

DOI: <http://dx.doi.org/10.15517/rce.v33i2.22216>

APLICACIÓN DEL ANÁLISIS FACTORIAL CONFIRMATORIO A UN MODELO DE MEDICIÓN DEL RENDIMIENTO ACADÉMICO EN LECTURA¹

Andrés Fernández Aráuz²

Recibido: 09/06/2015

Aprobado: 17/11/2015

RESUMEN

Mediante un ejemplo empírico, se muestran las diferencias entre las técnicas estadísticas de análisis factorial exploratorio y confirmatorio. En la investigación en Ciencias Sociales en nuestro país es muy frecuente el uso de variables observables para analizar la relación directa entre estas y la variable respuesta. Más aún, es también frecuente el uso de la técnica de análisis factorial exploratorio por el desconocimiento sobre la correcta aplicación del análisis factorial confirmatorio, debido a que esta última requiere del planteamiento de una hipótesis a ser evaluada, mientras que en el análisis exploratorio no es necesario que el investigador tenga una conciencia clara sobre la hipótesis que desea evaluar, sino que espera que los datos por sí mismos muestren las relaciones con la variable respuesta. Con un ejemplo empírico utilizando datos de la prueba internacional PISA 2009 se ilustra las diferencias entre ambas y las bondades de utilizar el análisis factorial confirmatorio para evaluar hipótesis de investigación.

PALABRAS CLAVE: ANÁLISIS CAUSAL, MODELOS ESTADÍSTICO, PISA.

ABSTRACT

Using an empirical example, I show the differences between the statistical techniques of exploratory and confirmatory factor analyzes. In the social science research in our country is very common the use of observable variables to analyze the direct relationship between these and the response variable. Moreover, the exploratory factor analysis is too frequently used due to the ignorance about the correct application of confirmatory factor analysis, because the latter requires the approach of a hypothesis to be tested, while in the exploratory factor analysis it is not required to have a clear idea about the hypothesis to be evaluated. With an empirical example using data from the 2009 PISA international test, I show the differences between the two approaches and the benefits of using confirmatory factor analysis to evaluate research hypotheses.

KEYWORDS: CAUSAL ANALYSIS, STATISTICAL MODEL, PISA.

- 1 Esto documento fue elaborado como trabajo final de investigación en el curso de Análisis Estadístico Multivariado (SP1628 – I Semestre 2015) del Posgrado de Estadística de la Universidad de Costa Rica.
- 2 Ministerio de Educación Pública, Departamento de Análisis Estadístico. Código postal 10102. San José Costa Rica; afa17@gmail.com

I. INTRODUCCIÓN

En Costa Rica, es muy común que en las investigaciones en Ciencias Sociales y en algunas otras áreas del conocimiento se utilicen datos observacionales para realizar análisis de relación o correlacionales, y aunque usualmente el objetivo de las investigaciones es probar hipótesis de causalidad y no de asociación, se utilizan técnicas estadísticas que no permiten al investigador realizar conclusiones en cuanto a dicha causalidad (aunque algunos se aventuran a hacerlas).

El análisis factorial confirmatorio, por el contrario, es una técnica que no es utilizada con fines exploratorios, sino que funciona como evaluación de un posible modelo de medición sobre la base de una teoría a ser probada.

A pesar de ser una herramienta tan potente, estos modelos no son tan frecuentemente utilizados en las investigaciones en ciencias sociales tanto en nuestro país como en otros debido en parte a la extendido uso de los modelos de regresión clásicos. Por lo anterior, el objetivo de este trabajo es presentar la técnica del análisis factorial confirmatorio, los aspectos a considerar para su correcta aplicación y la diferencia con el análisis factorial exploratorio, para mostrar las ventajas de la utilización de uno u otro enfoque en la investigación en ciencias sociales.

Si bien el artículo no realiza aportes metodológicos a la técnica estadística analizada, su aporte consiste en la comparación para un mismo set de datos de los resultados que se pueden obtener al utilizar un enfoque confirmatorio en lugar de uno exploratorio al momento de realizar análisis factoriales, con la intención de intentar llenar un vacío de técnicas estadísticas al alcance de los investigadores en Ciencias Sociales de nuestro país, cuyas investigaciones usualmente utilizan modelos de regresión (lineal, logit, probit, como los más comunes) que no permiten realizar conclusiones hacia la causalidad ni validar hipótesis teóricas previamente establecidas.

El análisis factorial confirmatorio se ha convertido en una de las técnicas estadísticas más utilizadas a nivel mundial, desde sus primeros exponentes (véase Joreskog, 1969 y 1971) que presentan el planteamiento teórico-matemático de estos modelos hasta recientes aplicaciones para la evaluación psicométrica de instrumentos, la validación de constructos, la estimación de efectos de posible causalidad, entre otros usos que se le pueden dar a estos modelos.

II. METODOLOGÍA

Análisis factorial

El análisis factorial es una técnica de reducción de datos que sirve para encontrar grupos homogéneos de variables a partir de un conjunto usualmente grande de variables. Esos grupos homogéneos se forman con las variables que correlacionan mucho entre sí (usualmente correlaciones superiores a 0,5 en valor absoluto) y procurando que unos grupos sean independientes de otros.

Mientras que la técnica de componentes principales trata de explicar la variabilidad total de las variables, el análisis factorial busca identificar la porción de la varianza total que es compartida por las variables, formulando un modelo que distinga fuentes de variación al describir cada variable como una combinación lineal de un número pequeño de factores comunes y un factor específico.

Desde sus primeros usos hace más de un siglo (Spearman, 1904 y 1927), el análisis factorial se ha convertido en uno de los procedimientos estadísticos multivariantes más utilizados en la investigación aplicada en una gran variedad de campos de aplicación (por ejemplo, la psicología, la educación, la sociología, la salud pública, entre otros). La principal intención del análisis factorial es determinar el número y la naturaleza de las variables latentes o factores que explican la variación y covariación entre un conjunto de medidas observadas, comúnmente conocidas como indicadores.

En concreto, un factor (también llamado variable latente, variable subyacente o constructo) es una variable no observable que influye en más de una medida observada y que da cuenta de las correlaciones entre estas medidas observadas. Por ejemplo, el estatus socioeconómico de un individuo no es algo que se le puede preguntar directamente en una encuesta, sino que es un concepto que involucra varios aspectos, como su nivel educativo, su puesto de trabajo, sus ingresos, entre otras variables. En este caso, el factor estatus socioeconómico es una variable no observable que se medie mediante la correlación entre algunas de las variables mencionadas, que sí son variables observadas.

Las observadas están intercorrelacionadas porque comparten un factor causal común (es decir, se encuentran influenciadas por el mismo constructo subyacente, en el ejemplo, el estatus socioeconómico). Por lo tanto, el análisis factorial intenta una modelación más parsimoniosa de la covariación entre un conjunto de indicadores debido a que el número de factores o constructos es menor que el número de variables observadas (es más parsimoniosa, porque es distinto incluir como parámetros independientes aquellos asociados al nivel educativo, el puesto de trabajo y el ingreso, que en este caso serían tres parámetros, mientras que si se incluyera el factor de estatus socioeconómico, que se mide mediante estas tres variables observadas, sólo se estaría incluyendo un parámetro).

Las concepciones iniciales sobre análisis factorial provienen del modelo de factor común (Thurstone, 1947), que postula que cada indicador en un conjunto de medidas observadas es una función lineal de uno o más factores comunes y un factor único. Por lo tanto, el análisis factorial segmenta la varianza de cada indicador en dos partes:

- (1) la varianza común o varianza explicada por el factor latente, que es estimada sobre la base de la varianza compartida con otros indicadores en el análisis; y
- (2) la varianza única, que es una combinación de la varianza que es específica para el indicador y la varianza del error aleatorio (es decir, error de medición en el indicador).

Hay dos tipos principales de análisis basados en el modelo de factor común: el análisis factorial exploratorio (EFA, por sus siglas en inglés) y el análisis factorial confirmatorio (CFA, por sus siglas en inglés; véase Joreskog 1969 y 1971). Tanto el EFA como el CFA pretenden reproducir las relaciones observadas entre un grupo de indicadores con un conjunto más pequeño de variables latentes, pero se diferencian fundamentalmente por el número y la naturaleza de las especificaciones a priori y restricciones hechas en el modelo de factores.

El análisis factorial exploratorio (EFA) es un enfoque impulsado o explorado por los datos, de tal manera que no se hacen especificaciones en relación con el número de factores latentes o al patrón de las relaciones entre los factores comunes y los indicadores (es decir, las cargas factoriales). Más bien, el investigador emplea el EFA como técnica exploratoria o descriptiva para determinar el número adecuado de factores comunes y descubrir cuáles variables de medición son indicadores razonables de las diversas dimensiones latentes.

En el análisis factorial confirmatorio (CFA), el investigador especifica el número de factores y el patrón de relación entre el indicador y las cargas factoriales de antemano, así como otros parámetros. El factor de solución pre-especificado se evalúa en términos de lo bien que este reproduce la matriz de varianzas y covarianzas de la muestra en las variables medidas.

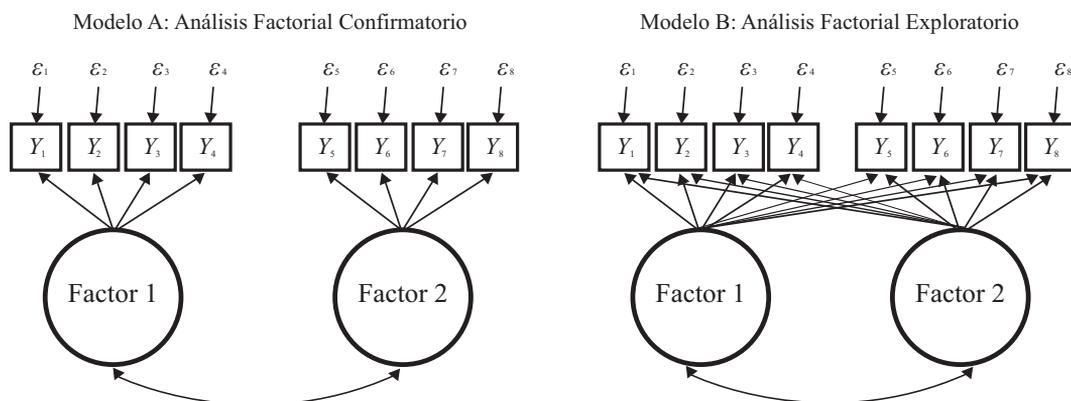
Por lo tanto, a diferencia del EFA, el CFA requiere una sólida base empírica o conceptual para guiar la especificación y evaluación del modelo de factores. En consecuencia, el EFA se utiliza en la etapa previa del proceso de desarrollo de la escala y construcción del constructo, mientras el CFA se utiliza en etapas posteriores, después de que la estructura subyacente ha sido previamente establecida a través bases teóricas y empíricas previas.

