

LOS MODELOS JERÁRQUICOS LINEALES APLICADOS A LA INVESTIGACIÓN SOBRE EFICACIA ESCOLAR

F. Javier Murillo Torrecilla*

El movimiento de investigación sobre Eficacia Escolar está dirigido al análisis de los factores que hacen que los alumnos de una escuela tengan mejores resultados educativos que los de otra. Desde su aparición hace ahora 30 años, ha tenido lugar una evolución en los modelos y las técnicas de análisis utilizadas, las variables y factores estudiados, los instrumentos de recogida de datos, etc. La principal dificultad metodológica siempre ha radicado en la imposibilidad de utilizar metodologías de carácter experimental con manipulación del tratamiento, en este caso la asistencia a un centro u otro, para comprobar los efectos de un centro y sus variables (Goldstein, 1997). Frente a ello, se han utilizado tanto estudios de carácter extensivo, con muestras formadas por una importante cantidad de alumnos y de centros, como intensivos, especialmente los trabajos sobre centros eficaces e ineficaces (Purkey y Smith, 1983). Aunque los estudios de centros prototípicos jugaron un papel relevante en la década de los 70 y principios de los 80, aportando diferentes «listas de factores» que destronaron la terrible frase «*School doesn't matter*» de Coleman *et al.* (1966) y Jenks *et al.* (1972), parece que la línea de investigación sobre eficacia escolar se decantó por trabajos de carácter cuantitativo.

Los modelos de investigación seguidos en este movimiento fueron evolucionando del famoso y criticado «Input-Output» (p.e. Brookover *et al.*, 1979) al modelo Contexto-Entrada-Proceso-Producto» (p.e. Scheerens, 1992), atravesando los distintos pasos

* (Centro de Investigación y Documentación Educativa (CIDE).
E-mail: javier.murillo@educ.mec.es

intermedios. Las variables analizadas pasaron de ser las variables de centro más fácilmente medibles, como el número de alumnos por aula o el presupuesto, a otras tales como el clima del centro y del aula, la autoestima del alumno, etc.

En cuanto a las técnicas de análisis, desde el informe Coleman, y hasta hace poco, los modelos por excelencia en eficacia escolar han sido los de regresión múltiple. Técnicamente, uno de los problemas de estos modelos es el generalizado incumplimiento de los supuestos (por ejemplo del error). Otro de los principales problemas surgía en el momento de seleccionar la unidad de análisis: el alumno o el centro. Quienes optaban por utilizar el alumno como unidad, estimaban el efecto de la escuela calculando la media de los residuales de los sujetos para cada centro (p.e. Jenks *et al.*, 1972 o Madaus *et al.*, 1980). Si, por el contrario, los investigadores optaban por un modelo de regresión múltiple con los datos agregados a nivel escolar (normalmente a través de las medias de las puntuaciones individuales), estimaban el efecto de la escuela mediante comparaciones realizadas directamente a partir de los residuales, utilizando los datos de los alumnos sólo para calcular las medias de cada escuela (Marks *et al.*, 1983; Prais, 1983). Una alternativa intermedia es el llamado Modelo de Cronbach (Cronbach y Webb, 1975), modelo de regresión donde las variables individuales están centradas en la media de su grupo y las variables de la escuela centradas en la media global.

Otra de las alternativas empleadas es el análisis de covarianza (Rutter *et al.*, 1979). Sin embargo, presenta muchos problemas, como por ejemplo la exigencia de que las unidades de análisis estén asignadas aleatoriamente a los tratamientos, o la necesidad de un alto número de grupos de cada condición, o que sólo muestra la existencia de diferencias, pero no las cuantifica. La alternativa del análisis causal ha sido menos utilizada, ya que tiene dos graves inconvenientes: uno de carácter práctico, que es la gran dificultad para que se ajusten los datos, y otro técnico: no tiene en cuenta la situación de anidamiento de los datos. De hecho, es una alternativa muy poco utilizada en la investigación sobre organizaciones y menos en el campo psicológico.

