



CUARTO INFORME DEL ESTADO DE LA EDUCACIÓN

Aporte Especial

**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**

Investigadora:

Eiliana Montero

2012



CONSEJO NACIONAL DE RECTORES

Nota: Las cifras de las ponencias pueden no coincidir con las consignadas por el IV Informe Estado de la Educación en el tema respectivo, debido a revisiones posteriores. En caso de encontrarse diferencia entre ambas fuentes, prevalecen las publicadas en el Informe.

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

Los modelos de ecuaciones estructurales o SEM, por sus siglas en inglés (Structural Equations Models) se han hecho populares en los últimos 25 años, en virtud de que representan una alternativa para explorar posibles planteamientos teóricos causales con datos que provienen de estudios observacionales o correlacionales (Bollen, 1989; Freedman, 2010).

En la investigación social y educativa resultan muy útiles, pues con frecuencia es imposible, por razones éticas o prácticas, realizar experimentos, que son los diseños óptimos para generar evidencias de causalidad, gracias a que la asignación de los sujetos a los grupos se realiza al azar.

En los estudios observacionales, por el contrario, se estudia el fenómeno en condiciones “naturales” y por tanto es imposible asegurar el control total de “terceras” variables o variables de confusión que amenazan la validez del argumento causal entre cualesquiera dos variables de interés.

Es en este contexto donde los modelos estructurales (SEM) ofrecen una propuesta estadística poderosa para representar planteamientos teóricos complejos, con múltiples constructos y con diversas hipótesis acerca de las dependencias (o causalidades) entre ellas. Los SEM se han usado en prácticamente todas las disciplinas científicas, aunque su uso está más difundido en las ciencias sociales y del comportamiento (Bollen, 1989; Freedman, 2010; Kaplan, 2009; Mulaik, 2009).

Contar de previo con un marco teórico sólido que sustente las hipótesis que conectan causalmente los diversos constructos involucrados en la explicación del fenómeno de interés es esencial para el empleo de estos modelos, así como una correcta operacionalización y medición de todos los constructos y variables relevantes.

En estudios de factores asociados al rendimiento en pruebas estandarizadas como PISA, los SEM son una vía para generar evidencias empíricas acerca de posibles planteamientos de causalidad entre las diversas variables y constructos de interés.

En este contexto se debe recordar que un modelo es una explicación teórica que describe cómo un sistema de funciones no observable, que produce predicciones de comportamiento, puede ser observado. Así, observaciones empíricas consistentes con lo predicho a partir del modelo dan evidencia a favor de su veracidad, mientras que observaciones no consistentes dan pie para refutarlo o redefinirlo. (Bollen, 1989; Freedman, 2010; Kaplan, 2009).

Otros nombres para SEM incluyen modelos de LISREL (Linear Structural Relations), que es el nombre del paquete de software más antiguo especializado en su estimación, modelación o modelamiento causal y estructuras de covariancia.

Las raíces de los SEM se pueden trazar a partir de otras tres técnicas multivariadas más antiguas, el análisis de factores confirmatorio que proviene de la psicología, el análisis de trayectoria o de senderos (path analysis) que proviene de la biología y la genética, y los sistemas de ecuaciones simultáneas que se desarrollaron en economía. Un modelo SEM completo integra estas tres técnicas (Kaplan, 2009; Mulaik, 2009).

En un modelo SEM se distinguen dos componentes: el modelo estructural, que establece las relaciones causales entre los constructos, y el modelo de medición, que describe con qué indicadores, son medidos los constructos (Bollen, 1989; Kaplan, 2009; Mulaik, 2009).

Un aspecto fundamental de la técnica consiste en que las variables observadas representan un número más pequeño de constructos o variables latentes que no pueden ser directamente medidos, solo inferidos a partir de las variables observadas.

