

BORDÓN

Revista de Pedagogía



Volumen 73
Número, 1
2021

SOCIEDAD ESPAÑOLA DE PEDAGOGÍA

ESTUDIO DE LA EFICACIA ESCOLAR EN CENTROS EDUCATIVOS DE PRIMARIA MEDIANTE EL USO DE MODELOS JERÁRQUICOS LINEALES

A Study of school effectiveness in Primary Schools using Hierarchical Linear Models

EMELINA LÓPEZ-GONZÁLEZ(1), ENRIQUE NAVARRO-ASENCIO(2),
MARÍA JOSÉ GARCÍA-SAN PEDRO(3), LUIS LIZASOAIN(4) Y JAVIER TOURÓN(5)

(1) Universidad de Valencia (España)

(2) Universidad Complutense de Madrid (España)

(3) (5) Universidad Internacional de La Rioja (España)

(4) Universidad del País Vasco (España)

DOI: 10.13042/Bordon.2021.80530

Fecha de recepción: 26/05/2020 • Fecha de aceptación: 20/03/2021

Autora de contacto / Corresponding author: Emelina López-González. E-mail: emelina.lopez@uv.es

INTRODUCCIÓN. Se presenta un estudio censal sobre la eficacia escolar en los centros educativos de primaria de la comunidad autónoma de La Rioja, un análisis secundario de una evaluación diagnóstica (ED) previa en esa comunidad. Se valoró el impacto de variables contextuales en los rendimientos de las asignaturas evaluadas no usando las puntuaciones medias brutas obtenidas en la ED, sino las puntuaciones medias ajustadas una vez controlado el efecto de variables contextuales. **MÉTODO.** El control se realizó por modelización estadística multinivel con modelos jerárquicos lineales en dos niveles: estudiantes y centros. Los residuos de los modelos se usaron como criterio de eficacia escolar. **RESULTADOS.** Se realizó un análisis descriptivo por niveles de las variables contextuales y de los rendimientos, además de un análisis detallado de los modelos jerárquicos por asignaturas. A partir de los residuos de los modelos se logró una ordenación de los centros según su eficacia. **DISCUSIÓN.** Los modelos jerárquicos lineales resultaron ser una herramienta adecuada para conocer el impacto de las variables contextuales en el aprendizaje escolar, más allá del rendimiento de los centros aportado previamente en la ED. Los modelos posibilitaron conocer los centros escolares con los mejores y peores resultados de eficacia controlado el efecto del contexto. La estrategia jerárquica detectó que en la relación de centros se encuentran escuelas de ambas redes y de muy diverso nivel socioeconómico, así como que en centros en los que en una asignatura obtienen un residuo muy alto (o muy bajo) no les ocurre lo mismo en otras asignaturas, informaciones imposibles de lograr con herramientas de análisis más convencionales.

Palabras clave: *Eficacia escolar, Evaluación educativa, Efectos de contexto, Modelos jerárquicos lineales.*

Introducción

Las evaluaciones educativas generales realizadas de forma periódica y orientadas a comprobar el funcionamiento de diferentes aspectos del sistema educativo tienen una larga trayectoria en España. El gran impulso a la actividad evaluadora del sistema educativo comenzó en 1990 con la Ley Orgánica de Ordenación General del Sistema Educativo (LOGSE), que creó el Instituto Nacional de Calidad y Evaluación (INCE), actualmente denominado Instituto Nacional de Evaluación Educativa (INEE), que continuó con la Ley Orgánica de Calidad de la Educación (LOCE), de 2002, y la Ley Orgánica de Educación (LOE, 2006). Esta última estableció el marco global de la evaluación general del sistema educativo y de los centros educativos que ha continuado desarrollándose hasta la actualidad. La primera evaluación general del sistema educativo se llevó sobre estudiantes de 14 a 16 años en el año 1995 (INECSE, 1998) y unos años después también se llevó a cabo en educación primaria (Instituto de Evaluación, 2007).

En este contexto se encuentran, por un lado, las evaluaciones generales de diagnóstico, de carácter muestral, que permiten obtener datos representativos tanto del alumnado y de los centros de las comunidades autónomas como del conjunto del Estado. Por otro lado, la LOE también estableció que las comunidades autónomas, dentro de este marco de referencia, pueden realizar en todos los centros una evaluación de diagnóstico, de carácter censal, de las competencias básicas alcanzadas por sus alumnos al finalizar el segundo ciclo de la EP y el segundo curso de la ESO. Dichas evaluaciones tienen un carácter formativo y orientador para los centros e informativo para las familias y para el conjunto de la comunidad educativa.

En este sentido, todas las evaluaciones mencionadas asumen que los estudiantes, las escuelas y, de forma más global, el sistema educativo tienen características específicas que determinan

los resultados educativos. Por tanto, se recomienda que los resultados se comparen utilizando las variables de contexto de los estudiantes y las escuelas (Instituto de Evaluación, 2009) y, de esta forma, estudiar las contribuciones de los centros educativos a los aprendizajes de los estudiantes, es decir, un análisis de la eficacia escolar. En consecuencia, en este trabajo se han optimizado los resultados de una evaluación diagnóstica (ED) realizada en el año 2017 para estudiar la eficacia escolar. Se trata de un estudio censal del curso de 6.º de primaria con los centros educativos de la comunidad autónoma de La Rioja para valorar el impacto de variables contextuales en los rendimientos de las asignaturas evaluadas. El interés estuvo en saber qué factores de contexto, susceptibles de intervención, han impactado en el aprendizaje escolar, logrando establecer qué centros promueven en sus estudiantes rendimientos superiores o inferiores a lo previsible en función de los valores esperables de cada una de las asignaturas.

Los estudios de eficacia escolar comenzaron como respuesta al conocido informe de Coleman de 1966, que evidencia la importancia del contexto socioeconómico en los resultados educativos (Coleman, 1988) por encima de variables escolares. Este informe fue criticado metodológicamente, ya que utilizó para el análisis de los datos la técnica de regresión múltiple paso a paso, introduciendo como primeros predictores las variables del contexto socioeconómico, dejando poca varianza por explicar a las variables escolares. El interés por rebatir los resultados del mencionado informe Coleman dio lugar a numerosos estudios que formaron el movimiento de eficacia escolar.