Diferencias entre el análisis factorial exploratorio y confirmatorio

Una diferencia fundamental entre ambos enfoques es que, tradicionalmente, en el EFA se estandarizan completamente todas las variables en el análisis (tanto observadas como latentes), para poder utilizar apropiadamente la matriz de correlaciones (es decir, que todas las correlaciones evidenciadas en la matriz sean comparables) y producir soluciones completamente estandarizadas; mientras que en el CFA, la mayor parte del análisis no requiere la estandarización de las variables observadas ni de las latentes, debido a que, en lugar de utilizar la matriz de correlaciones, en el CFA se analiza la matriz de varianzas y covarianzas, por lo que no solo se puede obtener una solución completamente estandarizada al igual que con el EFA (tanto indicadores como variables latentes son estandarizadas), sino también una solución no estandarizada donde la estimación de los parámetros se expresa en la métrica original de los indicadores, así como una solución estandarizada o no completamente estandarizada (indicadores no estandarizados pero variables latentes estandarizadas).

Otra diferencia importante entre ambos es que mientras que en el EFA todos los indicadores cargan libremente en todos los factores latentes y la solución es rotada para maximizar la magnitud de las cargas primarias y minimizar la magnitud de las cargas cruzadas, la rotación factorial no aplica en el CFA, debido a las restricciones identificadas previamente al fijar la mayor parte de las cargas cruzadas como iguales a cero. En otras palabras, la rotación no existe en el CFA porque una estructura simple es obtenida al especificar que los indicadores carguen en un solo factor latente³. Esta situación se ilustra en la siguiente figura 1.

FIGURA 1
ILUSTRACIÓN DE LA DIFERENCIA ENTRE EL ANÁLISIS FACTORIAL CONFIRMATORIO Y EXPLORATORIO PARA UN MODELO DE DOS FACTORES LATENTES



Fuente: Elaboración propia.

En la figura anterior, las 8 variables observadas utilizadas en el modelo de medición (Y1 hasta Y8) cargan libremente en ambos factores (factor 1 y factor 2) en el modelo B que ilustra el análisis factorial exploratorio, mientras que en el modelo A que ilustra al análisis factorial confirmatorio

3 También es permitido que una variable cargue en más de una factor latente en el análisis factorial confirmatorio, pero esto no es tan usual e igualmente requiere de la justificación teórica del porqué se puede dar esa doble-carga. Lo que no puede ocurrir en un CFA es que una variable cargue en todos los factores latentes (como en el caso del EFA).

solo las variables Y1 a Y4 cargan en el factor 1 y solo las variables Y5 a Y8 cargan en el factor 2. Esto significa que las cargas cruzadas de Y1 a Y4 en el factor 2 pueden estimarse libremente en el caso del EFA, mientras que esas mismas cargas cruzadas en el CFA son previamente definidas como iguales a cero (estos se ilustra al no haber flechas que conecten al factor 2 con las variables Y1 a Y4 en el modelo A).

Por esta razón es que el modelo de CFA es más parsimonioso que el modelo EFA: solo hay 8 cargas factoriales a ser estimadas en el modelo de CFA de la figura 1 (Y1 con factor 1, Y2 con factor 1, y así sucesivamente) mientras que hay 16 cargas factoriales a ser estimadas en el modelo de EFA ejemplificado en la figura 1 (Y1 con factor 1, Y1 con factor 2, Y2 con factor 1, Y2 con factor 2, y así sucesivamente). Debido a esto, la rotación no es requerida en el CFA.

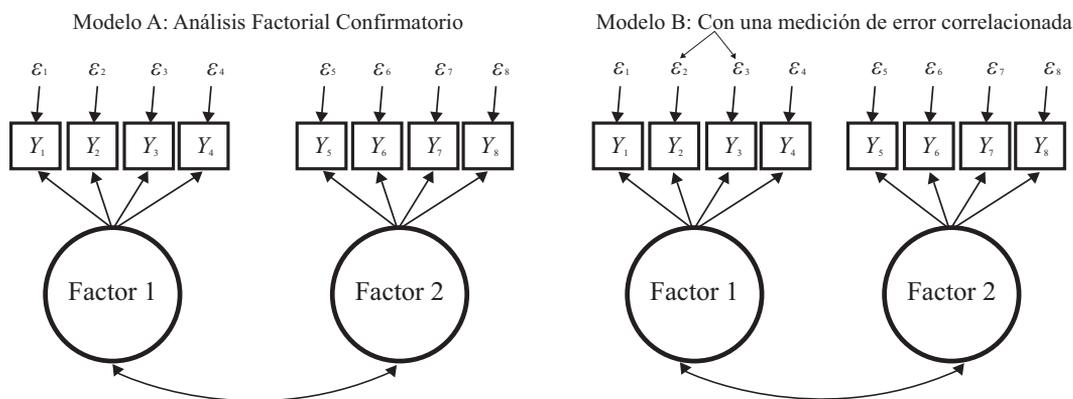
Por otra parte, a diferencia del EFA, el CFA permite al investigador especificar la naturaleza de las relaciones entre los errores de medición (varianzas únicas) de los indicadores.

Debido a las restricciones de identificación del EFA, los modelos factoriales deben ser especificados bajo el supuesto de que el error de medición es aleatorio (no correlacionan unos con otros), mientras que la correlación entre los errores de medición sí puede ser modelada en el CFA (aunque dicha especificación de correlación debe ser justificada sustancialmente en la teoría y evidencia empírica, así como todas las estructuras de un modelo CFA). La figura 2 ilustra lo dicho.

Por otra parte, a diferencia del EFA, el CFA permite al investigador especificar la naturaleza de las relaciones entre los errores de medición (varianzas únicas) de los indicadores.

Debido a las restricciones de identificación del EFA, los modelos factoriales deben ser especificados bajo el supuesto de que el error de medición es aleatorio (no correlacionan unos con otros), mientras que la correlación entre los errores de medición sí puede ser modelada en el CFA (aunque dicha especificación de correlación debe ser justificada sustancialmente en la teoría y evidencia empírica, así como todas las estructuras de un modelo CFA). La figura 2 ilustra lo dicho.

FIGURA 2
ILUSTRACIÓN DE UN MODELO DE ANÁLISIS FACTORIAL CONFIRMATORIO CON ERRORES DE MEDICIÓN CORRELACIONADOS



Fuente: Elaboración propia.

El modelo A de la figura 2 muestra un modelo de dos factores donde todos los errores de medición son modelados como aleatorios. Esto significa que la relación observada entre cualquier par de indicadores que cargan en el mismo factor (por ejemplo Y1 con Y3 en el factor 1 o Y5 con Y8 en el factor 2) se debe por completo a la influencia compartida en la dimensión latente. Por el

contrario, en el modelo B de la figura 2 se define una correlación de errores entre Y2 y Y3. Esta especificación asume que, mientras que Y2 y Y3 están relacionadas en parte por la influencia compartida en la dimensión latente (el Factor 1), una parte de la covariación es debida a otras fuentes además del factor común.

Otra importante diferencia entre EFA y CFA es la forma en que los resultados derivados del análisis factorial pueden ser relacionados con otras variables. En el caso del EFA, hacer esto requiere del cómputo de puntajes factoriales (factor scores) que sirvan como proxy de los factores latentes en análisis subsecuentes. Sin embargo, esta práctica es limitada por la indeterminación de los puntajes; esto es, que para un EFA dado existe un número infinito de sets de puntajes que podrían ser computados y que son igualmente consistentes con las cargas factoriales.

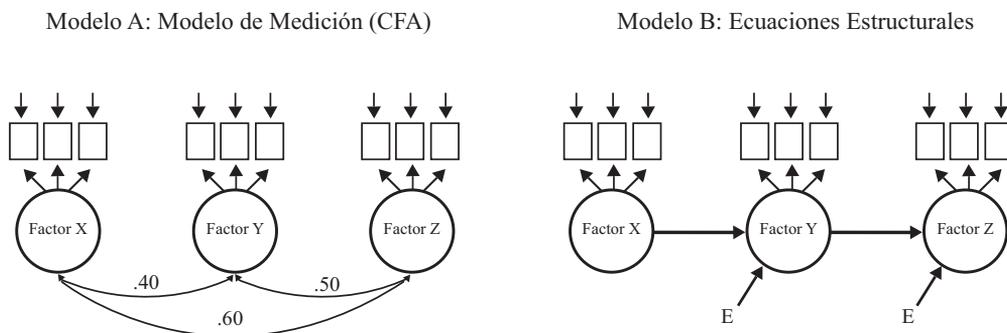
En el caso del CFA, esta indeterminación no es un problema porque la base conceptual y el sustento analítico previo elimina la necesidad de computar puntajes factoriales, es decir, los factores latentes son utilizados ellos mismos (y no puntajes) en los análisis, como en el caso de los modelos de ecuaciones estructurales, o SEM, por sus siglas en inglés.

Frecuentemente, el análisis factorial confirmatorio es utilizado como precursor para los modelos de ecuaciones estructurales (SEM) que especifican relaciones estructurales (por ejemplo, regresiones) entre las variables latentes. Los SEM se pueden separar en dos grandes componentes:

- El modelo de medición, que especifica el número de factores, cómo se relacionan los indicadores con las variables latentes, y las relaciones entre los errores de los indicadores (es decir, el análisis factorial confirmatorio) y,
- El modelo estructural, que especifica cómo se relacionan unas con otras las distintas variables latentes (efectos directos o indirectos, ninguna relación, relaciones espurias).

La figura 3 ilustra un método SEM.

FIGURA 3
ILUSTRACIÓN DE UN MODELO DE MEDICIÓN Y UNO DE ECUACIONES ESTRUCTURALES



Fuente: Elaboración propia.

Así, mientras las relaciones entre las variables latentes en el modelo de medición (modelo A de la figura 3) pueden inter-correlacionar libremente (factor X con factor Z, factor X con factor Y y factor Z con factor Y), en el modelo estructural la naturaleza exacta de dichas relaciones es especificada: en el modelo B de la figura 3 se indica que la relación entre el factor X y el factor Y está completamente mediada por el factor Z: el factor X tiene un efecto directo en factor Z, el factor Z tiene un efecto directo en el factor Y, y el factor X tiene un efecto indirecto en el factor Y.