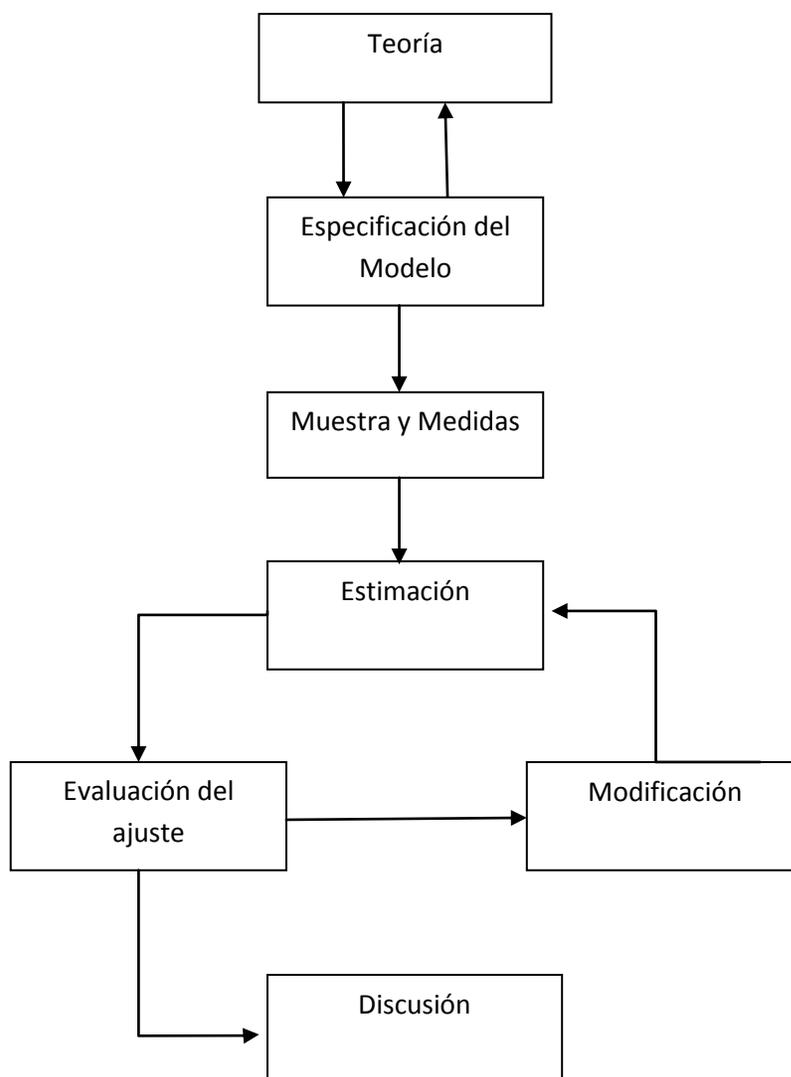
El marco de referencia para la fundamentación de estos modelos plantea que hay conceptos que no se pueden medir directamente -variables latentes o constructos- pero que sí se pueden aproximar mediante indicadores que son operacionalizaciones de esos conceptos, y se pueden establecer relaciones de causalidad entre las variables latentes, las llamadas ecuaciones estructurales.

El hecho de que un modelo SEM sea compatible o consistente con los datos observados, es decir que ajuste estadísticamente a éstos, solo brinda un apoyo empírico a las hipótesis de causalidad implicadas en el modelo, pero no prueba o valida el modelo causal. Se trata entonces fundamentalmente de una técnica de “falsación” en el sentido más clásico, pues las evidencias a favor del modelo no se prueban a partir del mismo modelo, sino eliminando posibles explicaciones alternativas.

Para el autor David Kaplan (2009) la secuencia de pasos involucrados en un estudio con modelos de ecuaciones estructurales se representa en el gráfico 1. Las variables identificadas como relevantes se miden e incorporan a un modelo de relaciones causales hipotéticas. Entonces, los modelos estructurales son una alternativa de análisis estadístico para establecer si datos empíricos que no fueron recolectados usando un diseño experimental, favorecen un cierto planteamiento teórico causal. Otras alternativas de análisis que cumplen propósitos similares son las redes bayesianas y las redes neuronales, aunque escapa al propósito de esta nota tocar estas temáticas.

Gráfico 1

El proceso de investigación que acompaña el uso de modelos de ecuaciones estructurales según David Kaplan



Fuente: Kaplan, D. (2009). Structural equation modeling: foundations and extensions. Segunda edición. Thousand Oaks, CA: Sage.

Es importante insistir en que los modelos SEM no prueban la causalidad, pero sí dan cuenta de evidencia empírica a favor o no de un planteamiento causal. Aunque siempre existirán teorías causales alternativas que explican igualmente bien las relaciones observadas, los SEM nos ayudan a evidenciar empíricamente relaciones causales hipotetizadas, orientadas por la teoría y especificadas a priori. (Bollen, 1989; Freedman, 2010; Kaplan, 2009; Ruiz, Pardo y San Martín, 2010).