Las investigaciones de eficacia escolar han ido evolucionando y aumentando su complejidad. Las técnicas estadísticas empleadas para el análisis de la información cambian desde los iniciales estudios correlacionales, como el modelo de los cinco factores de Edmons (1979), hasta la utilización de técnicas estadísticas avanzadas,

como los modelos jerárquicos lineales. Creemers *et al.* (2010) resumen en cuatro fases la evolución de la investigación sobre eficacia escolar:

1. La primera fase, a comienzos de la década de los ochenta, se centró en el análisis del efecto diferencial que profesores o escuelas podían producir en el rendimiento, mostrando que la escuela importa.
2. Los estudios de la segunda fase, a finales de los ochenta y principios de los noventa, trataron de buscar factores relacionados con la eficacia escolar, es decir, características asociadas a un mejor rendimiento académico. El modelo de cinco factores de Edmonds es uno de los ejemplos, y destacaban el liderazgo, las expectativas, la potenciación de las destrezas básicas, un buen clima y una evaluación frecuente del progreso de los estudiantes como elementos clave relacionados con el rendimiento.
3. La tercera fase, a finales de los noventa y principios del nuevo siglo, se centra en la elaboración de modelos de eficacia educativa que apoyen teóricamente por qué determinados factores son importantes en la explicación de la varianza de los resultados de logro de los estudiantes.
4. En la última fase, a partir del año 2000 principalmente, los modelos son más complejos, se centran en el cambio a lo largo del tiempo y analizan aspectos como la consistencia, la estabilidad o la eficacia diferencial. La educación es vista desde una perspectiva dinámica.

Este interés en la comparación de escuelas de forma cuantitativa, iniciado por las investigaciones de eficacia escolar, originó los primeros estudios que, durante la década de los ochenta y principio de los noventa, centran su atención en la elaboración de indicadores de rendimiento escolar utilizando, normalmente, las

puntuaciones medias de los resultados de los estudiantes de un centro educativo. Es decir, para analizar los efectos escolares emplearon las puntuaciones brutas de los test y clasificaron a las escuelas en función de ese rendimiento medio.

Estos trabajos recibieron críticas que apuntaban la necesidad de contextualizar las medidas del rendimiento escolar si se intentaba buscar los efectos que tienen los centros, independientemente de factores que escapan de su control, apareciendo iniciativas tempranas, como la investigación de Aitkin y Longford (1986), que analizaba y discutía cuestiones relativas a la modelización estadística en este tipo de estudios. Posteriormente, Goldstein (1997), en su trabajo sobre los métodos de la investigación en eficacia escolar, y De Maeyer *et al.* (2010), en el suyo sobre los criterios de eficacia usados en modelos multivariados, abordaron también detalladamente esta cuestión.

Estas medidas evolucionaron hacia indicadores contextualizados que tratan de aislar los efectos escolares de factores ajenos a los procesos que se producen en las escuelas, como lo que un estudiante ya conoce o factores socioeconómicos del contexto familiar y escolar. Para más información puede consultarse la revisión sistemática realizada por Murillo (2005), el manual de Townsend y Avalos (2007) o el estado de la cuestión que llevaron a cabo Reynolds *et al.* (2004).

En el presente trabajo no se han considerado las puntuaciones medias brutas obtenidas en la ED previa, sino las puntuaciones medias ajustadas una vez controlado el efecto de variables contextuales, como el nivel socioeconómico y cultural de las familias, si el estudiante proviene de una familia inmigrante, la satisfacción con el colegio, la recomendación del colegio y la ayuda prestada en casa. Se realizó modelización estadística multinivel con modelos jerárquicos lineales (HLM) (Raudenbush y Bryk, 2002; Goldstein, 2011; Snijders y Bosker, 2012),

modelos que trabajan bajo una estructura anidada en niveles. Se consideraron dos niveles: el nivel 1 del estudiante y el nivel 2 del centro, lo que permitió un ajuste más real a la situación de las escuelas.

El enfoque aquí empleado es similar al utilizado en otros trabajos sobre eficacia escolar. El estudio de Marchesi y Martín (2002) usó un modelo de dos niveles (estudiantes y centros) para evaluar la Educación Secundaria Obligatoria en España, o el trabajo de Murillo (2005) con datos de la evaluación general de Educación Primaria de 1995. También López *et al.* (2009) se sitúan en esta perspectiva, analizando datos de España en una evaluación de PISA. En Latinoamérica, Gaviria *et al.* (2004) estudiaron con estructuras anidadas los factores de eficacia escolar en el sistema brasileño, y Cervini (2004) en el sistema argentino. Huffman *et al.* (2006) trabajaron con tres niveles de anidamiento (estudiante, escuela y distrito) y un enfoque metodológico mixto con entrevistas a agentes educativos. En un contexto regional similar a la población aquí estudiada destacan los trabajos de Joaristi *et al.* (2014) y Lizasoain y Angulo (2014), con un modelo de dos niveles (estudiantes y características de los centros).

Así como ya se ha apuntado que existe abundante literatura sobre eficacia escolar, no son tan frecuentes los estudios sobre baja eficacia o ineficacia escolar, aspecto que también se analiza en esta investigación. Un ejemplo son los

trabajos de Hernández-Castilla *et al.* (2013) y los de Van De Grift y Houtveen (2006 y 2007).

Considerando todo lo anterior, el objetivo general de este estudio fue identificar los centros educativos que obtienen mejores y peores resultados de eficacia escolar controlando factores contextuales a partir de la ED de la población de centros de La Rioja. Primero se muestra un estudio descriptivo de las covariables y de las variables criterio empleadas en los modelos anidados, y después se construyen los modelos jerárquicos lineales que posibilitan la ordenación de los centros según su eficacia escolar.

Método

Participantes

El archivo original reunía 2.398 estudiantes agrupados en 82 centros (60% públicos y 40% concertados). Para evitar un posible incremento del error en los modelos multinivel, se estableció un punto de corte de, al menos, 11 estudiantes por escuela, eliminando 14 centros que escolarizan 110 estudiantes, siendo el total de estudiantes analizados de 2.288, agrupados en 68 centros. En la tabla 1 aparecen las frecuencias a ambos niveles según la red de centros. El total de estudiantes es 2.276 porque hubo 12 sin valor en esta variable; el número de centros es 67 (no 68), porque los 12 estudiantes pertenecen al mismo centro.

TABLA 1. Distribución de estudiantes y centros según la red, eliminados los centros con n<11

	Nivel 1 (estudiantes) n (% válido)	Nivel 2 (centros) n (% válido)
Público	1.367 (60.1%)	45 (67.2%)
Concertado	909 (39.9%)	22 (32.8%)
Total	2.276 (100.0%)	67 (100.0%)

Instrumentos y variables

Las variables criterio fueron los rendimientos en las asignaturas consideradas en la ED: Matemáticas, Lengua Española, Inglés y Ciencias y Tecnología. Las covariables eran variables estrictamente contextuales recogidas en la ED, entendiéndose por tales aquellas en las que la escuela no puede intervenir y que, por tanto, se refieren a cuestiones que al centro le vienen dadas, configurando el contexto en el que desarrolla su labor. De ahí el interés por controlar su efecto y lograr un criterio de eficacia equitativo.