La gran novedad en la investigación sobre Eficacia Escolar ha sido la aparición de los Modelos Multinivel, desarrollados, en gran medida, por las necesidades metodológicas de la investigación. Para analizar más detenidamente sus aportaciones, vamos a realizar un estudio comparando los modelos clásicos que utilizan el análisis de regresión con los modelos multinivel.

Muy pocos han sido los estudios empíricos sobre Eficacia Escolar llevados a cabo en España. Sobresalen tres: el realizado por Aurora Fuentes (1986) mediante la utilización de análisis causal; el llevado a cabo por el CIDE dentro del marco de una investigación coordinada por la OCDE (Muñoz-Repiso, Cerdán y Murillo, 1995; Murillo, 1996), metodológicamente relacionado con los estudios de escuelas prototípicas; y el coordinado por el profesor Castejón (Castejón, Navas y Sampascual, 1993; Castejón, 1996), que utiliza el análisis de regresión múltiple para estimar diversos índices de eficacia.

A partir de su investigación, Juan Luis Castejón escribió un interesante artículo para la revista RIE (Castejón, 1994) donde analizaba diferentes índices de eficacia. Concretamente comparaba cuatro índices de eficacia, los cuatro a partir del análisis de

regresión múltiple de alumnos y de centros, y ponderados o no. Su artículo concluía que, básicamente, no había diferencia entre los índices. En esta comunicación vamos a realizar un reanálisis de los mismos datos utilizando los modelos jerárquicos lineales o multinivel y, a partir de los resultados, discutiremos acerca de la adecuación de los índices propuestos.

La muestra de datos está formada por 1.924 alumnos de 69 aulas de 1º de BUP de los 24 centros de Bachillerato de la provincia de Alicante. Las variables analizadas son: rendimiento previo del alumno (x_1), inteligencia general (x_2), nivel socioeconómico familiar (x_3), nivel profesional de los padres (x_4) y, como variable de respuesta, las calificaciones finales de 1º de BUP (y) (Castejón, 1994).

El profesor Castejón calcula cuatro índices de eficacia, cada uno de ellos a partir de: a) Puntuaciones residuales a partir de la ecuación de regresión múltiple con datos de los centros obtenidos a partir de las medias de las puntuaciones de los sujetos; b) Puntuaciones residuales a partir de la ecuación de regresión múltiple con datos de los centros obtenidos a partir de la media ponderada de los sujetos; c) Medias de las puntuaciones residuales para cada alumno a partir de la ecuación de regresión; y d) Igual que el anterior, con la media residual estandarizada para cada alumno a partir de la ecuación de regresión.

Técnicamente, los cuatro índices comparados tienen serios problemas debidos a la utilización del análisis de regresión. Aquellos que utilizan las puntuaciones de los sujetos directamente pueden caer en la llamada falacia atomística, por la cual se otorgan las diferencias en las variables de los sujetos a los contextos. Al mismo tiempo, al compartir los sujetos un mismo contexto, se rechaza la hipótesis nula con mucha más frecuencia de lo que sugiere el α nominal. Para los dos índices que trabajan con los datos agrupados se pierde una gran cantidad de información, y disminuye la potencia del análisis estadístico. Además, se puede caer en la llamada falacia ecológica, por la cual se otorgan incorrectamente las características del contexto a los sujetos.

Los modelos multinivel, por su parte, proponen una estructura de análisis dentro de la cual se pueden reconocer los distintos niveles en que se articulan los datos, estando cada subnivel representado por su propio modelo (Draper, 1995). Cada uno de estos submodelos expresa la relación entre las variables dentro de un determinado nivel y especifica cómo las variables de ese nivel influyen en las relaciones que se establecen en otros niveles.

Tras el proceso de modelar los datos y estimar su ajuste, se ha llegado al modelo multinivel siguiente:

1º nivel (alumnos):

$$y_{ijk} = \beta_{0,jk} + \beta_{1,j}x_1 + 0,028x_{2,ijk} + 0,037x_{3,ijk} + 0,051x_{4,ijk} + \varepsilon_{ijk}$$

2º nivel (aulas):

$$\begin{aligned} \beta_{0,jk} &= \beta_{0k} + u_{0,jk} \\ \beta_{1,j} &= 0,906 + u_{1,jk} \end{aligned}$$