Como se dijo antes, es esencial contar un claro planteamiento teórico antes de recolectar los datos y tratar de estimar empíricamente el modelo estructural. El proceso

de estimación en sí mismo presenta frecuentemente problemas técnicos de naturaleza matemática-estadística que dificultan, o incluso, pueden impedir del todo su estimación empírica. Estos problemas tienen que ver con lo que se denomina, en este contexto, el tema de la “identificación” y van desde la situación extrema en donde no es posible estimar del todo la configuración planteada, hasta aquella en donde el modelo se ajusta, pero se generan parámetros con inconsistencias. Tales inconsistencias incluyen coeficientes estandarizados mayores a 1, y el caso en donde no se pueden calcular los errores estándar y por consiguiente tampoco las pruebas de significancia estadística individuales para cada parámetro estimado. Eventualmente se pueden resolver algunos de estos problemas, aplicando estrategias que en su mayoría, implican establecer como constantes o como nulos ciertos parámetros (Kaplan, 2009; Mulaik, 2009).

Una correcta modelación conlleva entonces un proceso iterativo que involucra, por un lado, una apropiada operacionalización de la teoría y, por otro, resolver problemas de estimación matemático-estadísticos (Freedman, 2010).

De hecho, muchos modelos de análisis estadístico más conocidos y antiguos son casos particulares de un modelo SEM, entre ellos: análisis factorial (exploratorio y confirmatorio), modelos de regresión múltiple, y, modelos de análisis de trayectoria o de senderos. En estos últimos se trabaja únicamente con variables observadas, y se supone que están medidas sin error, por tanto no hay modelo de medición y se estima solamente el modelo causal (Kaplan, 2009; Mulaik, 2009; Ruiz, Pardo & San Martín, 2010).

Una de las grandes ventajas a nivel conceptual que representan los modelos de ecuaciones estructurales sobre los modelos de regresión (ya sea clásicos o multinivel), es que por medio de ellos se puede generar evidencia empírica para planteamientos teóricos causales que establecen relaciones de mediación entre variables o constructos. (Kaplan, 2009; Mulaik, 2009; Ruiz, Pardo & San Martín, 2010). Para ilustrar el punto y como se verá más adelante con los datos de la prueba PISA de competencia lectora, la gestión del director no es un predictor directo del rendimiento en la prueba de Competencia Lectora, sino que su influencia sobre esta variable se da por medio de un indicador que se refiere a las dificultades para la enseñanza y clima escolar negativo del colegio. De esta manera, el efecto de la gestión del director sobre el puntaje en la prueba de Competencia Lectora está mediado por esta dimensión de dificultades para la enseñanza y clima escolar negativo. Se trata entonces de un efecto indirecto, que solo se presenta a través de la variable mediadora.

En este contexto se denominan variables exógenas a aquellas que no están determinadas por otras dentro del modelo y variables endógenas a aquellas que están explicadas por otras dentro del modelo. Usualmente se ajustan modelos lineales. Lo ideal para la estimación es que las variables sean métricas y que su distribución sea simétrica (parecida a la normal). El uso operativo de estos modelos involucra al menos cuatro pasos básicos (Bollen, 1989; Kaplan, 2009):

1. Identificación: establecer si el sistema de ecuaciones planteado tiene solución, utilizando los datos recolectados

2. Estimación: el método más popular sigue siendo el máximo verosímil
3. Bondad de ajuste: Se calculan diversas medidas, basadas en el grado de similitud entre la matriz observada de variancias y covariancias y la correspondiente matriz estimada con base en el modelo
4. Pruebas de hipótesis: acerca de los coeficientes del modelo

Finalmente, para emplear estos modelos con propiedad es altamente recomendable tener conocimientos sobre álgebra de matrices, solución de sistemas de ecuaciones lineales y métodos estadísticos de estimación de modelos y bondad de ajuste (Mulaik, 2009).

Solo como un ejercicio ilustrativo para ejemplificar el potencial de este enfoque de análisis de datos en investigación educativa hemos escogido la prueba de Competencia Lectora en PISA y algunas de las variables del estudiante y del centro educativo que resultan más relevantes para su predicción. Con ellas se desarrollaron diversos ensayos para la estimación de modelos estructurales que tuvieran sentido teórico, que pudieran estimarse de manera técnicamente correcta y que además arrojaran niveles relativamente aceptables de ajuste estadístico. Se utilizó el paquete LISREL que es el software especializado en SEM más antiguo y, discutiblemente, uno de los mejores.