Covariables del nivel 1 (de los estudiantes):

- Sexo.
- Índice socioeconómico y cultural familiar (ISEC): elaborado por la Consejería de Educación mediante un análisis de componentes principales de indicadores habituales (nivel de estudios de madres y padres, tipo de trabajo de ambos, bienes culturales, etc.).
- Condición de inmigrante: si el estudiante es inmigrante de reciente incorporación al sistema.
- Satisfacción escolar: índice agregado con la suma de ítems sobre diferentes aspectos de satisfacción de las familias con el colegio: *profesorado; dirección del colegio; trabajo en clase; relación con el/la tutor/a; compañeros de clase; utilización de recursos; nivel de aprendizaje; y satisfacción general con el centro*. Las respuestas se registraron en una escala Likert de cinco cuantificadores lingüísticos.
- Ayuda en casa: variable agregada, promedio de las respuestas de los padres a la pregunta *¿Con qué frecuencia hace usted u otra persona de la casa lo siguiente con su hijo/a?: hablar del trabajo que hace en clase; planificar tiempo para que estudie o haga deberes; ayudarle a hacer los deberes o a estudiar; comprobar que ha estudiado y ha hecho los deberes*. Las respuestas se registraron en una escala Likert de cuatro puntos.

Covariables del nivel 2 (características de centro):

- Red (pública o concertada).
- Tamaño (número de estudiantes en el curso evaluado).
- Promedio del ISEC.
- Proporción de estudiantes inmigrantes de reciente incorporación al sistema.
- Promedio de ayuda en casa.
- Promedio de satisfacción escolar.

Como puede observarse, excepto la red del centro el resto son agregaciones estadísticas por centro de las covariables de nivel 1.

Procedimiento

Los datos se analizaron con modelización multinivel en una estructura anidada en la que las puntuaciones de cada asignatura están asociadas al estudiante, elemento básico del nivel 1, y los estudiantes están agrupados en centros educativos, elementos básicos del nivel 2.

El modelo estadístico en el nivel 1 (estudiantes) fue:

$$Y_{ij} = \beta_{0j} + \sum_{q=1}^Q \beta_{qj} X_{qij} + r_{ij}$$

con parte probabilística definida por $r_{ij} \sim N(0, \sigma^2)$, donde:

Y_{ij} : puntuación obtenida en cada asignatura y su promedio por el estudiante i en el centro j de forma independiente, es decir, no multivariante.

β_{0j} : puntuación media de cada centro j en la asignatura correspondiente.

β_{qj} : influencia lineal de la covariable X_q del alumnado.

X_{qij} : puntuación del estudiante i en el centro j en la covariable X_q .

r_{ij} : residuo del estudiante i de la escuela j ; un indicador de cuánto se separa cada estudiante de lo esperado en su centro.

Como hipótesis se consideró que la influencia de las covariables del nivel 1 es análoga en todos los centros. Así, en el nivel 2 el modelo estructural que corresponde a cada asignatura y que refleja la variación entre los centros fue:

$$\beta_{0j} = \gamma_{00} + \sum_{s=1}^S \gamma_{0s} W_{sj} + u_{0j}$$

con parte probabilística definida por $u_{0j} \sim N(0, \tau_{00})$, donde:

- β_{0j} : rendimiento medio de cada centro j .
- γ_{00} : efecto común a todos los centros.
- γ_{0s} : efecto lineal de la covariable W_s en el rendimiento medio de los centros.
- W_{sj} : valor que toma la escuela j en la covariable de centro W_s .
- u_{0j} : variación residual entre los centros una vez controlados todas las variables individuales y de centro incluidas en el modelo.

Como la influencia de las covariables no difiere significativamente entre los centros, las pendientes β_{qj} no varían, en consecuencia, el coeficiente β_{qj} responde a la expresión: $\beta_{qj} = \gamma_{q0}$.

Cuando las variables contextuales se introducen como covariables de cada nivel, el modelo estima, en cada centro y asignatura, un valor que ha de entenderse como la puntuación media que obtienen los centros educativos que operan en las mismas —o en muy similares— condiciones contextuales, es decir, que obtienen similares valores en dichas variables. La diferencia entre la puntuación obtenida y la estimada es el residuo (u_{0j}), que tomamos como

indicador de la eficacia relativa del centro. Si para un centro el residuo es alto y positivo, la escuela obtiene puntuaciones por encima de las que obtienen el conjunto de centros cuyo contexto es similar al suyo. Se trata de un centro de alta eficacia, pues está funcionando por encima de lo que estadísticamente cabría esperar, habida cuenta de las condiciones contextuales. Si en un centro el residuo es alto y negativo, la escuela obtiene puntuaciones por debajo de lo estadísticamente esperable, por debajo de las que obtienen el conjunto de centros cuyas condiciones contextuales son similares. Se trata, entonces, de un centro de baja eficacia.

Análisis de datos

Se realizó un estudio descriptivo de las covariables y de las distribuciones de las variables criterio. Los resultados se presentan en los dos niveles objeto de estudio. Después se construyeron los modelos jerárquicos lineales y se tomaron los residuos como indicadores del grado de eficacia de cada centro.

Resultados

Estudio descriptivo de las covariables

Índice socioeconómico y cultural familiar (ISEC): se analizó en valores estandarizados. Al nivel de los estudiantes, la distribución del ISEC es asimétrica negativa y ligeramente platicúrtica (tabla 2). Aunque la concentración de casos se sitúa por encima de la media, se extienden hacia los valores negativos (asimetría negativa de $-.731$): algunos estudiantes tienen un ISEC bastante más bajo en relación con el grupo. En los centros la distribución se aproxima a una normal mesocúrtica. La variabilidad de los centros es menor que en el nivel de estudiantes: los valores se distribuyen en niveles superiores e inferiores a la media de manera proporcional.

