables que estamos empleando y la teoría que sustenta la inclusión de dichas variables no tiene suficiente capacidad predictiva. Con lo cual, erróneamente podríamos haber concluido que la teoría debería ser reemplazada por otra con mayor capacidad predictiva.

Tabla 2. Resultados de la simulación modelo causal

Parámetro poblacional y su valor			$\phi_{21}$	$\gamma_1$	$\gamma_2$	VE
			.300	.600	.300	.558
Ti	Tipo de análisis					
S	Regresión	Media	.211	.432	.241	.291
		P.Sesgo	<b>-.297</b>	<b>-.280</b>	<b>-.197</b>	<b>-.478</b>
	SEM	Media	.305	.603	.295	.562
		P.Sesgo	.017	.005	-.017	.007
M <sub>P</sub>	Regresión	Media	.192	.413	.235	.266
		P.Sesgo	<b>-.360</b>	<b>-.312</b>	<b>-.217</b>	<b>-.523</b>
	SEM	Media	.300	.605	.303	.573
		P.Sesgo	.000	.008	.010	.027
M <sub>P<sub>y</sub>N</sub>	Regresión	Media	.201	.407	.235	.262
		P.Sesgo	<b>-.330</b>	<b>-.322</b>	<b>-.217</b>	<b>-.530</b>
	SEM	Media	.309	.601	.300	.567
		P.Sesgo	.030	.002	.000	.016

Nota:  $\phi_{21}$ : correlación entre los factores exógenos (variables independientes, F1 y F2).  $\gamma_1$  : magnitud del efecto de F1 sobre el factor endógeno (variable dependiente, F3).  $\gamma_2$  : magnitud del efecto de F2 sobre el factor endógeno (variable dependiente, F3). VE: varianza explicada global del modelo. Ti: tipo de ítems usados para medir los factores. S: ítems simétricos. M<sub>P</sub>: ítems con asimetría positiva moderada. M<sub>P<sub>y</sub>N</sub>: ítems con asimetría positiva y negativa moderada. P.Sesgo: proporción de sesgo de la estimación del parámetro.

Estas consecuencias negativas pueden redundar en importantes retrasos en el conocimiento científico en ciencias sociales, pero afortunadamente son evitables, pues conforme a los resultados obtenidos en este estudio, el uso de modelos SEM adaptados para trabajar con variables observadas de tipo ordinal y variables latentes continuas subyacentes, permite obtener una representación más precisa de la realidad que se está investigando y, por lo tanto, obtener mayor seguridad respecto de los resultados obtenidos en una investigación.

## Conclusiones

Como se ha mostrado en los ejemplos precedentes, el emplear técnicas estadísticas para variables observadas de intervalo, en lugar de estadística para variables latentes (no observadas) y variables ordinales, tiene consecuencias negativas para los modelos de medida y los modelos de estructura con los que se trabaja en sociología y ciencias sociales en general. En los modelos de medida, emplear la estadística tradicional puede llevar a distorsiones en la interpretación de los constructos medidos, obteniéndose representaciones con un número de dimensiones diferente a las presentes realmente en los datos y con un ordenamiento de ítems que puede llevar a los investigadores a reestructurar erróneamente sus caracterizaciones conceptuales. Y en los modelos de estructura, puede tener como consecuencia una importante disminución de los tamaños de los efectos detectados, con lo que algunas relaciones presentes en los datos a nivel poblacional, no serán detectadas como significativas, mientras que aquellas que sean detectadas como significativas, serán percibidas como de fuerza más baja que lo real.

Esta pérdida de potencia y precisión de los análisis estadísticos sociales, no sólo es relevante por implicar una posible incapacidad de detectar las relaciones existentes entre las variables a nivel poblacional y el consecuente retraso que esto puede implicar en el avance de la construcción de conocimiento sociológico basado en la evidencia, sino porque tiende a producir que los sociólogos se conformen con la detección de relaciones significativas entre variables, abdicando –posiblemente por no ser capaces de detectarlas– en la búsqueda de relaciones fuertes con alta capacidad explicativa de las variables dependientes. Quizá ello explique que tantas investigaciones en sociología pasen por alto o simplemente no incluyan en sus resultados el reporte de los *tamaños de efecto* (i.e., estadísticos que evalúan la fuerza de la relación) encontrados, y en cambio pongan exclusiva atención a la detección de relaciones significativas.