Por problemas de estimación matemático-estadísticos fue imposible ajustar un modelo integral y completo que incluyera simultáneamente constructos del estudiante y del centro educativo, con sus respectivos indicadores. Por ello fue necesario estimar de manera separada un modelo estructural con variables latentes para los datos del estudiante y un modelo de trayectoria o senderos (que solo trabaja con variables observadas) para los datos provenientes de la institución

Modelo estructural con las variables del estudiante para predecir Competencia Lectora

Descripción de las variables

Se establecen tres constructos como variables exógenas EXTRACT, ANT_ED y ESTRATE y tres constructos como variables endógenas K_CULTU, LECTURA Y REND. Sus descripciones y correspondientes indicadores se presentan en el Cuadro 1. El Gráfico 2, presenta, por su parte el modelo estructural estimado.

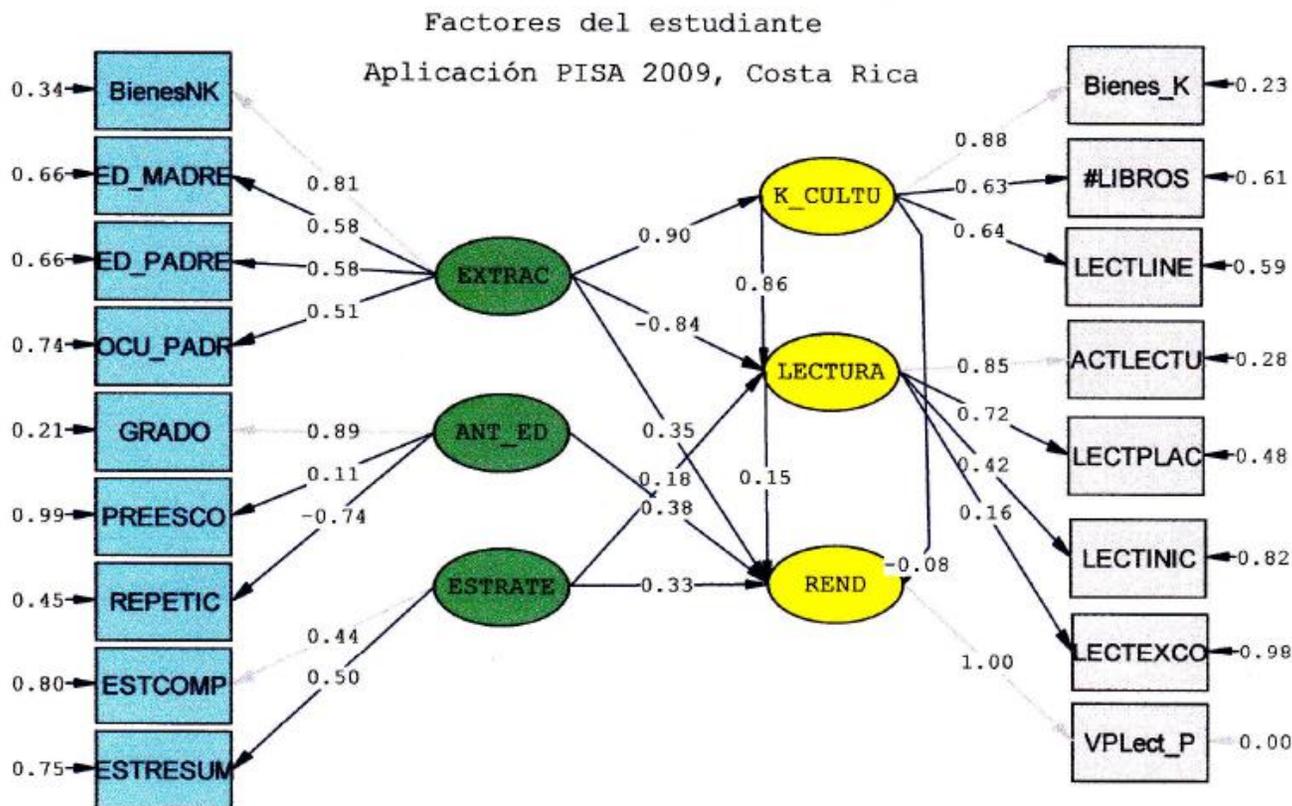
Cuadro 1

Descripción de los constructos e indicadores utilizados en el modelo estructural para predecir el puntaje en la prueba PISA 2009 de Competencia Lectora, Costa Rica