TABLA 2. Estadísticos descriptivos de las covariables

	Población total N=2288, % (n)		Muestra total N=68
	Nivel de estudiante (N1)		Nivel de centro (N2)
ISEC			
N	1982		68
Asimetría	-.731		-.224
Error estándar de asimetría	.055		.051
Curtosis	-.003		-.266
Error estándar de curtosis	.11		.102
Condición de inmigrante (estudiante)			
N	2276		
No inmigrante	95.0	(2163)	
Público	58.7	(1271)	
Concertado	41.2	(892)	
Inmigrante	5.0	(113)	
Público	85.0	(96)	
Concertado	15.0	(17)	
Satisfacción con el colegio			
N	2277		68
Mínimo	9		29
Máximo	40		36.86
Rango	31		4.86
Media	33.21		33.36
Mediana	33		33.31
Moda	40		29
Desviación típica	5.074		1.625
Coeficiente de variación	0.15		0.048
Recomendación del colegio			
N	97.9	(2239)	68
No	4.7	(106)	
Sí	95.3	(2133)	
Mínimo			.73
Máximo			1.0
Media			.98
Mediana			.97
Moda			1.0
Desviación típica			.051
Ayuda en casa			
N	100	(2288)	68
Mínimo	1.0		2.97
Máximo	9.0		3.78
Media	3.44		3.39
Mediana	3.50		3.40
Moda	4.0		3.54
Desviación típica	.83		.151
Asimetría	2.07		.10
Error típico de asimetría	.051		.29

Condición de inmigrante: se tuvo en cuenta si el estudiante nació en otro país y, además, el padre o la madre también nacieron en otro país. Hay 113 (5%) estudiantes inmigrantes y 2.163 (95%) no inmigrantes (tabla 2). De los 1.367 estudiantes de la red pública, 96 (7%) son inmigrantes; en la red concertada, de 909 hay 17 (1.8%). De los 113 estudiantes inmigrantes, 96 (85%) están en centros públicos y 17 (15%) en concertados.

Satisfacción con el colegio: el valor medio obtenido en el nivel de estudiantes (N1) es 33 (de un mínimo teórico de 8 y un máximo de 40) y coincide con la mediana (tabla 2). La tendencia central es alta: los estudiantes están satisfechos con el colegio, con una distribución asimétrica negativa y bien concentrada, dado que la desviación típica es 5 unidades de 40, con un coeficiente de variación del 15%. La asimetría negativa explica que la media se encuentre disminuida por pocos casos con opiniones muy bajas. En el nivel de los centros (N2), la media y la mediana también coinciden y las opiniones son muy favorables. Aquí las opiniones son, incluso, más homogéneas: el coeficiente de variación apenas refleja un 4.85% de dispersión.

Recomendación del colegio: han respondido 2.239 (97.9%) estudiantes, de los cuales 2.133 (95.3%) sí recomendarían el colegio (tabla 2). Resultados similares se encuentran en el nivel de los centros (N2): lo confirman la media alrededor de 1 (.98), al igual que la mediana (.97) y el bajo valor de la desviación típica (.051).

Ayuda en casa: se trabaja a partir de una variable agregada de cuatro ítems, utilizando el valor medio de las respuestas. La tendencia central se sitúa en algo más de 3, es decir, la ayuda en casa se produce *más de dos veces por semana*, sin llegar a ser *todos los días*. La distribución al nivel de estudiantes es asimétrica negativa, con respuestas relativamente homogéneas: el coeficiente de variación refleja

dispersión moderada (CV=18.6%). Debe tenerse en cuenta también el sentido de la ayuda: puntuaciones elevadas en esta variable implican una mayor ayuda en casa, lo que muy probablemente pueda explicarse por dificultades del estudiante.

Estudio descriptivo de las asignaturas

La tabla 3 recoge los estadísticos descriptivos de las asignaturas evaluadas en los dos niveles. En el nivel de estudiantes se aprecian valores medios similares en las cuatro asignaturas: una media algo mayor en Inglés (510.77) y menor en Ciencias y Tecnología (506.85). La correspondencia entre la media y la mediana en las cuatro asignaturas informa sobre la simetría de las distribuciones. Las asignaturas de Lengua Española e Inglés presentan una cierta asimetría negativa (media inferior a la mediana), lo que implica que la tendencia general de los sujetos es haber obtenido valores altos, pero con algunos casos con valores bajos diferenciados de los demás que hacen disminuir los promedios. Por ejemplo, en Inglés llama la atención el valor mínimo (35.17), cuyo peso en el cálculo de la media habrá sido importante al diferenciarse de forma notable del resto de puntuaciones. En Ciencias y Tecnología y Matemáticas (la media es superior a la mediana), la tendencia de los estudiantes es bastante similar, con algunos casos que aumentan el promedio por tener puntuaciones más elevadas. En cuanto a la dispersión de los estudiantes, las puntuaciones en las cuatro asignaturas son relativamente homogéneas: los coeficientes de variación no superan el 20% y son parecidos. En el nivel de centros las medidas de centralización son similares en las cuatro asignaturas. No se aprecian diferencias en cada asignatura entre media y mediana, lo que aporta información de la normalidad de las distribuciones. Las puntuaciones son homogéneas en las cuatro asignaturas: prácticamente no hay dispersión y los coeficientes de variación son muy bajos.

TABLA 3. Estadísticos descriptivos de las asignaturas según niveles

	Nivel de estudiantes (N1)				Nivel de centros (N2)			
	Matemáticas	Lengua	Inglés	Ciencias y Tecnología	Matemáticas	Lengua	Inglés	Ciencias y Tecnología
N	2288	2288	2288	2288	68	68	68	68
Mínimo	127.48	154.85	35.17	154.01	431.94	427.01	428.11	443.87
Máximo	967.59	949.82	886.64	844.17	650.73	578.35	594.59	563.86
Media	508.33	509.19	510.77	506.86	505.29	504.64	504.38	505.31
Mediana	502.18	512.48	521.35	497.89	506.51	504.17	508.59	504.93
Moda	441.38	535.20	507.59	519.24	431.94	427.01	428.11	443.87
Desv. típ.	100.39	95.83	94.71	96.55	37.54	32.27	39.18	23.39
CV	.20	.19	.19	.19	.07	.06	.08	.05

Nota: CV=coeficiente de variación.

Modelos jerárquicos lineales de las asignaturas

Los modelos jerárquicos analizan cómo varían los resultados de rendimiento en los niveles de anidamiento: se estima la variación explicada por las características de los estudiantes y la variación explicada por los centros. El índice de correlación intraclase (ICC) muestra la proporción entre ambas cantidades (tabla 4): es la parte (porcentaje) de la variación del rendimiento que es debida a los centros.

La información de la tabla 4 está referida a los modelos nulos, sin predictores. La asignatura en la que la escuela tiene un mayor efecto en los resultados es en Inglés, lo que sucede en ella determina el 13.42% de la variación de los resultados. Ese efecto es del 10.88% en Matemáticas, del 7.6% en Lengua y solo del 2.77% en Ciencias y Tecnología.