Esta estrategia centrada en la detección y reporte de significación estadística no sólo constituye una limitación para el desarrollo del conocimiento sociológico, sino que tiene connotaciones epistemológicas, pues es una herramienta de construcción de conocimiento más limitada y potencialmente sesgada que aquella centrada en los tamaños del efecto, ya que es proclive a confundir la presencia de significación con la importancia teórica o relevancia social de la relación detectada (Chow, 1988; Fernández-Cano, 2009), tiende a desechar la heurísticamente pobre hipótesis nula de no existencia de relación entre variables, sin avanzar en el contraste de modelos alternativos en base a su capacidad predictiva efectiva, lo cual podría explicar el limitado avance de las disciplinas de las ciencias sociales (Ioannidis, 2012; Meehl, 1978).

Creemos que el análisis de variables latentes, al centrarse en determinar el ajuste y capacidad predictiva de modelos teóricos, es una gran oportunidad para avanzar en la forma en que realizamos investigación en metodología cuantitativa en ciencias sociales, mejorando la relación entre la construcción de teorías y la búsqueda de evidencia empírica que las sustente o refute.

## Referencias

- Abdi, H., Valentin, D., & Edelman, B. (1999). *Neural Networks (Quantitative Applications in the Social Sciences)*. London: Sage.
- Asún, R. A., Rdz-Navarro, K., & Alvarado, J. M. (2016). Developing multidimensional Likert scales using item factor analysis: The case of four-point items. *Sociological Methods and Research*, 45(1), 109-133.
- Batista-Foguet, J. M., & Coenders, G. (2000). *Modelos de ecuaciones estructurales*. Madrid: La Muralla.
- Bernstein, I. H., & Teng, G. (1989). Factoring items and factoring scales are different: Spurious evidence for multidimensionality due to item categorization. *Psychological Bulletin*, 105(3), 467-477.
- Bollen, K. A., & Barb, K. H. (1981). Pearson's r and coarsely categorized measures. *American Sociological Review*, 46(2), 232-239.
- Borsboom, D. (2008). Latent variable theory. *Measurement: Interdisciplinary Research and Perspectives*, 6(1), 25-53.
- Carifio, J., & Perla, R. J. (2007). Ten common misunderstandings, misconceptions, persistent myths and urban legends about Likert scales and Likert response formats and their antidotes. *Journal of Social Sciences* 3(3), 106-116.
- Chow, S. L. (1988). Significance test or effect size? *Psychological Bulletin*, 103(1), 105-110.
- Cilliers, P. (2000). What can we learn from a theory of complexity? *Emergence*, 2(1), 23-33.
- Cupani, M. (2012). Análisis de ecuaciones estructurales: conceptos, etapas de desarrollo y un ejemplo de aplicación. *Revista Tesis Facultad de Psicología*, 2(1), 186-199.
- DiStefano, C. (2002). The impact of categorization with confirmatory factor analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(3), 327-346.
- Epstein, J. M. (2002). Modeling civil violence: An agent-based computational approach. *Proceedings of the National Academy of Sciences*, 99(suppl 3), 7243-7250.
- Fernández-Cano, A. & Fernández-Guerrero, I. M. (2009). *Crítica y alternativas a la significación estadística en el contraste de hipótesis*. Madrid: La Muralla.
- Ferrando, P. J., & Anguiano-Carrasco, C. (2010). El análisis factorial como técnica de investigación en psicología. *Papeles del Psicólogo*, 31(1), 18-33.
- Garson, G. D. (1998). *Neural networks: An introductory guide for social scientists*. London: Sage.
- Gilbert, N., & Terna, P. (2000). How to build and use agent-based models in social science. *Mind & Society*, 1(1), 57-72.
- Green, S. B. (1983). Identifiability of spurious factors using linear factor analysis with binary items. *Applied Psychological Measurement*, 7(2), 139-147.
- Holgado-Tello, F. P., Chacón-Moscó, S., Barbero-García, I., & Vila-Abad, E. (2010). Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. *Quality & Quantity*, 44(1), 153-166.
- Ibáñez, J. (1986). Perspectivas de la investigación social: el diseño en las tres perspectivas. En M. García-Ferrando, J. Ibáñez & F. Alvira (Eds.), *El análisis de la realidad social: métodos y técnicas de investigación* (pp. 42-75). Madrid: Alianza.
- Ioannidis, J. P. (2012). Why science is not necessarily self-correcting. *Perspectives on Psychological Science*, 7(6), 645-654.
- Jamieson, S. (2004). Likert scales: How to (ab)use them. *Medical Education*, 38(12), 1212-1218.
- Kelava, A., Werner, C. S., Schermelleh-Engel, K., Moosbrugger, H., Zapf, D., ... West, S.G. (2011). Advanced nonlinear latent variable modeling: distribution analytic LMS and QML