De esta manera, es claro por qué una solución de un CFA es una antesala muy útil a un modelo de ecuaciones estructurales, que intenta reproducir las relaciones entre variables latentes con un set de parámetros estructurales más parsimonioso (es decir, menos parámetros) que el mismo CFA.

Parámetros de un modelo factorial confirmatorio

Los modelos factoriales confirmatorios analizan la estructura de varianzas y covarianzas de los datos. Por esta razón, los parámetros a estimar (cargas factoriales, varianzas y covarianzas de error, varianzas y covarianzas factoriales) intentan reproducir la matriz de varianzas y covarianzas observada de los datos.

Las variables latentes en CFA pueden ser exógenas o endógenas. Una variable exógena es una variable que no es causada por otras variables en la solución. Una variable endógena es causada por una o más variables en el modelo (es decir, otras variables en la solución ejercen efectos directos en la variable).

De esta manera, las variables exógenas pueden ser vistas como las X, las variables independientes, o variables predictoras (casuales). Las variables endógenas son equivalentes a las Y, variables dependientes o variables resultado. Sin embargo, en el caso de los modelos estructurales, una variable endógena podría ser la causa de otra variable endógena. Por ejemplo, en la figura 3, el factor X es un factor exógeno, mientras que, tanto el factor Z como el factor Y, son variables endógenas.

Para poder realizar un CFA, la variable latente debe tener su escala definida. La variable latente es inobservable y por ende no tiene métrica definida, por lo que la forma más común de darle una escala es fijar la métrica de la variable latente para que sea la misma que uno de sus indicadores, que será llamado indicador de referencia⁴.

Otro aspecto a tomar en cuenta es la identificación del modelo al momento de su especificación. Un modelo está sub-identificado cuando el número de parámetros desconocidos (estimados libremente) excede el número de piezas de información conocida (el número de elementos de la matriz de varianzas y covarianzas). Un modelo sub-identificado no puede ser resuelto, porque hay un infinito número de parámetros estimados que resultan en ajuste perfecto del modelo. En este caso, el número de grados de libertad es negativo.

Se dice que un modelo está identificado si el número de parámetros desconocidos es igual al número de parámetros conocidos, que también muestra perfecto ajuste y 0 grados de libertad.

El objetivo es que el modelo esté sobre identificado. Un modelo está sobre identificado si el número de parámetros desconocidos es menor al número de elementos conocidos. La diferencia entre ambos números de parámetros es por ende positiva y son los grados de libertad. Como recomendación usual, para evitar problemas, un mínimo de tres indicadores por cada variable latente es sugerido.

Aunque en esta sección no han sido explicados en detalle todos los aspectos de los parámetros a estimar y los métodos de estimación, estos serán abordados más adelante en la estimación del modelo que se pretende ajustar con los datos que a continuación se presentarán.

Datos utilizados en el ejemplo empírico

Para ejemplificar la estimación de los modelos de análisis factorial exploratorio y confirmatorio, se utilizarán datos provenientes de una muestra de estudiantes costarricenses participantes en la prueba PISA 2009.

4 Esto aplicada para la solución no estandarizada del modelo, que será explicada más adelante.

El Programa Internacional de Evaluación de Estudiantes (PISA- Programme for International Student Assessment) es un esfuerzo conjunto entre los países miembros de la OCDE que busca medir la capacidad que tienen los estudiantes en edad de 15 años y que se acercan a concluir el ciclo de escolaridad obligatoria, para afrontar los retos de la sociedad actual.

La evaluación tiene una visión a futuro: en lugar de centrarse en evaluar conocimientos específicos de acuerdo con los currículos o planes de estudio de cada centro educativo, se evalúa la capacidad que tienen los estudiantes para utilizar su conocimiento y habilidades para afrontar retos de la vida real.

Las encuestas y evaluaciones de PISA se realizan cada tres años y aunque todas miden habilidades y conocimientos en lectura, ciencias y matemáticas, cada evaluación se enfatiza en una de estas tres áreas, dedicando cerca del 60% de los ítems a la evaluación del área en la que se enfoca la prueba. La primera de estas evaluaciones se aplicó en el año 2000 y tuvo énfasis en lectura; la segunda prueba se aplicó en el año 2003 y se concentró en matemáticas; la tercera ocasión fue en el año 2006 y evaluó habilidades en ciencias. Con estas tres pruebas concluyó el primer ciclo completo de evaluaciones de PISA.

El segundo ciclo de pruebas se inició en el año 2009 con la evaluación que enfatiza lectura. Esta fue la primera evaluación en la que Costa Rica tomó parte dentro del programa de PISA. La segunda prueba en la que participó Costa Rica se realizó en el año 2012 y tuvo énfasis en las habilidades matemáticas, mientras que recientemente en mayo del año 2015 fue aplicada en nuestro país la prueba PISA 2015, cuyo énfasis son las habilidades en ciencias.

Además de la medición de habilidades y conocimientos de los estudiantes mediante una prueba escrita, PISA también aplica cuestionarios a los estudiantes y directores de colegios para recolectar información que permita analizar y contextualizar adecuadamente los resultados de las pruebas.

La muestra de PISA está diseñada sobre una población de estudiantes con edades entre los 15 años y tres meses y 16 años y dos meses que asisten a las instituciones educativas oficiales y se encuentran cursando un grado equivalente al séptimo o superior para todos los países participantes. Por conveniencia, se hace referencia a la población como estudiantes de 15 años.

El diseño de la muestra para estudios educativos normalmente no se basa en un muestreo aleatorio simple sobre la población objetivo. En su lugar, se utiliza un muestreo en dos etapas: en la primera etapa una muestra de centros educativos es seleccionada de una lista completa de los colegios que incluyen a la población de estudiantes de interés. En el caso de PISA, todos los centros educativos son seleccionados con una probabilidad que es proporcional a su tamaño. PISA maneja como estándar una tasa de participación mínima de los centros del 85% para disminuir el sesgo por no respuesta. En el caso de Costa Rica, la tasa de participación efectiva de los colegios muestreados en PISA 2009 fue del 99,43%.

En la segunda etapa, se selecciona aleatoriamente una muestra de estudiantes para los centros educativos elegidos. En los estudios realizados por PISA, se selecciona a un mínimo de 35 estudiantes en edad de 15 años para cada centro educativo participante. En Costa Rica 4.578 estudiantes fueron efectivamente evaluados en la prueba PISA 2009, lo que da una tasa de participación del 94.45% (la tasa mínima requerida es establecida por PISA en 80%).

El cuadro 1 muestra algunas estadísticas de la muestra de estudiantes participantes en PISA 2009. Del total de estudiantes evaluados un 53% son mujeres, y el 85% se encontraban enrolados en colegios públicos. Aunque el nivel oficial que debe cursar un individuo de 15 años en Costa Rica es el noveno año, fueron evaluados estudiantes que se encontraban en casi todos los niveles, pero predominantemente en noveno y décimo años (75%). La proporción de estudiantes cursando el séptimo año se debe, en su mayoría, a estudiantes que perdieron uno o más años en la escuela o en el colegio.

CUADRO 1
 COSTA RICA: ESTADÍSTICAS DESCRIPTIVAS DE LA MUESTRA DE PISA SEGÚN CATEGORÍA. 2009
 (CIFRAS EN PORCENTAJES)

Categoría	Porcentaje	
Género		
Mujeres	53,0	(0,6)
Hombres	47,0	(0,6)
Sector educativo al que asiste		
Público	84,6	(1,4)
Privado	15,4	(1,4)
Año de Colegio que cursa		
Séptimo	8,5	(0,8)
Octavo	16,0	(1,0)
Noveno	34,1	(1,2)
Décimo	40,9	(1,8)
Undécimo	0,4	(0,1)
Duodécimo	0,0	(0,0)

Fuente: Elaboración propia con datos de PISA 2009.
 Error estándar entre paréntesis.

Las bases de datos de una evaluación internacional como la prueba PISA contienen más de 300 variables a disposición (entre las variables observadas medidas directamente en los cuestionarios aplicados a los estudiantes y a los directores de los centros educativos, y las variables latentes que son previamente calculadas por PISA para la creación de índices, como por ejemplo el índice de estatus económico social y cultural). Si bien este número tan elevado de variables son “ideales” para la realización de un análisis factorial exploratorio, donde el investigador puede no tener idea de cómo correlacionan las variables o puede no tener en mente una hipótesis de investigación a ser probada, para el presente trabajo se trabajará únicamente con ocho variables, que serán testeadas en el análisis factorial confirmatorio y que a su vez servirán de ejemplo para ilustrar las diferencias entre ambos enfoques.

Producto de la evidencia empírica previa, se probará la validez de un modelo del rendimiento educativo de los estudiantes en la prueba de lectura de la evaluación PISA 2009. Por lo tanto, la hipótesis de partida es que existen tres constructos latentes inobservables que dan buena cuenta de la explicación de los determinantes de dichos factores latentes en el rendimiento educativo. Estas variables latentes serán las siguientes:

Extracción socioeconómica: En diversos países, ha sido ampliamente probada la hipótesis de que un mayor estatus socioeconómico del estudiante se asocia con un mejor rendimiento educativo. Según las evaluaciones de PISA, en los países donde existe mayor nivel de repetición, el nivel socioeconómico suele ser menor, lo que sugiere que los grupos de menor nivel socioeconómico son más propensos a ser afectados negativamente por el grado de repetición. En muchos países, los centros educativos que compiten por los estudiantes tienden a tener un mayor rendimiento, pero esto es a menudo explicado por el mayor nivel socioeconómico de los estudiantes que asisten a estos centros. Esto sucede porque los padres con un nivel socioeconómico alto, al momento de elegir a

cuál centro enviar a sus hijos, tienden a considerar más el rendimiento académico de los centros educativos que los padres de un nivel socioeconómico bajo (Montero, 2012).

Si bien la base de datos de PISA incluye una variable latente denominada Índice de Estatus Económico, Social y Cultural que fue ampliamente evaluada a la luz de la teoría y de los datos para todos los países en dicha evaluación, dicha variable contiene las puntuaciones factoriales de un análisis factorial exploratorio, por lo que no será ingresada de esta forma en el modelo de medición, sino que se utilizarán sus tres principales variables observadas para la construcción de la variable latente.

Estas tres variables son en realidad índices previamente calculados por PISA: El índice posesiones del hogar (HOMEPOS), el índice de estatus ocupacional más alto de los padres (HISEI), y el índice del nivel educativo más alto de los padres (PARED).