En las tablas 5 y 6 se presentan los resultados de las asignaturas analizadas con los modelos jerárquicos, incluyendo ahora el efecto de los

TABLA 4. Índice de correlación intraclase (ICC)

Asignatura	Varianza entre estudiantes	Varianza entre centros	Varianza total	ICC
Matemáticas	8991.83	1097.35	10089.18	10.88%
Lengua	8524.38	701.21	9225.59	7.60%
Inglés	7842.97	1215.38	9058.35	13.42%
Ciencias y Tecnología	9063.51	258.02	9321.53	2.77%

Nota: proporción de varianza en N2 con respecto a la varianza total.

predictores de contexto. Esta información se completa con la figura 1 acerca de las relaciones entre las puntuaciones de las asignaturas de los distintos centros y sus ISEC: se visualiza qué centros están rindiendo de forma distinta de lo que cabe esperar por su estatus socioeconómico. Con un punto negro se destacan los centros

con una diferencia (residuo) positiva, es decir, escuelas con un resultado por encima de lo esperado. Con un aspa, centros con residuo negativo y que, por tanto, obtienen resultados inferiores a lo esperable por su nivel socioeconómico. El resto, con un punto hueco, se podría decir que rinden como se espera.

TABLA 5. Estimaciones de los modelos de efectos fijos para las asignaturas

Modelo	Estimación	Error típico	gl	t	p-valor (IC 95%)
Matemáticas N1					
Intersección	517.79	17.20	1847.06	30.11	$p < .000$ (484.06, 551.52)
Sexo N1	21.97	4.12	1935.73	5.34	$p < .000$ (13.90, 30.04)
Satisfac. colegio N1	1.64	.42	1960.11	3.90	$p < .000$ (.81, 2.47)
ISEC N1	30.96	2.26	1957.34	13.72	$p < .000$ (26.54, 35.39)
Ayuda en casa	-20.27	2.71	1942.74	-7.50	$p < .000$ (-25.56, -14.97)
Lengua N1					
Intersección	485.81	24.39	209.74	19.92	$p < .000$ (437.74, 533.89)
Sexo N1	-29.58	4.009	1910.94	-7.39	$p < .000$ (-37.43, -21.73)
Satisfac. colegio N1	1.84	.41	1964.45	4.53	$p < .000$ (1.04, 2.64)
ISEC N1	27.39	2.18	1927.075	12.60	$p < .000$ (23.13, 31.66)
Ayuda en casa	-17.66	2.62	1946.78	-6.75	$p < .000$ (-22.78, -12.53)
Sexo N2	81.23	36.96	72.84	2.20	$p = .030$ (7.58, 154.88)
Inglés N1					
Intersección	527.74	16.14	1662.22	32.70	$p < .000$ (496.08, 599.39)
Red	25.50	8.32	62.28	3.07	$p = .003$ (8.87, 42.12)
Sexo N1	-24.33	3.78	1924.62	-6.43	$p < .000$ (-31.75, -16.91)
Satisfac. colegio N1	1.37	.39	1948.27	3.53	$p < .000$ (.61, 2.12)
ISEC N1	27.52	2.09	1956.96	13.19	$p < .000$ (23.42, 31.61)
Ayuda en casa N1	-17.17	2.48	1931.57	-6.92	$p < .000$ (-22.04, -12.31)
Ciencias y Tecnología N1					
Intersección	512.29	16.96	1861.71	30.20	$p < .000$ (479.02, 545.55)
Sexo N1	14.06	4.17	1959.13	3.37	$p = .001$ (5.88, 22.22)
Satisfac. colegio N1	1.33	.42	1940.16	3.15	$p = .002$ (.50, 2.15)
ISEC N1	23.86	2.27	1689.96	10.52	$p < .000$ (19.41, 28.31)
Inmigrante N1	-36.21	11.28	1965.98	-3.21	$p = .001$ (-58.32, -14.09)
Ayuda en casa N1	-15.15	2.738	1964.99	-5.55	$p < .000$ (-20.51, -9.79)

Nota: IC=intervalo de confianza 95%.

TABLA 6. Explicación de la varianza de los modelos jerárquicos propuestos para las asignaturas

Modelo	Varianza N1	Varianza N2	Total
Matemáticas			
Nulo	8991.83	1097.35	10089.18
Propuesto	8043.77	865.47	8909.24
Varianza explicada	10.54%	21.13%	11.70%
Lengua			
Nulo	8524.38	701.21	9225.59
Propuesto	7555.76	245.17	7800.93
Varianza explicada	11.36%	65.04%	15.44%
Inglés			
Nulo	7842.97	1215.38	9058.35
Propuesto	6763.17	743.51	7506.68
Varianza explicada	13.77%	38.82%	17.13%
Ciencias y Tecnología			
Nulo	9063.51	258.02	9321.53
Propuesto	8358.43	176.40	8534.83
Varianza explicada	7.78%	31.63%	8.44%

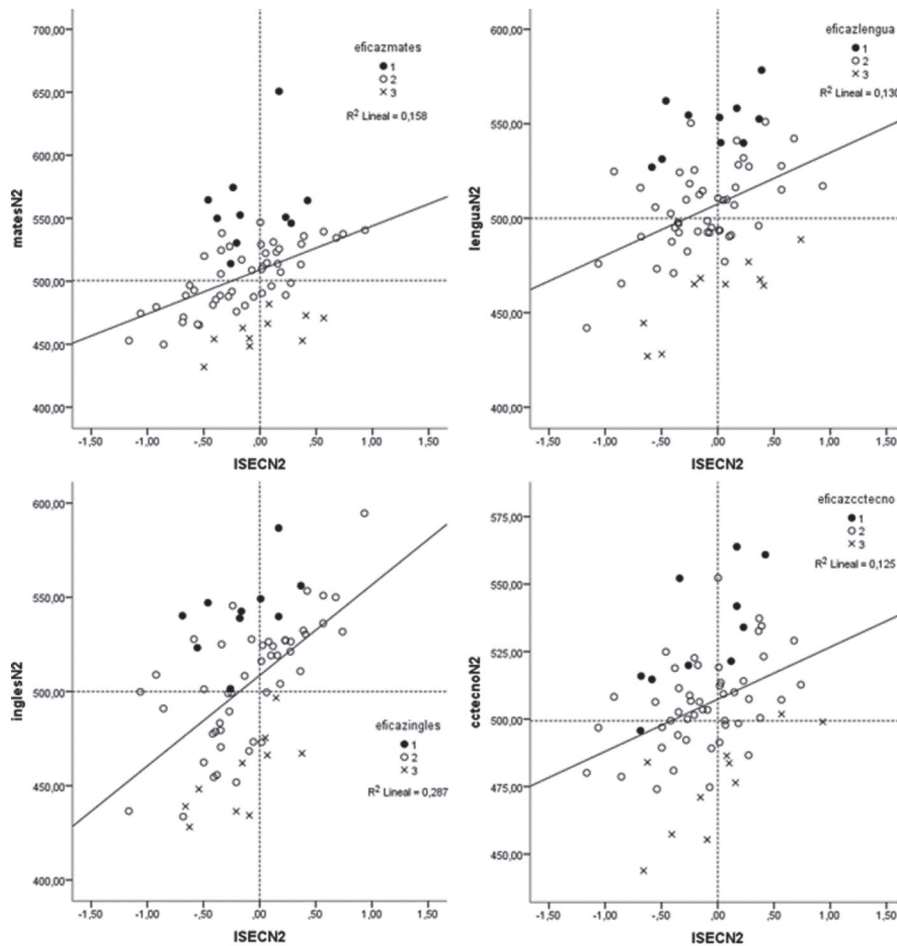
Modelo de eficacia para Matemáticas: el modelo estimado para el rendimiento en Matemáticas indica que las variables con efecto significativo y positivo son el sexo, la satisfacción con el centro y el ISEC y, con efecto negativo, la ayuda en casa (tabla 5).