Nombre del constructo	Indicadores	Nombre en la base de datos
EXTRACT (exógeno) Extracción Social	Tenencia de bienes no directamente asociados a c capital cultural	BienesNK
	Nivel educativo de la madre	ED_MADRE
	Nivel educativo del padre	ED_PADRE
	Clasificación de ocupación del padre	OCU_PADR
ANT_ED (exógeno) Antecedentes Educativos	Grado que cursa	GRADO
	Asistió a preescolar	PREESCO
	Frecuencia de repetición	REPETIC
ESTRATE (exógeno) Estrategias de Lectura	Escala de percepción de la eficacia de estrategias para comprender texto	ESTCOMP
	Escala de percepción de la eficacia de estrategias para resumir un texto	ESTRESUM
K_CULTU (endógeno) Capital Cultural	Tenencia de bienes asociados a capital cultural	Bienes_K
	Cantidad de libros en la casa	#LIBROS
	Frecuencia de lectura en línea	LECTLINE
LECTURA (endógeno) Actitudes y hábitos de Lectura	Tiempo de lectura por placer	LECTPLAC
	Frecuencia de lectura por iniciativa propia	LECTINIC
	Frecuencia de lectura de textos para el colegio	LECTEXCO
	Escala de actitudes hacia la lectura	ACTLECTU
REND (endógeno)	Puntaje en la prueba de PISA de Competencia Lectora	VPLect_P

Gráfico 2

Modelo de Ecuaciones Estructurales para predecir el puntaje en Competencia Lectora



Chi-Square=1516.10, df=108, P-value=0.00000, RMSEA=0.063

Interpretación de resultados

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).

Solo en el caso de Asistencia a Pre-escolar como indicador de Antecedentes Educativos se presenta una relación relativamente débil (0.11).

Por otra parte, a algún lector pudiera sorprenderle el signo negativo que muestra para este constructo del indicador de Repitencia, pero dada la naturaleza de esta variable tiene sentido que correlacione negativamente con el constructo (y por ende con los otros dos indicadores), dado que a mayor repitencia menor nivel en términos de Antecedentes Educativos.

En la dimensión de hábitos de Lectura, todos los indicadores presentan valores óptimos, excepto el relacionado con Lectura de Textos del colegio que tiene una relación relativamente débil (0.16).

Es importante notar que el constructo rendimiento (REND) solo cuenta con un indicador, que es precisamente el puntaje en la prueba PISA de Competencia Lectora y por tanto se debe asumir que está medido sin error. Esta es una debilidad metodológica, pues lo ideal sería que cada constructo sea representado por al menos 3 indicadores (Ruiz, Pardo y San Martín, 2010)

En relación al modelo estructural, es decir, las trayectorias causales que conectan variables latentes exógenas y endógenas, vemos que los coeficientes que tienen sentido teórico para el constructo Extracción Social son aquellos que lo conectando causalmente con Capital Cultural y con Rendimiento en la prueba (0.90 y 0.35 respectivamente), mientras que el valor de gran magnitud pero negativo para la relación entre Extracción Social y Actitudes y Hábitos hacia la Lectura merece estudiarse con más detenimiento, pues parece ser contra intuitivo.

La magnitud de la relación entre Antecedentes Educativos y la nota en la prueba PISA de Competencia Lectora parece razonable (0.38), mientras que la relación entre el constructo Estrategias de Lectura (comprender y resumir un texto) y Actitudes y Hábitos hacia la Lectura (0.18) también resulta defendible, al igual que la relación estimada entre estas estrategias y la nota en la prueba (0.33).

Hablando ahora de las relaciones entre las variables endógenas, la fortísima relación entre Capital Cultural y Actitudes y Hábitos hacia la Lectura es también claramente defendible desde la teoría (0.86), al igual que la más débil relación entre Actitudes y Hábitos hacia la Lectura y el rendimiento en la prueba (0.15).

El capital Cultural tiene así un efecto indirecto de importancia sobre el rendimiento en la prueba de PISA de Competencia Lectora, a través de Actitudes y Hábitos de Lectura, el efecto directo no parece tener importancia práctica (-0.08).

En cuanto a la bondad del ajuste de esta primera propuesta de modelo causal tentativo se debe indicar que no es totalmente satisfactoria, dado que el valor óptimo ideal para la medida RMSEA es 0.05 o menos (Kaplan, 2009; Mulaik, 2009), aunque algunos autores son más laxos, sugiriendo el valor de 0.08 como umbral aceptable (Ruiz, Pardo & San Martín, 2010).