El primero de estos índices, el índice de posesiones familiares o del hogar, es a su vez un índice que incorpora la utilización de otros tres índices: El índice de bienestar familiar (WEALTH) que incorpora dentro de su cálculo artículos como habitación propia, Internet, computadores, televisores, automóviles, entre otros. El índice de posesiones culturales (CULTPOS) que incluye literatura clásica, libros de poesía, obras de arte presentes en el hogar, entre otros y el índice de recursos educativos del hogar (HEDRES) que incluye escritorio y lugar tranquilo para estudiar, computadora con software educativo, libros y diccionarios.

El segundo índice da cuenta del estatus ocupacional más alto de alguno de los padres o tutores de los estudiantes. El estatus ocupacional es medido mediante el Índice ISEI (International Socio-Economic Index of Occupational Status). Este índice varía entre 16 y 90 puntos donde un mayor índice se asocia con un mayor estatus ocupacional.

En el caso del nivel educativo, este es medido mediante años de educación de acuerdo con el ISCED (International Standard Classification of Education), y también da cuenta del nivel educativo más alto de los padres o tutores de los estudiantes.

Para una explicación más detallada sobre el análisis de componentes principales utilizado en el cálculo del índice de Estatus Económico, Social y Cultural, así como del proceso de ponderación utilizado para cada país puede consultar OECD (2012).

Por otra parte., diversos trabajos sugieren que las estrategias de aprendizaje de los estudiantes, su motivación, compromiso y confianza en sí mismos, así como el tiempo dedicado al estudio están relacionados con el éxito académico. De acuerdo con la OECD (2011), los estudiantes con afinidad a la lectura y mejores herramientas que les permiten entender, recordar o resumir información compleja, son quienes obtienen los mejores puntajes en la evaluación de lectura de PISA. Para el estudio de estas relaciones, los expertos de PISA definen dos categorías de análisis: compromiso por la lectura y estrategias de lectura (o enfoques de aprendizaje).

Estrategias de lectura: esta sería una variable latente que mediría las estrategias para estudiar la materia de lectura de los estudiantes. En el cuestionario aplicado a los estudiantes se obtuvo información sobre las estrategias para memorizar, comprender y recordar, resumir, y sobre las técnicas de control y elaboración.

Las estrategias de elaboración se refieren a la capacidad del estudiante para transferir la nueva información proporcionada por la lectura de un texto e incorporarla al stock de información previa que maneja el estudiante para ser utilizada en contextos fuera del centro educativo y en experiencias personales. Las estrategias de control se refieren a las técnicas utilizadas por los estudiantes para auto-controlar sus actividades de lectura.

Con base en una investigación previa que analizó a fondo estos factores mediante un modelo logístico (Fernández, A. 2013) se logró determinar que:

“De los índices donde se evidencia un efecto positivo, la mayoría se relaciona con características o actitudes propias, que desarrolla cada estudiante, en pro de mejorar la calidad de su educación. En especial, el índice de habilidad para resumir presenta el mayor

efecto marginal, de un 76%, lo que significa que un incremento de una desviación estándar en la medida de este índice -para un estudiante desaventajado- producirá un incremento en la razón de probabilidades de un 76%.

Intuitivamente, esto implica que es más probable que un estudiante desaventajado sea resiliente si utiliza técnicas para escribir resúmenes como subrayar oraciones relevantes y principalmente redactar las ideas del resumen en sus propias palabras, en lugar de copiar textualmente algunas oraciones del texto.

El índice de habilidades para comprender y recordar figura como la cuarta variable, y se encuentra muy relacionada, al igual que el índice de habilidades para resumir, con las técnicas en donde se prioriza el resumir ideas y explicarlas en sus propias palabras, en lugar de copiar o recitar una oración textualmente”

Por esta razón, basado en la evidencia empírica previa, para la construcción de la variable latente de estrategias de lectura se utilizará el Índice de habilidades para resumir (METASUM, por su código en la base de datos de PISA) y el Índice de habilidades para comprender y recordar (UNDREM)⁵.

Personal: Para medir el compromiso por las actividades que involucren la lectura, en el cuestionario aplicado a los estudiantes se recopiló información sobre el gusto por leer, el tiempo destinado a dicha afición por disfrute, la diversidad de los materiales leídos, la variedad de las actividades de lectura por medio de Internet.

En la misma investigación (Fernández, A. 2013) se señala que:

“Como segunda y tercera variables con mayor impacto para propiciar la resiliencia, se encuentran el gusto del estudiante por la lectura y su actitud hacia el centro educativo. Ambas están relacionadas con las actitudes personales del alumno, la forma en que valora el proceso educativo, su colegio, la importancia del estudio y en general el gusto por la lectura”.

De esta manera, para la medición de esta variable latente se utilizarán dos variables observadas: su disfrute propio de la lectura por placer (JOYREAD) y la diversidad de lecturas que realiza fuera del centro educativo (DIVREAD).

En lugar de agregar otro indicador relacionado con este constructo de actitud personal hacia la lectura, se incluirá un indicador no relacionado directamente con este constructo, que será la actitud del estudiante hacia su centro educativo (ATSCHL). Note que si bien se trata de la misma idea de actitud personal del estudiante, este último indicador mide otra dimensión, y no la actitud hacia la lectura, sino hacia el proceso educativo como un todo. La intención es mostrar qué ocurre en un CFA cuando se detecta una variable observada que realmente no es una buena medida para el constructo.

En síntesis, las variables latentes que serán utilizadas junto con sus respectivas variables indicadoras se describen en el cuadro 2.

5 Todos los índices calculados por PISA son computados para tener media 0 y desviación estándar de 1.

CUADRO 2
FACTORES LATENTES Y VARIABLES OBSERVADAS DEL MODELO DE TRES FACTORES

Factor latente	Variables observadas	
	Código	Descripción
Extracción Socioeconómica	HOMEPOS	Índice de posesiones materiales del hogar
	HISEI	Índice de estatus ocupacional más alto de los padres
	PARED	Índice del nivel educativo más alto de los padres
Estrategias de lectura	METAUM	Índice de habilidades para resumir
	UNDREM	Índice de habilidades para comprender y recordar
Actitud Personal	JOYREAD	Lectura por placer
	DIVREAD	Diversidad de lecturas realizadas fuera del horario escolar
	ATSCHL	Actitud del Estudiante hacia el centro educativo

Fuente: Elaboración propia.

Método de estimación

Para la estimación de los modelos de factores latentes existen varios métodos como los mínimos cuadrados en dos etapas, mínimos cuadrados no ponderados, mínimos cuadrados generalizados, máxima verosimilitud, máxima verosimilitud robusta, mínimos cuadrados ponderados, entre otros.

El primer método es no-iterativo y sirve para proporcionar las soluciones iniciales de los métodos iterativos, aunque sus soluciones pueden ser tomadas como la estimación final.

Los métodos iterativos proporcionan estimaciones basadas en la minimización de una función de ajuste, la cual modela las discrepancias entre las matrices de variancias y covariancias observadas y estimadas. La función de ajuste del método de máxima verosimilitud (ML) es la más utilizada, y se expresa de la siguiente manera:

$$F_{ML} = \ln|S| - \ln|\Sigma| + tr(S)(\Sigma^{-1}) - p$$

Donde $|S|$ es el determinante de la matriz de variancias y covariancias observada, $|\Sigma|$ es el determinante de la matriz de variancias y covariancias estimada o predicha, p es el orden de la matriz "input" (el número de indicadores o variables manifiestas) y tr es una expresión para representar al operador de la traza una matriz.

El principio subyacente de la estimación vía máxima verosimilitud es encontrar las estimaciones de los parámetros del modelo que maximizan la probabilidad de observar los datos disponibles si los datos hubieran sido recolectados de la misma población en muestreos sucesivos. En otras palabras, la estimación por máxima verosimilitud tiene como propósito encontrar los valores de los parámetros que hacen que los datos observados sean más probables.

Los programas computacionales como LISREL, MPLUS y otros comienzan con un set inicial de parámetros conocidos como los valores iniciales (que pueden ser generados automáticamente por el software) y repetidamente refina esta estimación en un esfuerzo por reducir el valor de F_{ML} de la ecuación anterior (Brown, 2006). La convergencia del modelo se alcanza cuando el programa llega a un set de parámetros estimados que no pueden ser mejorados (es decir, que no pueden reducir F_{ML} más allá de lo ya alcanzado)⁶.

6 Los algoritmos de estimación más utilizados son el Algoritmo Newton-Raphson y los Métodos Cuasi-Newton (Mulaik, 2009).

Una razón por la que la estimación vía máxima verosimilitud es la más utilizada es porque posee propiedades estadísticas deseadas, como la habilidad de proveer errores estándar para cada uno de los parámetros estimados del modelo, que pueden ser utilizados para realizar pruebas de significancia estadística de los parámetros estimados y para determinar la precisión de estas estimaciones. Además, F_{ML} es utilizado en el cálculo de una gran cantidad de índices de bondad de ajuste (Brown, 2006).

De acuerdo con la teoría justificada en los párrafos precedentes, se probará el modelo de medición mediante un análisis factorial confirmatorio, pero es importante recordar que para el análisis factorial exploratorio dicha teoría y el planteamiento de la hipótesis no es requerido, sino que todas las variables son analizadas a la vez en todas las variables latentes, que no necesariamente son los mismos tres constructos que serán probados, sino que existen técnicas para escoger cuántas posibles variables latentes puede generar el análisis factorial exploratorio.

Por último, es importante indicar que para la estimación del análisis factorial exploratorio se utilizará la función *factanal* de R y el módulo de *Análisis factorial exploratorio* del SPSS.

Para el análisis factorial confirmatorio se utilizará como base el LISREL (en inglés LInear Structural RELations) así como la librería *SEM* de R. Todas las estimaciones serán computadas vía máxima verosimilitud, es importante tener presente que existen otros métodos de estimación, como máxima verosimilitud robusta, o más recientemente vía enfoques bayesianos. Todos los códigos o sintaxis se muestran al final de este documento.

III. RESULTADOS

Análisis factorial exploratorio:

Para proceder con el análisis factorial exploratorio, es importante recordar que este tipo de análisis no requiere de una base teórica previa ni de un modelo hipotético a evaluar, sino únicamente tomar los datos y esperar a que ellos hablen por sí solos.