Los datos muestran que el rendimiento promedio para las chicas con un ISEC medio, que no están satisfechas con su centro y no reciben ayuda en casa, es de 517.8 puntos aproximadamente. Este resultado aumenta unos 22 puntos en el caso de los chicos, y también aumenta 1.6 puntos por cada unidad de aumento en su satisfacción con el colegio. En cambio, los resultados se reducen 20 puntos por cada nivel de aumento de la ayuda en casa. Respecto al impacto del nivel socioeconómico, aquellos que están en una desviación típica por encima de la media del ISEC obtienen unos 31 puntos más

(30.96) en Matemáticas. Conviene destacar que todas las variables que afectan de forma significativa a esta asignatura son características del estudiante y su familia.

En la tabla 6 se indica la cantidad de variación de los resultados que se consigue explicar con la inclusión de los predictores, es decir, la diferencia entre los modelos nulo y propuesto. Estos predictores explican un 11.7% de los resultados de Matemáticas. Si se pone la atención en los niveles, el impacto de los centros en los resultados se reduce en un 21.13% al considerar las características contextuales de los estudiantes. Controlando estas variables se lleva a cabo una estimación más veraz de la aportación de los centros a los resultados en Matemáticas y, en consecuencia, los residuos escolares que se calculan para determinar la eficacia de los centros educativos están libres de los efectos de las

FIGURA 1. Relación entre el ISEC de los centros y sus rendimientos en cada asignatura



características individuales de los estudiantes. Por último, el ISEC de los centros tiene un 15.8% de varianza compartida con el rendimiento en Matemáticas ($R^2=0.158$) (figura 1).

Modelo de eficacia para Lengua Española: los predictores con impacto significativo en los resultados de Lengua son los mismos que para Matemáticas (tabla 5). En esta asignatura se suma al modelo la proporción de chicos que hay en el centro educativo, que también tiene un efecto significativo.

Los resultados promedio en Lengua para las chicas con un ISEC medio, que no están satisfechas con su centro escolar, que no reciben ayuda en casa y cuyos centros son con la totalidad de estudiantes de sexo femenino, es de 485.8 puntos. De forma opuesta a lo que ocurría en Matemáticas, los chicos obtienen unos 29.5 puntos en promedio; no obstante, aquellos centros que están formados únicamente por varones obtendrían 81 puntos más. El resto de predictores (satisfacción, nivel socioeconómico y ayuda en casa) mantienen el sentido e importancia de su impacto.

Tal y como indica la tabla 6, la inclusión de estos predictores consigue explicar el 15.44% de la variación de los resultados de Lengua. En el caso de los centros educativos, los predictores explican más de la mitad de las diferencias que existen en el rendimiento en Lengua, concretamente el 65%. Por último, el nivel socioeconómico de los centros comparte el 13% de la varianza con los resultados de Lengua (figura 1), cifra muy similar al impacto que tiene el ISEC en Matemáticas.

Modelo de eficacia para Inglés: en Inglés también se mantiene estable el conjunto de predictores del modelo vinculados al nivel de estudiante y su contexto (sexo, satisfacción, ayuda en casa y nivel socioeconómico) (tabla 5). También la titularidad tiene un impacto significativo en la explicación de los resultados de Inglés.

El modelo estimado señala que las chicas con un ISEC medio, que no están satisfechas con el centro, no reciben ayuda en casa y pertenecen a centros públicos, obtienen 527.7 puntos en promedio; los chicos obtienen 24 puntos menos en promedio. Además, los centros concertados obtienen de media 25.5 puntos más. El resto de predictores (satisfacción, ISEC y ayuda en casa) mantienen el sentido e importancia de su impacto.

Con la inclusión de los predictores en el modelo se consigue explicar el 17% de la variación de los resultados en Inglés (tabla 6). Si ponemos la atención en los centros, el conjunto de variables explica aproximadamente el 38% de sus diferencias. El impacto del nivel socioeconómico de los centros es mucho más importante en Inglés que en Lengua Española y Matemáticas (figura 1). La varianza compartida es del 28%: hay una fuerte vinculación entre el indicador socioeconómico de las escuelas y el rendimiento en Inglés. Aun así, hay centros que destacan respecto a los que tienen características similares.

Modelo de eficacia para Ciencias y Tecnología: para interpretar este modelo debe tenerse en

cuenta que Ciencias y Tecnología es la asignatura con una menor aportación de los centros a los resultados: solo explican el 2.77% de las diferencias (ver el ICC en la tabla 4). De forma similar al resto de asignaturas, en el modelo estimado hay variables del estudiante y su contexto que tienen un efecto significativo (sexo, satisfacción, ayuda en casa y nivel socioeconómico). Se suma también otro predictor vinculado a las características de los estudiantes: la condición de inmigrante.

El modelo indica que las chicas no inmigrantes, con un nivel socioeconómico medio, que no están satisfechas con su centro y no reciben ayuda en casa obtienen 512 puntos en promedio (tabla 5). Los chicos obtienen 14 puntos más y los estudiantes emigrantes 36 puntos menos de media. El resto de predictores (satisfacción, nivel socioeconómico y ayuda en casa) mantienen el sentido e importancia de su impacto.