3^{er} nivel (centros): $\beta_{ok} = 2,948 + v_{ok}$

$$\begin{aligned} [v_{ok}] &\approx N(0, \Omega_v) : \Omega_v = [0,026] \\ \begin{bmatrix} u_{0,jk} \\ u_{1,jk} \end{bmatrix} &\approx N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0,064 & \\ 0,006 & 0,010 \end{bmatrix} \\ [e_{0ijk}] &\approx N(0, \Omega_e) : \Omega_e = [1,030] \end{aligned}$$

Del análisis de este modelo podemos sacar algunas ideas:

1. Se reconocen tres niveles de análisis: alumno (*i*), aula (*j*) y centro (*k*). Con ello conseguimos, fundamentalmente, que el modelo se ajuste mejor a los datos, pero también estudiamos lo que ocurre en el aula, nivel que está recobrando importancia en los estudios de eficacia escolar (p.e. Creemers, 1994 o Scheerens, 1997). Sin embargo, el centro explica muy poca variabilidad en los resultados de los alumnos; el aula explica algo más, pero no demasiado. Esto significa que el efecto de la escuela, en estos datos, es muy pequeño, lo que puede ser debido a que la variable respuesta es la calificación media obtenida en 1° de BUP, por lo que los centros han tenido poco tiempo para dejar su impronta.
2. El modelo se ajusta bastante bien a los datos, por lo que tanto el rendimiento previo como la inteligencia y el nivel socio-económico y cultural ayudan a explicar parte del rendimiento. Con más variables predictoras, fundamentalmente de aula y de centro, se podría conseguir un ajuste aún más perfecto.
3. Cada centro tiene un punto de corte diferente pero las mismas pendientes. Por tanto, gráficamente se representarían por hiperplanos paralelos. Sin embargo, en el nivel del aula aparece la pendiente del rendimiento previo, lo cual indica que la incidencia del rendimiento previo de los alumnos en sus puntuaciones finales sería diferente en cada aula.

Contrastando este modelo con el obtenido tras un análisis de regresión clásico, en cualquiera de las variedades analizadas por Castejón, hay que destacar que, en primer lugar, el modelo multinivel se ajusta mejor a los datos. Esto se debe, entre otras razones, a que una de las aportaciones de los modelos multinivel es explicar en mayor medida la parte aleatoria del modelo. Se realiza una distinta partición de la varianza, convirtiendo una parte de la misma en variación sistemática y descomponiendo la otra en función de sus fuentes de variación: alumno, aula o centro. De ahí obtenemos otra ventaja: se conoce cuál es el origen de los residuales, con lo que podemos obtener de manera más precisa los residuales de los alumnos, del aula y de los centros. Con ello, las cuatro propuestas que analiza Castejón se ven claramente superadas.

Sin embargo, el mayor problema que tiene la utilización de los residuales para estimar un «índice de eficacia de los centros» es de carácter conceptual. En ellos subyace la idea de que una escuela es «eficaz» cuando la calificación media de los alumnos que finalizan el curso es superior a la prevista, teniendo en cuenta su rendimiento previo, inteligencia y origen socio-económico y cultural. Esta definición tiene

algunos inconvenientes: en primer lugar, la valoración del logro del objetivo de un centro no se mide exclusivamente a través de las calificaciones del alumnado, sino que es necesario considerar otros aspectos de su desarrollo personal y social, cuestión suficientemente reiterada. Igualmente, hay que considerar que un centro puede ser más eficaz en Ciencias Naturales y menos en Matemáticas, diferencias que no contempla este índice homogeneizador. Por último, entendemos que un centro es eficaz si se mantiene así durante varios años y no en un momento coyuntural, por lo que no pueden ser definitivos los resultados tomados en un único momento. Estos problemas, sin embargo, tienen que ver con los datos recogidos y, por tanto, no entran en nuestra consideración ni en la del artículo de Castejón.