En este caso, al solo utilizar ocho variables se restringe en cierta medida las bondades del EFA, que son más evidentes cuando se utiliza un gran número de variables (más de 20, por ejemplo). Sin embargo, el EFA será computado con estas ocho variables para poder mostrar las diferencias entre el EFA y el CFA.

El cuadro 3 muestra los resultados del análisis factorial exploratorio aplicado con las ocho variables de medición. La primer parte de este cuadro muestra las cargas factoriales rotadas (con rotación varimax) y se encuentran sombreadas en rojo las cargas mayores a 0,55. Se muestran únicamente los primeros cuatro factores, es decir, los factores o variables latentes que de acuerdo con el EFA son los que más aportan en la explicación de la variabilidad de los datos.

CUADRO 3
CARGAS FACTORIALES DEL EFA CON DATOS DE PISA 2009

	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4
HISEI	0,767	-0,016	0,005	0,115
PARED	0,690	0,002	0,005	0,092
HOMEPOS	0,717	0,056	0,024	0,125
METASUM	0,160	0,042	0,021	0,633
UNDREM	0,073	0,069	0,021	0,532
JOYREAD	-0,006	0,345	0,101	0,134
DIVREAD	0,049	0,995	-0,018	-0,054
ATSCHL	0,027	0,127	0,988	0,045
Eigenvalues	1,61	1,14	0,99	0,74
% Varianza individual	20,1	14,2	12,4	9,3
% Varianza acumulada	20,1	34,4	46,8	56,1

Fuente: cálculos propios con datos de PISA 2009.
Rotación Varimax.

Para poder seleccionar un número de factores apropiado para posteriores análisis, existen algunas técnicas descriptivas, como el gráfico de sedimentación o el análisis de los autovalores iniciales (donde se escogen los factores con autovalores iniciales mayores a uno) o el porcentaje de varianza explicado por cada factor (donde se seleccionan por convención aquellos factores con un porcentaje de varianza individual igual o mayor al 10%).

Si se utilizará la técnica de autovalores iniciales, se escogerían solo dos factores, en tanto que de acuerdo con el porcentaje de varianza, se escogerían tres factores.

Las cargas factoriales resaltadas en rojo son las que más tienen peso en la conformación de los factores, y por tanto se intenta relacionar el factor con las variables que tienen cargas importantes. Así por ejemplo, se nota que en el factor 1 tienen más peso las variables HISEI, PARED y HOMEPOS, que coincide exactamente con una de las variables latentes que se evaluarán mediante el CFA, pero nuevamente, como en el EFA en realidad no se dispone de una teoría como la del CFA, el investigador debe analizar cuáles variables tienen más peso en el factor para poder “explicar” qué es lo que este factor está midiendo.

En el factor 2 tiene una carga factorial importante únicamente la variable DIVREAD. Cuando esto sucede, en el EFA se permite que el factor se llame igual que esa variable, por lo que el factor 2 mediante el EFA mediaría la diversidad de lecturas de los estudiantes.

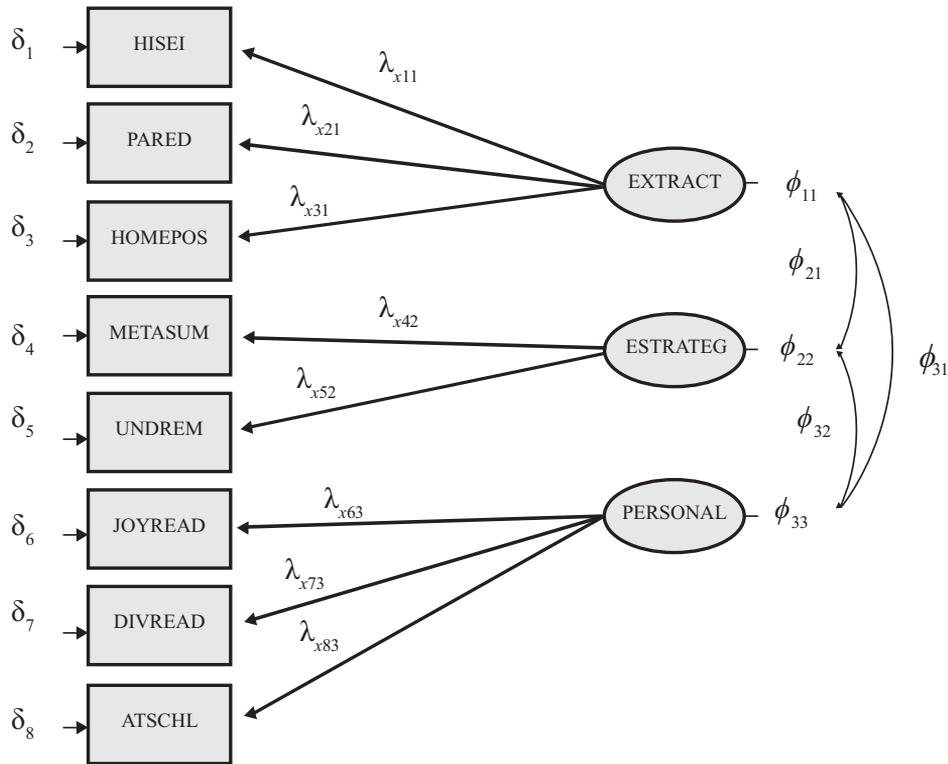
En el factor 3 únicamente tiene un peso importante la variable ATSCHL, de actitud hacia el centro educativo.

Con estos tres factores, se explica el 46,8% de la variabilidad total del modelo, y esta es la intención del EFA: usar el menor número posible de factores que expliquen la mayor variabilidad posible de los datos.

Análisis factorial confirmatorio

En el caso del análisis factorial confirmatorio, el modelo de medición a estimar es el que se muestra en la figura 4.

FIGURA 4
 DIAGRAMA DE TRAYECTORIA PARA UN MODELO DE TRES FACTORES CON ANÁLISIS FACTORIAL CONFIRMATORIO SIN ERRORES DE MEDICIÓN CORRELACIONADOS



Fuente: Elaboración propia con datos de PISA 2009.

De acuerdo con el modelo anterior, se asume (al igual que con el EFA) que todas las medidas de error (δ_i) no están correlacionadas unas con otras. Esto implica que para los indicadores que cargan en la misma variable latente, la covarianza observada entre las mediciones puede ser explicada enteramente por el constructo subyacente. Los datos iniciales o inputs básicos se muestran en el siguiente cuadro.

CUADRO 4
MATRIZ DE CORRELACIONES Y DESVIACIÓN ESTÁNDAR DE LAS VARIABLES OBSERVADAS

	HISEI	PARED	HOMEPOS	METASUM	UNDREM	JOYREAD	DIVREAD	ATSCHL
HISEI	1,0000							
PARED	0,5405	1,0000						
HOMEPOS	0,5632	0,5051	1,0000					
METASUM	0,1929	0,1672	0,2003	1,0000				
UNDREM	0,1134	0,1027	0,1238	0,3514	1,0000			
JOYREAD	0,0268	0,0069	0,0109	0,1003	0,0984	1,0000		
DIVREAD	0,0151	0,0309	0,0836	0,0149	0,0431	0,3337	1,0000	
ATSCHL	0,0285	0,0282	0,0560	0,0586	0,0553	0,1491	0,1074	1,0000
SD	18,4007	4,2404	1,1940	0,9990	0,9587	0,7990	1,0887	1,0329

Fuente: cálculos propios con datos de PISA 2009.

SD: Desviación estándar muestral observada de cada variable.

El cuadro 4 muestra la matriz de correlaciones observadas entre las variables. Es importante indicar que si bien el análisis factorial confirmatorio basa sus cálculos en la matriz de varianzas y covarianzas, esta matriz es inferible desde la matriz de correlaciones y la desviación estándar de los datos, y por esta razón, la desviación estándar muestral observada se incluye en el cuadro 4⁷. Ambos datos (la matriz de correlaciones y la desviación estándar) se indican en la sintaxis de LISREL.

Primero se hará un breve comentario en cuanto a la *identificación* del modelo. Según lo expuesto en secciones anteriores, se busca que el modelo esté sobreidentificado para poder estimarlo. Un modelo está sobreidentificado si el número de piezas de información disponibles es mayor al número de parámetros a estimar.

Comenzando por la primer parte, el número de piezas de información son las entradas en la matriz de correlaciones (covarianzas). Dicha matriz es de tamaño 8x8 (simétrica) y por lo tanto contiene 36 distintas entradas de varianzas y covarianzas ($8*(8+1)/2$).

¿Cuántos parámetros serán estimados? La figura 4 muestra 8 varianzas de error a ser estimadas (δ_i), 8 cargas factoriales (λ_i), 3 covarianzas factoriales (ϕ_{ij}) y tres varianzas factoriales (ϕ_{ii}). Por lo tanto, habría 22 parámetros desconocidos a ser estimados. Sin embargo, como fue mencionado, es requerido brindarle una métrica de medición a las tres variables latentes inobservables. Esto se hace fijando la carga factorial de una variable observada con su variable latente.

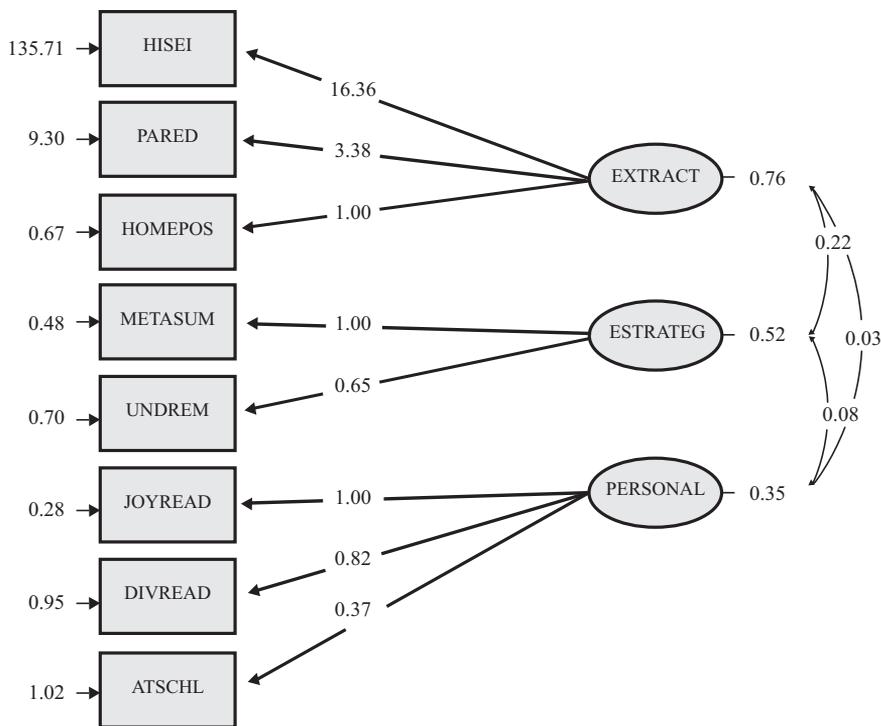
En este caso, λ_{x31} será fijada igual a 1 para el factor 1 (EXTRACT); λ_{x42} será fijada igual a 1 para el factor 2 (ESTRETEG) y, λ_{x63} será fijada igual a 1 para el factor 3 (PERSONAL). De esta manera, el número de parámetros a estimar se reduce a 19 (22-3).