Como se indica en la tabla 6, la inclusión de los predictores consigue explicar un 8.44% de las diferencias en los resultados. Si ponemos el foco en los centros, la reducción de la variación es del 31.6%, es decir, hay poco margen en la aportación de los centros, pero incluso controlando esas características del contexto individual, las escuelas influyen en los resultados de Ciencias y Tecnología, hay varianza significativa en el nivel de centro. Por último, el nivel socioeconómico de los centros educativos comparte el 12.5% de la varianza con los resultados de Ciencias y Tecnología (figura 1).

Selección de los centros más y menos eficaces

Además de la información aportada, los modelos estimados posibilitan detectar las escuelas que más destacan por su residuo positivo o negativo, lo que conforma dos grupos de centros de interés desde el punto de vista de la eficacia escolar: centros en los que, habiendo controlado las variables contextuales, la eficacia escolar se ha logrado con claro éxito sobre los demás

centros, y escuelas donde no se alcanzan los niveles deseados, en las que se precisaría de intervenciones adecuadas.

Obtenidos los residuos de cada centro como la diferencia entre el rendimiento medio observado en cada asignatura y el esperado tras la intervención de las covariables en el modelo anidado, se ordenan por rangos y se seleccionan los diez centros con mayor residuo positivo (tabla 7)

y los diez con menor residuo negativo (tabla 8). La concreción de esta cifra corresponde a los centiles 85 y 15, respectivamente, y es arbitraria; dependiendo de los intereses del estudio podrían fijarse otros puntos de corte. No obstante, estos centiles garantizan que el tamaño del residuo sea suficientemente grande o pequeño para que el supuesto de alta o baja eficacia sea plausible. Puntos de corte en centiles intermedios comprometerían tal asunción.

TABLA 7. Ordenación de los diez centros de más alta eficacia por asignaturas

Cod_centro	RedN2	TamañoN2	ISECN2	InmigN2	MatesN2	Residuo
Matemáticas						
64	1	44	.17	.02	65.73	11.83
49	1	24	-.24	.00	574.41	44.05
27	0	27	-.46	.07	564.52	38.80
31	0	26	-.18	.00	552.52	35.15
72	0	17	-.38	.06	549.91	31.42
16	0	37	.23	.03	55.81	25.70
55	1	17	.42	.06	564.03	23.02
39	0	17	-.21	.00	53.38	2.12
34	0	38	-.26	.16	513.91	19.37
60	1	20	.28	.00	546.06	18.82
Lengua						
34	0	38	-.26	.16	554.54	38.63
44	0	40	.39	.03	578.35	38.44
64	1	44	.17	.02	558.20	33.86
67	1	43	.03	.02	539.99	32.39
27	0	27	-.46	.07	562.09	29.63
3	0	31	-.50	.06	531.29	24.01
59	1	19	-.58	.16	527.02	22.19
58	1	20	.02	.00	553.32	22.16
12	0	62	.23	.06	539.82	21.75
25	0	63	.37	.03	552.50	21.29
Inglés						
64	1	44	.17	.02	586.73	5.32
32	0	37	-.69	.27	54.25	45.14
27	0	27	-.46	.07	547.10	38.76
21	0	37	-.55	.11	523.23	38.69
25	0	63	.37	.03	556.11	37.94
13	0	42	-.16	.07	542.57	37.71
31	0	26	-.18	.00	538.92	29.92
34	0	38	-.26	.16	501.46	28.22
63	1	29	.01	.03	549.22	27.36
10	0	51	.17	.00	539.78	25.46

TABLA 7. Ordenación de los diez centros de más alta eficacia por asignaturas (cont.)

Cod_centro	RedN2	TamañoN2	ISECN2	InmigN2	MatesN2	Residuo
Ciencias y Tecnología						
64	1	44	.17	.02	563.86	21.68
48	0	32	-.68	.13	515.95	13.43
10	0	51	.17	.00	541.84	11.19
34	0	38	-.26	.16	519.93	1.32
78	0	20	-.34	.00	552.16	9.19
59	1	19	-.58	.16	514.71	9.13
55	1	17	.42	.06	56.90	8.31
16	0	37	.23	.03	534.03	7.85
54	1	67	.12	.01	521.45	7.72
32	0	37	-.69	.27	495.75	7.52

TABLA 8. Ordenación de los diez centros de más baja eficacia por asignaturas

Cod_centro	RedN2	TamañoN2	ISECN2	InmigN2	MatesN2	Residuo
Matemáticas						
42	0	17	-.50	.18	431.94	-25.01
69	1	45	.08	.07	481.88	-28.63
2	0	38	-.41	.05	454.12	-32.73
11	0	19	.07	.05	466.38	-33.58
15	0	19	-.09	.00	448.44	-34.42
24	0	22	.41	.14	472.79	-38.19
19	0	60	-.15	.02	462.90	-38.42
30	0	35	-.09	.06	454.61	-42.03
66	1	85	.57	.00	47.76	-51.07
68	1	27	.38	.00	452.80	-53.44
Lengua						
62	1	21	.27	.05	476.91	-21.97
41	0	25	-.66	.20	444.53	-23.30
4	0	32	-.21	.13	465.15	-23.32
19	0	60	-.15	.02	468.25	-28.68
11	0	19	.07	.05	465.08	-28.73
42	0	17	-.50	.18	428.08	-3.35
70	1	89	.74	.00	488.75	-33.42
24	0	22	.41	.14	464.40	-36.81
68	1	27	.38	.00	467.56	-37.46
20	0	16	-.62	.13	427.01	-47.55

TABLA 8. Ordenación de los diez centros de más baja eficacia por asignaturas (cont.)