Otras dificultades, más interesantes en este caso, tienen que ver con el procedimiento de análisis. Podemos afirmar que un centro no es eficaz si no consigue el desarrollo máximo de *todos* sus alumnos. La equidad es, por tanto, un factor íntimamente ligado a la eficacia; de hecho, no hay eficacia sin equidad. Como se ha visto, con los modelos jerárquicos lineales se elabora un modelo para cada nivel, es decir, para centros y aulas, en los que se recoge no sólo el punto de intersección (los resultados de un hipotético alumno medio de esa escuela o aula), sino cómo varía la pendiente para cada variable predictora; de tal forma, se observa si el centro es igualmente eficaz para todos los alumnos y no hay discriminaciones por sus características personales o sociales. Otra ventaja es la posibilidad de incorporar variables de los distintos niveles que minimicen la parte aleatoria del modelo.

Frente a los modelos clásicos, los modelos multinivel ofrecen a la investigación sobre Eficacia Escolar la posibilidad de poder recoger la estructura anidada de los datos educativos en sus niveles (alumno, aula, centro y contexto), así como poder distinguir con mayor precisión los efectos debidos a cada uno de esos niveles. De la misma forma, y esto es conceptualmente muy destacable, el modelo incluye la posibilidad de reconocer la equidad del centro, elemento fundamental para una correcta concepción de los términos «eficacia escolar» y «calidad de la educación».

REFERENCIAS BIBLIOGRÁFICAS

- Brookover, W.B., Beady, C., Flood, P., Schweitzer, J. y Wisenbaker, J. (1979). *School social systems and student achievement: Schools can make a difference*. New York: Praeger.
- Castejón Costa, J.L. (1994). Estabilidad de diversos índices de eficacia de centros educativos. *Revista de Investigación Educativa*, 24, 45-60.
- Castejón Costa, J.L. (1996). *Determinantes del rendimiento académico y de los centros educativos: factores y modelos*. Alicante: Editorial Club Universitario.
- Castejón, J.L., Navas, L. y Sampascual, G. (1993). Investigación sobre la eficacia de centros de enseñanza secundaria. Un modelo de identificación y funcionamiento. *Revista de Educación*, 301, 221-224.
- Coleman, J.S. et al. (1966). *Equality of educational opportunity*. Washington: US Government Printing Office.
- Creemers, B. (1994). *The effective classroom*. London: Cassell.

- Cronbach, L.J. y Webb, N. (1975). Between class and within class effects in a reported aptitude x treatment interaction: A reanalysis of a study by G.L. Anderson. *Journal of Educational Psychology*, 67, 717-724.
- Draper, D. (1995). Inference and hierarchical modeling in the Social Sciences. *School Effectiveness and School Improvement*, 20, 115-147.
- Fuentes Vicente, A. (1986). *Procesos funcionales y eficacia de la escuela. Un modelo causal*. Madrid: Universidad Complutense.
- Goldstein, H. (1997). Methods in School Effectiveness Research. *School Effectiveness and School Improvement*, 8(4), 369-395.
- Jenks, C.S. (1972). *A reassessment of the effect of family and schooling in America*. New York: Basic Books.
- Madaus, G.F., Afrasian, P.W. y Kellaghan, F. (1980). *School effectiveness: a reassessment of evidence*. New York: Mc Graw-Hill.
- Marks, J., Cox, C. y Pomian-Szrednicki, M. (1983). *Standards in English schools*. London: National Council for Educational Standards.
- Muñoz-Repiso, M., Cerdán, J. y Murillo, F.J. (1995). *Calidad de la educación y eficacia de la escuela. Estudio sobre la gestión de los recursos educativos*. Madrid: CIDE.
- Murillo, F.J. (1996). ¿Son eficaces nuestras escuelas? *Cuadernos de Pedagogía*, 246, 66-72.
- Prais, S. J. (1983). Formal and informal teaching: a further reconsideration of Professor Bennet's statistics. *Journal of the Royal Statistical Society - A*, 146, 163-169.
- Purkey, S.C. y Smith, M.S. (1983). Effective schools: a review. *Elementary School Journal*, 4, 427-452.
- Rutter, M. et al. (1979). *Fifteen thousand hours*. London: Open Books.
- Scheerens, J. (1992). *Effective schooling: research, theory and practice*. London: Cassell.
- Scheerens, J. (1997). Conceptual models and theory-embedded principles on effective schooling. *School Effectiveness and School Improvement*, 8(3), 269-310.