Dado que se tienen 36 piezas de información conocidas y solo se requiere estimar 19 parámetros, el número de grados de libertad es de 17 (36-19) y se dice que el modelo está sobre identificado (tendrá solución única).

7 Para una matriz de correlaciones dada de tamaño $p \times p$ (M) y una matriz diagonal que contiene las p desviaciones estándar (D), la matriz de varianzas y covarianzas (S) se obtiene al pre multiplicar y post multiplicar dicha matriz diagonal por la matriz de correlaciones: $S=D*C*D$.

Al estimar el modelo mediante la sintaxis LISREL que se muestra en el anexo, se obtiene lo siguiente:

FIGURA 5
SOLUCIÓN ESTIMADA NO ESTANDARIZADA DEL MODELO DE TRES FACTORES PARA EXPLICAR EL RENDIMIENTO EN LECTURA DE LOS ESTUDIANTES COSTARRICENSES EN PISA 2009



Chi - Square = 77.03, df = 17, P-value = 0.00000, RMSEA = 0.030.

Fuente: Elaboración propia con datos de PISA 2009.

En la solución que se muestra en la figura 5, se utilizaron todas las variables en su métrica original (es decir, no es la solución estandarizada). Es de notar que, tal y como fue especificado en el modelo, las cargas factoriales que relacionan HOMEPOS con EXTRACT, METASUM con ESTRATEG y JOYREAD con PERSONAL fueron fijadas para ser iguales a 1 (es decir, no eran parámetros a ser estimados).

Para evaluar el ajuste del modelo, existen varias medidas que dan cuenta del ajuste global del mismo. La mayoría se basan en analizar las diferencias entre la matriz de varianzas y covarianzas observada y la matriz de varianzas y covarianzas reproducida por el modelo. La intención es que la matriz estimada sea lo más cercana posible a la matriz observada. Ambas se muestran en el cuadro 5.

CUADRO 5
MATRIZ DE VARIANZAS Y COVARIANZAS OBSERVADA Y ESTIMADA
PARA EL MODELO DE TRES FACTORES

OBSERVADA								
	HISEI	PARED	HOMEPOS	METASUM	UNDREM	JOYREAD	DIVREAD	ATSCHL
HISEI	338,501							
PARED	42,165	17,976						
HOMEPOS	12,370	2,557	1,425					
METASUM	3,544	0,708	0,239	0,998				
UNDREM	2,000	0,417	0,142	0,336	0,919			
JOYREAD	0,395	0,023	0,010	0,080	0,075	0,638		
DIVREAD	0,303	0,143	0,109	0,016	0,045	0,290	1,185	
ATSCHL	0,541	0,123	0,069	0,060	0,055	0,123	0,121	1,067
ESTIMADA								
	HISEI	PARED	HOMEPOS	METASUM	UNDREM	JOYREAD	DIVREAD	ATSCHL
HISEI	338,586							
PARED	41,961	17,981						
HOMEPOS	12,402	2,565	1,426					
METASUM	3,539	0,732	0,216	0,998				
UNDREM	2,309	0,478	0,141	0,337	0,919			
JOYREAD	0,436	0,090	0,027	0,082	0,054	0,638		
DIVREAD	0,356	0,074	0,022	0,067	0,044	0,289	1,185	
ATSCHL	0,160	0,033	0,010	0,030	0,020	0,130	0,106	1,067

Fuente: cálculos propios con datos de PISA 2009.
Estimaciones mediante máxima verosimilitud en LISREL.

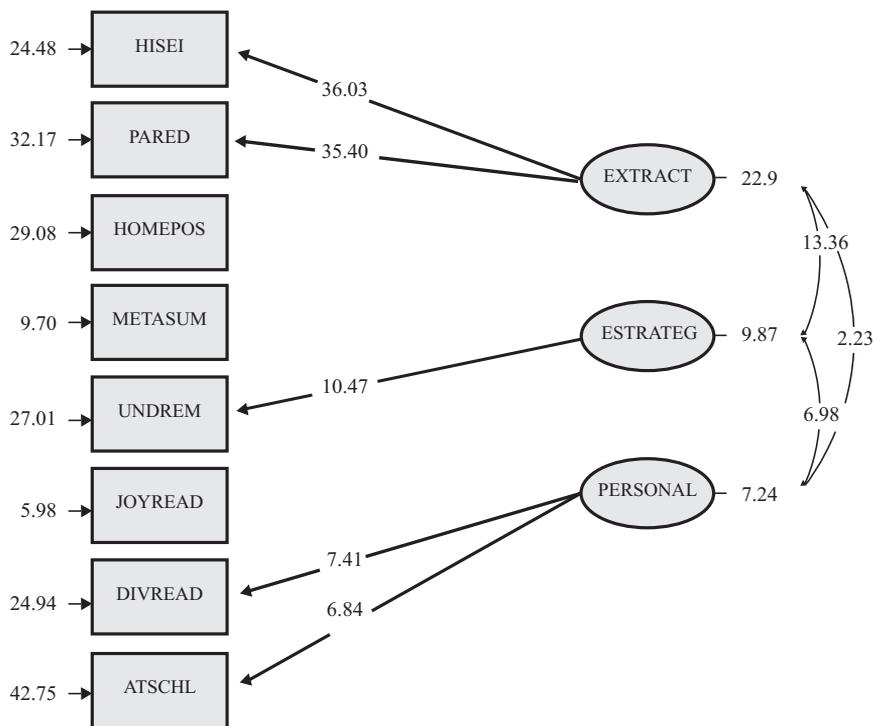
A simple vista ambas podrían parecer muy similares (ver cuadro 5), pero lo mejor es recurrir a algunos índices de bondad de ajuste global del modelo; sin embargo, note de antemano que las mayores diferencias “observadas” entre ambas matrices se dan en la variable ATSCHL (actitud del estudiante hacia el centro educativo) que fue incluida en el constructo PERSONAL con la intención de mostrar qué puede ocurrir cuando un indicador no mide realmente lo que se desea en el constructo.

Uno de los índices de bondad de ajuste es el de la raíz de error cuadrático medio de la aproximación (RMSEA, por sus siglas en inglés). Es un índice que evalúa el “error de aproximación” porque evalúa que tan bien se ajusta el modelo a la población (o qué tan buen ajuste se obtiene dados los valores observados). Por lo tanto, un valor de cero indica un ajuste perfecto, mientras que valores alejados del cero para este índice muestran un ajuste menor. En la práctica, el valor óptimo ideal para la medida RMSEA es 0,05 o menos (Kaplan, 2009; Mulaik, 2009). Según se muestra en la Figura 5, el RMSEA=0,0297 (con un intervalo al 90% de confianza de 0,0231 y 0,0366), que significa un buen ajuste.

Otro índice muy utilizado es el de Ajuste Comparativo (CFI), que evalúa el ajuste del modelo estimado contra una solución restringida o de modelo nulo (típicamente un modelo en el que las covarianzas entre todos los indicadores son fijadas en cero). Un valor cercano a uno de este índice muestra un mejor ajuste. En este caso, este índice fue de 0,9872, que muestra que el modelo ajustado es mejor con respecto a un modelo nulo.

Existen muchas otras medidas de bondad de ajuste global del modelo que no serán expuestas, pero que el lector interesado puede consultar en Brown (2006). Sin embargo, el ajuste global del modelo es solo la primer parte para poder concluir si el modelo realmente cumple con la hipótesis que se está probando. Es importante mirar también el ajuste de los parámetros estimados. La medida típica de ajuste es la prueba z (o t-value) proveniente de la llamada prueba de Wald, que consiste en dividir el coeficiente estimado en la solución completamente estandarizada y dividirlo entre el error estándar de dicho coeficiente. Un valor mayor a 2 (1,96 estrictamente al 95%) en un parámetro específico muestra que el parámetro es realmente distinto del cero (es decir, el parámetro tiene un efecto real sobre la solución). La figura 6 muestra el t-value de cada parámetro estimado.

FIGURA 6
SIGNIFICANCIA ESTADÍSTICA DE LOS PARÁMETROS DE LA SOLUCIÓN ESTIMADA PARA EL MODELO DE TRES FACTORES



Fuente: Elaboración propia según la estimación el LISREL.

Todos los parámetros son estadísticamente significativos de acuerdo con la prueba de Wald; sin embargo, es importante notar que algunos son más pequeños que otros, es decir, existe variabilidad en los t-value que deben ser considerados apropiadamente, debido a que para tamaños de muestras grandes (como es el caso, con 4003 individuos) la prueba de Wald tiende a producir resultados de aceptación del coeficiente en lugar de rechazos.

Por lo tanto, es importante continuar con algunas otras pruebas que evalúen más de cerca la matriz objetivo, que es la matriz de varianzas y covarianzas, para analizar los residuos o diferencias entre los parámetros observados y los estimados.

Mientras que las medidas de bondad de ajuste proveen de una idea general del ajuste del modelo, el análisis de la matriz residual de varianzas y covarianzas estandarizada provee de información específica de cuán bien fueron ajustados los parámetros del modelo para reproducir la matriz inicial (observada).

Debido a la imposibilidad de comparar diferencias en las matrices de varianzas y covarianzas cuando se utilizan los indicadores en su métrica inicial, es que deben utilizarse las matrices estandarizadas, que son calculadas dividiendo los residuos ajustados por los errores estándar de sus estimaciones. De esta manera, los residuales estandarizados son análogos a las puntuaciones estándar en una distribución de muestreo y por lo tanto pueden ser interpretadas en el sentido de las puntuaciones *z*. En otras palabras, los valores pueden ser considerados como el número de desviaciones estándar por el cual los residuos ajustados difieren de los residuos de valor cero que estarían asociados con un modelo de ajuste perfecto.