A manera de reflexión general, y bajo el entendido de que se trata de resultados tentativos, que deben ser objeto de mayor profundidad en el análisis, parece preocupante que no sean los antecedentes educativos sino la extracción social del estudiante la variable que determina con más fuerza predictiva los resultados de la Prueba de PISA de Competencia Lectora, debido conjuntamente, a su influencia directa (0.35) y su influencia indirecta por medio del Capital Cultural y las Actitudes y Hábitos hacia la Lectura ($0.90 \times 0.86 \times 0.15$). Mientras que, sorprendentemente, los antecedentes educativos no están prediciendo de manera relevante las Actitudes y Hábitos hacia la

Lectura. Este conjunto de resultados parece ser consistente con la segregación en las oportunidades educativas que las crecientes desigualdades sociales y económicas están generando en nuestro país.

Los hallazgos brindan también evidencia de posibles deficiencias en el proceso de la entrega de los programas educativos sobre una dimensión de impacto fundamental, las Actitudes y Hábitos hacia la Lectura. Desde una perspectiva de lo que “debería ser” a nivel educativo, podemos afirmar que en una sociedad menos segregada social y económicamente, la extracción social del estudiante debería tener un peso relativamente menor en el rendimiento en pruebas estandarizadas y en otras variables de impacto educativo. Por el contrario, los antecedentes educativos de los estudiantes (algo que se supone es más fácil de modificar con acciones de política) deberían tener una mayor influencia. Por otra parte, la predicción tan precisa que se puede realizar en la dimensión de Actitudes y Hábitos hacia la Lectura a partir del Capital Cultural del estudiante preocupa también desde una perspectiva de equidad. La relación directa de Capital Cultural hacia el puntaje de PISA no es de relevancia (-0.08), sino que se presenta a través de Actitudes y Hábitos de Lectura.

Por último, es importante recalcar la influencia de las Estrategias de Lectura (para comprender y para resumir) tanto en los Actitudes y Hábitos hacia la Lectura, como en el mismo puntaje de la prueba PISA (coeficientes de 0.18 y 0.33 respectivamente). A partir de estos resultados la recomendación evidente al sistema educativo sería promover la enseñanza explícita de estas estrategias.

Modelo estructural con las variables del colegio para predecir Competencia Lectora

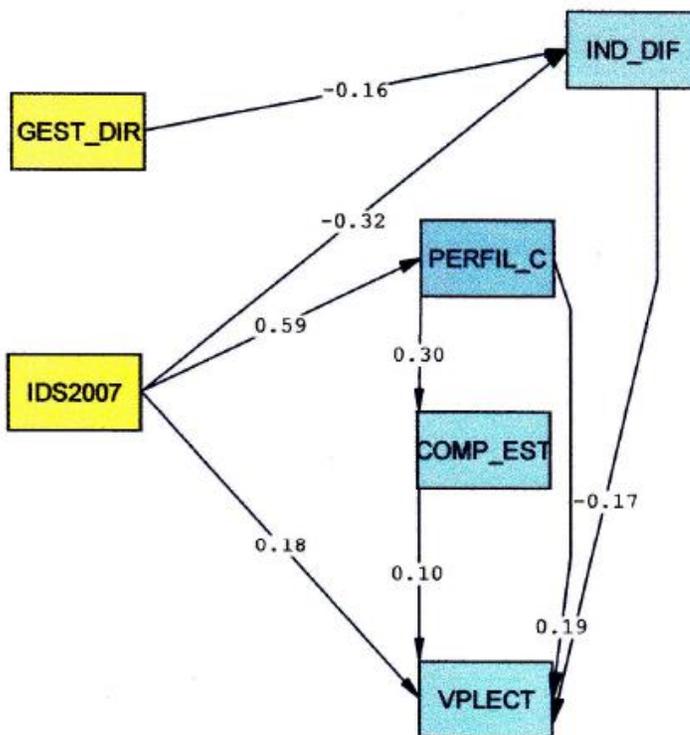
En el caso de los factores institucionales (variables del director y del centro educativo) por problemas de estimación matemático-estadísticos, muy comunes en el contexto de SEM, solamente fue posible ajustar un modelo de análisis de trayectoria o de senderos, en donde no existe modelo de medición y se asume que las variables se miden sin error. Los resultados se presentan en el Gráfico 3.