- estimators of interaction and quadratic effects. *Structural Equation Modeling: A Multidisciplinary Journal*, 18(3), 465-491.
- Lewin, R. C. (2002). *El caos como generador del orden*. Barcelona: Tusquets Editores.
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 22(140), 44-55.
- Likert, R., Roslow, S., & Murphy, G. (1934). A simple and reliable method of scoring Thurstone attitudes scales. *The Journal of Social Psychology*, 5(2), 228-238.
- Luhmann, N. (1998). *Sistemas sociales: lineamientos para una teoría general*. Anthropos Editorial.
- Martínez, M. (2006). La investigación cualitativa (síntesis conceptual). *Revista de investigación en psicología*, 9(1), 123-146.
- Meadows, D. H., Randers, J., & Meadows, D. L. (2006). *Los límites del crecimiento: 30 años después*. Galaxia Gutenberg.
- Macy, M. W., & Willer, R. (2002). From factors to actors: Computational sociology and agent-based modeling. *Annual Review of Sociology*, 28(1), 143-166.
- Medina, J. I. G. V. (2011). La simulación basada en agentes: una nueva forma de explorar los fenómenos sociales/Agent-based Modelling: A New Way of Exploring Social Phenomena. *Reis*, 91-109.
- Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of consulting and clinical Psychology*, 46(4), 806-834.
- Michell, J. (2009). The psychometricians' fallacy: Too clever by half? *British Journal of Mathematical Statistical Psychology*, 62, 41-55.
- Mulaik, S. A. (2009). *Foundations of factor analysis*. London: CRC Press.
- Muthén, B. O. (2002). Beyond SEM: General latent variable modeling. *Behaviormetrika*, 29(1), 81-117.
- Muthén, B., & Asparouhov, T. (2015). Causal effects in mediation modeling: An introduction with applications to latent variables. *Structural Equation Modeling: A Multidisciplinary Journal*, 22(1), 12-23.
- Muthén, L.K. and Muthén, B.O. (1998-2015). *Mplus user's guide. Seventh edition*. Los Angeles, CA: Muthén & Muthén.
- Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education*, 15(5), 625-632.
- Ragin, C. C., & Pennings, P. (2005). Fuzzy sets and social research. *Sociological Methods & Research*, 33(4), 423-430.
- Ramos-Zincke, C. (2005). Cómo investigan los sociólogos chilenos en los albores del siglo XXI: Paradigmas y herramientas del oficio. *Persona y Sociedad*, 19(3), 85 - 119.
- Rdz-Navarro, K., & Alvarado, J. M. (2015). Reexamining nonlinear structural equation modeling procedures: The effect of parallel and congeneric measures. *Multivariate Behavioral Research*, 50(6), 645-641.
- Stevens, S. S. (1946). On the theory of scales of measurement. *Science*, 103(2684), pp. 677-680.
- Vivanco, M. (2010). *Sociedad y complejidad: del discurso al modelo*. Santiago: Lom Ediciones.
- Wirth, R. J., & Edwards, M. C. (2007). Item factor analysis: current approaches and future directions. *Psychological Methods*, 12(1), 58-79.