Los residuos estandarizados pueden tener valores positivos o negativos. Un residuo estandarizado positivo sugiere que los parámetros del modelo subestiman la relación de orden cero entre dos indicadores en algún grado. Para residuos muy grandes, estos pueden indicar que se necesitan parámetros adicionales en el modelo para medir de una mejor manera la covarianza entre los indicadores. Por el contrario, un residuo estandarizado negativo sugiere que los parámetros del modelo sobreestiman la relación entre dos indicadores en cierta medida.

En la práctica, se considera que los residuos estandarizados que son iguales o mayores que 1,96 (en valor absoluto) son problemáticos. Sin embargo, el tamaño de los residuos estandarizados está influenciado por tamaño de la muestra. En general, tamaños de muestra más grandes están asociados con residuos estandarizados más grandes. Por esta razón, algunos investigadores recomiendan el uso de valores de corte más grandes (por ejemplo, 2,58, que corresponde al nivel de alfa de 0,01; véase Byrne, 1998). El cuadro 6 muestra la matriz de residuales estandarizados.

CUADRO 6
MATRIZ DE RESIDUOS ESTANDARIZADOS PARA EL MODELO DE TRES FACTORES

	HISEI	PARED	HOMEPOS	METASUM	UNDREM	JOYREAD	DIVREAD	ATSCHL
HISEI	0,000							
PARED	0,152	0,000						
HOMEPOS	-0,057	-0,087	0,000					
METASUM	0,018	0,037	1,184	0,000				
UNDREM	-1,097	0,011	1,012	0,000	0,000			
JOYREAD	-0,180	-1,327	-1,080	-0,175	2,085	0,000		
DIVREAD	-0,169	0,958	4,313	-3,402	0,065	0,093	0,000	
ATSCHL	1,272	1,309	1,645	-1,948	2,382	-0,495	-0,346	0,000

Fuente: cálculos propios con datos de PISA 2009.
Estimaciones mediante máxima verosimilitud en LISREL.

Según las celdas sombreadas en rojo, hay una subestimación de la relación entre cada par de indicadores involucrado (HOMEPOS con DIVREAD, UNDREM con ATSCHL y UNDREM con JOYREAD). Esto sugiere la necesidad de parámetros adicionales, tal y como podrían ser la inclusión de correlaciones entre errores. Esto puede ser evidencia de errores en el planteamiento de los constructos, posiblemente influenciadas por la inclusión de la variable ATSCHL en un constructo en el

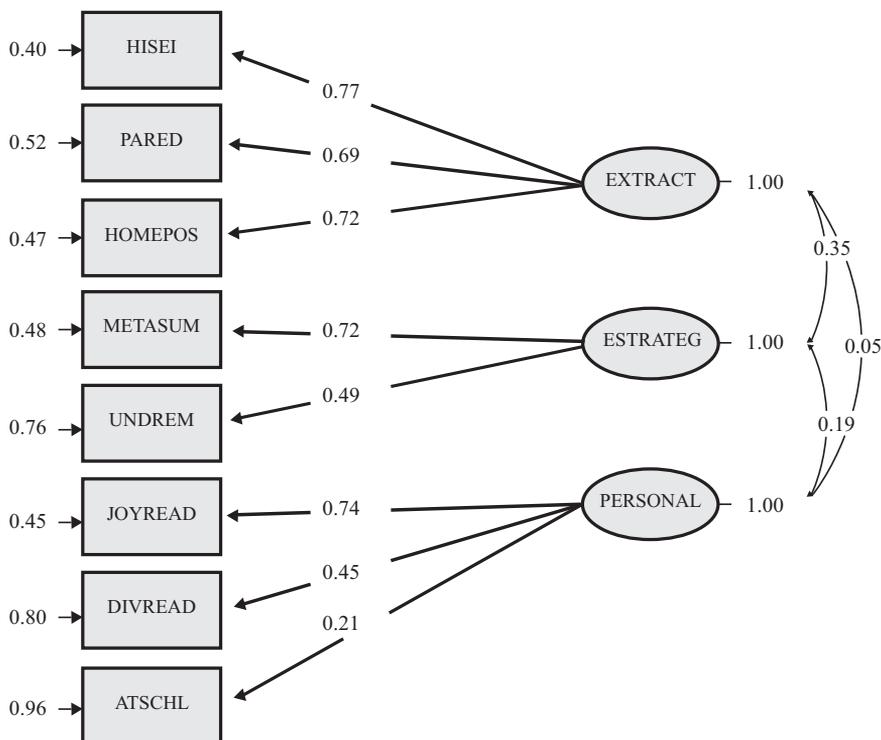
cuál teóricamente no tiene sustento, sino que fue incluido para mostrar posibles patrones de error en el modelo de medición.

Existen algunas formas de corregir estos problemas para intentar obtener un modelo con un mejor ajuste y significancia. De hecho, el mismo programa de LISREL muestra las posibles trayectorias que deberían ser incluidas y que tendrían un efecto de mejora en el modelo; sin embargo, es importante tener presente que el análisis factorial tiene como objetivo probar una hipótesis previamente especificada a la luz de la teoría, el conocimiento y la evidencia empírica previa, por lo que modificar patrones de trayectoria podría no tener un sustento teórico e incluso podría cambiar por completo la hipótesis que quería ser evaluada.

En vista de lo anterior, en este artículo el modelo permanecerá como el anteriormente estimado, y a continuación se compararán estas estimaciones con las obtenidas previamente para el análisis factorial exploratorio.

Para esto, a continuación en la figura 7 se introduce la estimación del mismo análisis de trayectoria pero en su solución estandarizada, ya que esta solución es comparable con la del análisis factorial exploratorio (que sólo muestra soluciones estandarizadas).

FIGURA 7
SOLUCIÓN ESTANDARIZADA ESTIMADA DEL MODELO DE TRES FACTORES



Fuente: Elaboración propia con datos de PISA 2009.

En cuanto al modelo de medición (el sistema que conecta a los constructos con los indicadores propuestos para medir cada uno de ellos) se observa que mayoritariamente todos los indicadores presentan cargas factoriales altas, es decir, los coeficientes que representan el grado de relación entre el constructo con sus respectivos indicadores muestran valores apropiados (cargas mayores a 0,3 en

valor absoluto se consideran óptimas). Sin embargo, el coeficiente que explica la relación entre la actitud del estudiante hacia el centro educativo (ATSCHL) y el factor PERSONAL es muy bajo, lo que evidencia poca relación entre el constructo medido y dicha variable (y que ratifica la noción de que esta variable no mide de manera correcta el constructo de actitud personal hacia la lectura).

En la solución completamente estandarizada, la métrica de ambos, los indicadores y los factores latentes están estandarizadas, por lo que las cargas factoriales pueden ser interpretadas como coeficientes de regresión estandarizadas; por ejemplo, un incremento de un punto estándar en el índice de posesiones del hogar (HOMEPOS) se asocia con un incremento de 0,73 puntos estándar en el factor de extracción socioeconómica.

Además, cuando un modelo de medición no incluye dobles cargas de un mismo indicador (es decir, ninguna variable carga en más de un factor, como en este caso) las cargas factoriales de la solución completamente estandarizada también pueden interpretarse como correlaciones del indicador con el factor latente porque este factor es el único predictor del indicador. De esta manera, el cuadrado de las cargas factoriales provee la proporción de la varianza en el indicador que es explicada por el factor latente. El cuadro 7 muestra estos resultados.

CUADRO 7
CARGAS FACTORIALES DE CADA VARIABLE CON SU FACTOR LATENTE Y PROPORCIÓN DE VARIANZA EXPLICADA POR EL FACTOR LATENTE

Variable	Factor latente	Carga Factorial	Proporción de varianza	Proporción Error
HISEI		0,77	0,60	0,40
PARED	EXTRACT	0,69	0,48	0,52
HOMEPOS		0,73	0,53	0,47
METASUM		0,72	0,52	0,48
UNDREM	ESTRATEG	0,49	0,24	0,76
JOYREAD		0,74	0,55	0,45
DIVREAD	PERSONAL	0,45	0,20	0,80
ATSCHL		0,21	0,04	0,96

Fuente: cálculos propios con datos de PISA 2009.

El cuadro anterior significa que, por ejemplo, el factor EXTRACT da cuenta del 53% de la variabilidad en el indicador de HOMEPOS, de 60% en el indicador de HISEI y de 48% en el indicador de PARED.

La diferencia entre el 100% y este porcentaje es, en consecuencia, la proporción de la varianza no explicada por el factor latente. A este porcentaje se le llama error de varianza único. Por ejemplo, mientras que el factor PERSONAL solo explica un 4% de la variabilidad en ATSCHL, un 96% es varianza única o no explicada por el factor latente.

Por último, y solo para ilustrar mejor las diferencias entre el EFA y el CFA, a continuación se muestran las cargas factoriales obtenidas con cada uno (aunque se debe tomar en cuenta que el factor 2 y el factor 3 del EFA no miden lo mismo que los factores ESTRATEG y PERSONAL en el CFA).

CUADRO 8
DIFERENCIAS ENTRE LAS CARGAS FACTORIALES PARA LOS MODELOS DE ANÁLISIS FACTORIAL
EXPLORATORIO Y CONFIRMATORIO CON DATOS DE PISA 2009

Variable	Análisis Factorial Exploratorio			Análisis Factorial Confirmatorio		
	FACTOR 1	FACTOR 2	FACTOR 3	EXTRACT	ESTRATEG	PERSONAL
HISEI	0,77	-0,02	0,00	0,77		
PARED	0,69	0,00	0,01	0,69		
HOMEPOS	0,72	0,06	0,02	0,73		
METASUM	0,16	0,04	0,02		0,72	
UNDREM	0,07	0,07	0,02		0,49	
JOYREAD	-0,01	0,34	0,10			0,47
DIVREAD	0,05	0,99	-0,02			0,45
ATSCHL	0,03	0,13	0,99			0,21

Fuente: cálculos propios con datos de PISA 2009.

Para el exploratorio se muestran las cargas con Rotación Varimax.

Nuevamente, es importante resaltar que debido a que el análisis factorial exploratorio (EFA) no se realiza sobre una base teórica de las relaciones entre los datos, todas las variables cargan en todos los factores, mientras que con el análisis factorial confirmatorio se busca probar una teoría causal, por lo cual solo las variables evaluadas en cada factor presentan cargas factoriales en esos factores.

IV. CONCLUSIONES

El objetivo de este trabajo consistió en ilustrar de una manera empírica con datos del Programa para la Evaluación Internacional de Estudiantes las diferencias entre el análisis factorial exploratorio y el análisis factorial confirmatorio, poniendo un énfasis en la ventaja del confirmatorio con respecto al exploratorio.