Gráfico 3

Modelo de Trayectoria para predecir el puntaje en Competencia Lectora

Factores institucionales

Aplicación PISA 2009, Costa Rica



Chi-Square=135.00, df=4, P-value=0.00000, RMSEA=0.085

Como se observa en el Gráfico 3 se plantearon dos variables exógenas, la gestión del director que se mide con una escala, y el IDS (Índice de Desarrollo Social) del distrito en donde está ubicado el colegio. Hay cuatro variables endógenas: un índice de dificultades para la enseñanza y clima escolar negativo, el perfil del colegio que es una escala ordinal compuesta por la combinación de las variables dependencia (público o privado), modalidad (académico o técnico) y zona (urbano o rural), toma el valor más bajo en el caso de público rural académico y el más alto en el caso de privado urbano académico, el promedio de computadoras por estudiante que presenta el colegio y, finalmente, el puntaje en Competencia Lectora de PISA.

Aparte de las dificultades técnicas antes mencionadas para estimar un modelo SEM completo, el modelo de trayectoria finalmente estimado para los factores institucionales dista de presentar niveles óptimos de bondad de ajuste, por tanto los resultados deben considerarse como preliminares y tentativos. Debe indicarse, eso sí, que todos los

coeficientes tienen el signo esperado de acuerdo con la teoría y esto permite hacer una interpretación más directa de los hallazgos.

Nuevamente, desde un enfoque de equidad, resulta preocupante que el IDS se identifique como una variable predictora relevante para 3 de las 4 variables endógenas (índice de dificultades para la enseñanza y clima escolar negativo, perfil del colegio y puntajes en Competencia Lectora). Quizá la forma más injusta de inequidad es esta que relaciona variables educativas individuales del estudiante con el nivel de desarrollo del distrito en que está ubicado su colegio, pues se trata de desigualdades de tipo estructural, que deben enfrentarse también estructuralmente.

Por otra parte, tal como se había mencionado antes, se documenta un efecto indirecto de la gestión del director sobre los puntajes de la prueba de PISA de Competencia Lectora a través del índice de dificultades para la enseñanza y clima escolar negativo. Entonces la recomendación de política educativa es mejorar las destrezas de gestión del director para así disminuir el nivel de problemas para la enseñanza y clima negativo, lo que a su vez incidirá en mejores desempeños en PISA.

Resulta interesante discutir el efecto directo encontrado del número per cápita de computadoras por estudiante sobre el puntaje de la prueba PISA de Competencia Lectora. Quizá se trate de un resultado espurio debido al efecto de otra variable asociada a estas dos pero no controlada, como el acceso a material de lectura que permite una computadora conectada a internet. Parece pertinente, eso sí, recomendar, como medida de política educativa, que se aumente el número de computadoras por estudiante en los colegios, como una forma de coadyuvar a que los estudiantes tengan mayor oportunidad de acceso a materiales de lectura disponibles en internet.

Finalmente, reiteramos que lo ideal, a nivel de análisis, sería integrar simultáneamente en un solo “mega-modelo” estructural, variables del estudiante y variables de la institución, pues la manera en que debimos proceder aquí, estimando cada uno por separado es posible que genere sesgos de estimación en los parámetros, situación que se denomina falacia atomística cuando el modelo solo utiliza variables del estudiante, y falacia ecológica, cuando solo se incluyen variables de la institución (Gaviria, 2005).

Bibliografía

- Bollen, K. A. (1989). *Structural equations with latent variables*. New York, NY: John Wiley & Sons.
- Freedman, D.A. (2010). *Statistical Models and Causal Inference: A dialogue with the Social Sciences*. New York: Cambridge University Press.
- Gaviria, J.L. & Castro M. (2005). *Modelos jerárquicos lineales*. Madrid: La Muralla.
- Kaplan, D. (2009). *Structural equation modeling: foundations and extensions*. Segunda edición. Thousand Oaks, CA: Sage.
- Mulaik, S.A. (2009). *Linear causal modeling with structural equations*. New York: CRC Press Taylor & Francis Group.
- Ruiz, M.A., Pardo, A., San Martín, R. (2010). Modelos de Ecuaciones Estructurales. *Papeles del Psicólogo*, Vol. 31, Núm. 1, enero-abril, 2010, pp. 34-45. España: Consejo General de Colegios Oficiales de Psicólogos.