El análisis factorial exploratorio es muy utilizado en estadística porque muchas veces, cuando el investigador se enfrenta a una base de datos, no conoce a fondo lo que estos miden o no tiene claro qué podrían evaluar. Sin embargo, sí se tiene una clara idea de lo que podrían mediar los datos y se plantea una hipótesis de medición, el análisis factorial confirmatorio se puede aplicar a datos no experimentales para validar el poder predictivo de dicha hipótesis.

Como la teoría lo sugiere, los resultados muestran que las variables de nivel educativo máximo de los padres de familia, estatus ocupacional más alto de los padres de familia y de posesiones materiales en el hogar, son variables que en conjunto explican el constructo inobservable de extracción socioeconómica del estudiante.

Los índices de habilidades del estudiante para resumir un texto y de habilidades para comprender y recordar las partes esenciales de un texto dan buena cuenta del factor latente de estrategias de lectura, aunque el índice de habilidades para comprender y recordar muestra una carga factorial más baja que el índice de habilidades para resumir, es decir, la proporción de la varianza que explica el factor latente sobre este indicador, es más baja.

Por último, el índice de disfrute propio del estudiante de la lectura por placer, la diversidad de materiales que lee fuera del sistema educativo y la actitud personal hacia el centro educativo son, en conjunto, indicadores aceptables para medir el constructo de actitud personal del estudiante hacia la lectura. Sin embargo el tercer indicador muestra una carga factorial muy baja, que puede

evidenciar poca relación con este constructo o posibles correlaciones de error con otros indicadores; este índice de actitud personal del estudiante hacia el centro educativo muestra una carga factorial tan baja que la variable latente solo explica el 4% de la varianza en este indicador, por lo que el 96% es varianza única.

A la luz de estos resultados, podría considerarse el variar algunas relaciones de trayectoria entre indicadores y variables latentes, así como definir efectos entre las mismas variables latentes para evaluar un modelo de ecuaciones estructurales sobre el rendimiento en lectura.

V. REFERENCIAS

- Brown, T. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford Press. doi:10.1198/tas.2008.s98.
- Byrne, B. (1998). *Structural equation modeling with ISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Mahwah, NJ: Erl-baum.
- Fernández, A. (2013). Análisis de la Resiliencia Educativa de los Estudiantes Costarricenses con datos de la Prueba de Lectura de la Evaluación PISA 2009. *Revista de Ciencias Económicas*, 31(2), 75–99.
- Joreskog, K. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, 34(2), 183-202. doi: 10.1007/BF02289343.
- Joreskog, K. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika*, 36(2), 109-133. doi: 10.1007/BF02291393.
- Kaplan, D. (2009). *Structural equation modeling: foundations and extensions*. Segunda edición. Thousand Oaks, CA: Sage. doi:10.4135/9781452226576.
- Montero, E. (2012). *Los modelos de ecuaciones estructurales como herramienta para explorar posibles relaciones causales en investigación educativa: Una ilustración con datos de PISA 2009 en Costa Rica*. Aporte Especial. Cuatro Informe del Estado de la Educación. Programa Estado de la Nación en Desarrollo Humano Sostenible.
- Mulaik, S. A. (2009). *Linear causal modeling with structural equations*. New York: CRC Press Taylor & Francis Group. doi: 10.1201/9781439800393.
- Organización para la Cooperación y el Desarrollo Económico. (2011). *Against the Odds: Disadvantaged students who succeed in school*. OECD Publishing. doi:10.1787/9789264090873-en.
- Organización para la Cooperación y el Desarrollo Económico. (2012). *PISA 2009 Technical Report*. OECD Publishing. doi:10.1787/9789264167872-en.
- Spearman, C. (1904). *General Intelligence, objectively determined and measured*. *American Journal of Psychology*, 15, 201-293. doi:10.2307/1412107.
- Spearman, C. (1927). *The abilities of man*. New York: Macmillan.
- Thurstone, L. (1947). *Multiple-factor analysis*. Chicago: University of Chicago. Press.

ANEXOS

ANEXO 1

SINTAXIS DE LISREL PARA LA ESTIMACIÓN DEL MODELO CONFIRMATORIO DE TRES FACTORES.

TITLE MODELO DE TRES FACTORES DE LECTURA

\$Especificar la estructura de mis datos con DA, NI para el número de variables en el archivo de datos analizado. NO para especificar el número de observaciones en el archivo. MA para especificar la matriz a ser analizada. Se especifica como CM (Covariance Matrix)

DA NI=8 NO=4004 MA=CM

\$Especificar las etiquetas de cada variable (LA="labels")

LA

HISEI PARED HOMEPOS METASUM UNDREM JOYREAD DIVREAD ATSCHL

\$Definir la matriz de correlaciones de Pearson (KM)

KM

1.0000

0.5405 1.0000

0.5632 0.5051 1.0000

0.1929 0.1672 0.2003 1.0000

0.1134 0.1027 0.1238 0.3514 1.0000

0.0268 0.0069 0.0109 0.1003 0.0984 1.0000

0.0151 0.0309 0.0836 0.0149 0.0431 0.3337 1.0000

0.0285 0.0282 0.0560 0.0586 0.0553 0.1491 0.1074 1.0000

\$Especificar la desviación estándar muestral de las variables observadas (SD)

SD

18.4007 4.2404 1.1940 0.9990 0.9587 0.7990 1.0887 1.0329

\$Con MO se especifica el modelo LISREL a ser estimado

\$NX para especificar el número de variables X en el modelo (exógenas)

\$NK para especificar el número de variables independientes latentes en el modelo LISREL

\$PH para especificar la forma y modo de la matriz PHI (PHI especifica las varianzas y covarianzas de las variables latentes independientes)

\$SY para decir que la forma es simétrica

\$FR para decir que el modo es libre

\$LX para especificar la forma y el modo de la matriz Lambda_X (Lambda_X especifica la relación lineal entre las variables latentes independientes y sus variables indicadoras)

\$FU para decir que la forma es completo

\$FR para decir que el modo es libre

\$TD para especificar la forma y el modo de la matriz Theta_Delta (Theta_Delta especifica las varianzas y covarianzas de los errores de medición de las indicadoras)

\$SY para decir que la forma es simétrica

\$FR para decir que el modo el libre

MO NX=8 NK=3 PH=SY,FR LX=FU,FR TD=SY,FR

\$Con LK se especifica las etiquetas de las variables latentes independientes, que llamaré EXTRACT (extracción social), ESTRATEG (técnicas y estrategias de lectura) y PERSONAL (factores personales).

LK

EXTRACT ESTRATEG PERSONAL

PA LX

1 0 0

1 0 0

0 0 0

0 0 0

0 1 0

0 0 0

0 0 1

0 0 1

\$Con VA se especifican los valores que serán los elementos fijos de las matrices parámetro del modelo LISREL

VA 1.0 LX(3,1) LX(4,2) LX(6,3)

PA TD

1

0 1

0 0 1

0 0 0 1

0 0 0 0 1

0 0 0 0 0 1

0 0 0 0 0 0 1

0 0 0 0 0 0 0 1

PA PH

1

1 1

1 1 1

\$Con PD se genera automáticamente el gráfico de trayectoria

PD

\$Con OU especifico los métodos a ser usados y los resultados a ser producidos. ME para especificar el método de estimación, en este caso máxima verosimilitud (ML). RS para decir que muestre los residuos, los residuos estandarizados, QQ-PLOT y la matriz de covarianzas ajustada. MI para que muestre el número de valores perdidos. SC para que muestre la solución completamente estandarizada. ND para que use 4 decimales en los resultados.

OU ME=ML RS MI SC ND=4

ANEXO 2

SINTAXIS DE SPSS PARA LA ESTIMACIÓN DEL ANÁLISIS FACTORIAL EXPLORATORIO:

FACTOR

```

/VARIABLES hisei pared homepos metasum undrem joyread divread atschl
/MISSING LISTWISE /ANALYSIS hisei pared homepos metasum undrem joyread
divread atschl
/PRINT UNIVARIATE INITIAL CORRELATION SIG DET KMO INV REPR EXTRACTION
ROTATION
/PLOT EIGEN
/CRITERIA FACTORS(4) ITERATE(25)
/EXTRACTION PC
/CRITERIA ITERATE(100)
/ROTATION VARIMAX
/METHOD=CORRELATION .

```

ANEXO 3

SINTAXIS DE R PARA LA ESTIMACIÓN DEL ANÁLISIS FACTORIAL EXPLORATORIO Y CONFIRMATORIO

```

#####EXPLORATORIO
rm(list=ls(all=T))
getwd()
base=read.table("datos.txt", header=T)
fit=factanal(base,4,rotation="none")
print(fit,digits=2,cutoff=0.0,sort=TRUE)
load=fit$loadings[,1:4]
load
plot(load,type="n")
text(load,labels=names(base),cex=0.7)
fit

```

```

####CONFIRMATORIO
library(sem)
basecov <- cov(base)
model.base <- specifyModel()
F1 -> HISEI, lam1, NA
F1 -> PARED, lam2, NA
F1 -> HOMEPOS, lam3, NA
F2 -> METASUM, lam4, NA
F2 -> UNDREM, lam5, NA
F3 -> JOYREAD, lam6, NA
F3 -> DIVREAD, lam7, NA
F3 -> ATSCHL, lam8, NA
HISEI <-> HISEI, e1, NA
PARED <-> PARED, e2, NA
HOMEPOS <-> HOMEPOS, e3, NA

```

```
METASUM <-> METASUM, e4, NA
UNDREM <-> UNDREM, e5, NA
JOYREAD <-> JOYREAD, e6, NA
DIVREAD <-> DIVREAD, e7, NA
ATSCHL <-> ATSCHL, e8, NA
F1 <-> F1, NA, 1
F2 <-> F2, NA, 1
F3 <-> F3, NA, 1
F1 <-> F2, F1F2, NA
F1 <-> F3, F1F3, NA
F2 <-> F3, F2F3, NA
```

```
basesem <- sem(model.base, basecov, nrow(base))
# Resultados
summary(basesem)
# Coeficientes estandarizados (cargas)
stdCoef(basesem)
```



Todos los derechos reservados. Universidad de Costa Rica. Este artículo se encuentra licenciado con Creative Commons Reconocimiento-NoComercial-SinObraDerivada 3.0 Costa Rica. Para mayor información escribir a revista.iice@ucr.ac.cr