Cod_centro	RedN2	TamañoN2	ISECN2	InmigN2	MatesN2	Residuo
Inglés						
56	1	77	.15	.03	496.67	-28.52
19	0	60	-.15	.02	462.04	-28.72
53	1	21	.05	.00	475.41	-28.73
11	0	19	.07	.05	466.23	-29.34
41	0	25	-.66	.20	439.03	-33.24
15	0	19	-.09	.00	434.30	-35.41
20	0	16	-.62	.13	428.11	-42.32
4	0	32	-.21	.13	436.42	-46.29
68	1	27	.38	.00	467.19	-5.35
36		12	.00	.08	57.98	
Ciencias y Tecnología						
20	0	16	-.62	.13	484.00	-8.91
65	1	26	.10	.00	483.70	-9.03
66	1	85	.57	.00	501.85	-1.36
41	0	25	-.66	.20	443.87	-11.08
52	1	23	.93	.00	499.00	-11.19
69	1	45	.08	.07	486.42	-13.38
30	0	35	-.09	.06	455.35	-13.77
71	1	50	.16	.04	476.45	-16.72
19	0	60	-.15	.02	471.04	-17.10
2	0	38	-.41	.05	457.34	-18.79

Discusión y conclusiones

Los modelos jerárquicos lineales han resultado una herramienta adecuada para lograr el objetivo de este estudio: saber qué factores contextuales han impactado positiva o negativamente en el aprendizaje escolar, más allá del rendimiento de los centros aportado previamente en la ED. Estos modelos también han permitido detectar, controlado el efecto de las variables contextuales, los centros escolares con los mejores y peores resultados de eficacia, por medio del residuo del modelo como indicador para su

selección. Advertimos que esta definición de eficacia escolar es forzosamente parcial (Scheerens, 2017) al estar basada en las puntuaciones obtenidas por los estudiantes en las competencias que se evalúan en la ED, es decir, en cuestiones de orden instructivo. Sabemos que la labor educativa de un centro es compleja y va más allá de aspectos meramente instructivos, lo cual no debe olvidarse a la hora de tomar en consideración los resultados presentados. Otra cuestión importante es la eliminación en los análisis de 110 estudiantes que cursan en centros muy pequeños. Las estimaciones sobre

ellos no hubieran sido muy fiables, lo cual no implica que carezcan de interés, muy al contrario. Es probable que se trate de centros ubicados en pequeños núcleos rurales con dinámicas específicas de trabajo.

En cuanto al análisis descriptivo de las variables contextuales, hemos detectado un cierto nivel de segregación en la condición de estudiante emigrante. La proporción de estudiantes con este perfil fue baja, solo de un 5% (114), pero el 85% de ellos están escolarizados en centros públicos. Es cierto también que la distribución de estudiantes en las dos redes (pública/concertada) es de 60-40 en el nivel de estudiantes y de 67-33 en el nivel de centros, lo que indica una clara diferencia en el número de estudiantes por centro entre ambas redes. En cuanto a la satisfacción parental con el centro y la recomendación de este a otras familias, los resultados han sido altos. La implicación parental también resultó alta, aunque aquí la interpretación debe hacerse en relación con el rendimiento académico: un nivel alto de implicación parental puede ser indicador de dificultades de los hijos en los estudios.

Con respecto a la descripción de las asignaturas, la proporción entre la variabilidad que hay entre estudiantes y la que se da entre centros educativos (ICC) es muy baja, lo que puede considerarse como un indicador de altos niveles de igualdad en el funcionamiento de los centros. Dicho de otro modo, los centros de La Rioja difieren muy poco entre sí en los resultados académicos. La única excepción se produce en Inglés, lo que también es habitual en el sistema educativo español (Madrid, 2010; Rascón y Bretones, 2018). Es esta una cuestión a la que habría que prestar atención específica, pues en un mundo globalizado como el de hoy, podría ser causa de segregación por razones socioeconómicas. En definitiva, la baja variabilidad entre centros educativos tiene como efecto que los modelos estadísticos generados tengan una limitada capacidad explicativa de la varianza total. Para poder incrementar la misma en estudios futuros sería conveniente considerar más covariables.

En relación con el análisis de las asignaturas, hemos encontrado que Matemáticas y Lengua Española operan básicamente bajo el mismo esquema. Las covariables significativas fueron del nivel de los estudiantes y no de los centros. Destacan el ISEC, el sexo y el nivel de satisfacción; la ayuda en casa apareció también, pero con una relación negativa. La discrepancia con Lengua Española es que, como en la mayoría de los estudios internacionales al respecto (Meinck y Brese, 2019; Reilly *et al.*, 2019), son las chicas las que tienden a puntuar más alto. En Inglés se ha generado un modelo similar, con la variedad de que la red sí tiene mucha importancia, señalando una diferencia notable a favor de los centros concertados. En Ciencias y Tecnología la variabilidad entre los centros fue mínima (solo el 2.77%), y resultó significativo el hecho de que el estudiante sea inmigrante.

Respecto a la selección de los centros de alta y baja eficacia, el modelo asumido con la estrategia jerárquica origina que en la relación obtenida se encuentren escuelas de ambas redes y de muy diverso nivel socioeconómico, así como que centros que en una asignatura obtienen un residuo muy alto (o muy bajo), en otras asignaturas no les sucede igual. En algunos trabajos se ha discutido la posible inestabilidad de estos indicadores de eficacia (Newton *et al.*, 2010), observando que suele producirse en evaluaciones realizadas en cursos sucesivos (Joaristi *et al.*, 2014). También la OCDE apunta como una posible razón la inestabilidad de los resultados de las escuelas pequeñas (OCDE, 2011, p. 137), aunque aquí estos centros no se han considerado. En todo caso, tomar en cuenta los residuos como indicadores de eficacia es, sin duda, más equitativo que basarse solo en las meras puntuaciones brutas, al contemplar el contexto en el que cada escuela desarrolla su labor. Por último, son de especial interés los centros que aparecen en todas o en casi todas las ordenaciones por ser claros candidatos a un análisis cualitativo posterior, lo que podría formar parte de una segunda fase del estudio.

En definitiva, como mencionaron Reynolds *et al.* (2014), las dos preguntas básicas de la investigación sobre eficacia escolar son ¿qué es lo que hace buena a una escuela? y ¿cómo logramos más buenas escuelas? Dar una adecuada

respuesta a qué se entiende por un buen centro educativo, por un colegio de alta eficacia, es una cuestión crucial porque de la misma va a depender en gran medida el tipo de resultados que se obtengan.

Referencias bibliográficas

- Aitkin, M. y Longford, N. (1986). Statistical modelling issues in school effectiveness studies. *Journal of the Royal Statistical Society*, 149(1), 1-43.
- Cervini, R. (2004). Influencia de los factores institucionales sobre el logro en matemática de los estudiantes en el último año de la educación media de Argentina. *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 2(1).
- Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94, 95-120.
- Creemers, B., Kyriakides, L. y Sammons, P. (2010). *Methodological advances in Educational Effectiveness Research*. Routledge.
- De Maeyer, S., Van den Bergh, H., Rymenans, R., Van Petegem, P. y Rijlaarsdam, G. (2010). Effectiveness criteria in school effectiveness studies: further research on the choice for a multivariate model. *Educational Research Review*, 5, 81-96. <https://doi.org/10.1016/j.edurev.2009.09.001>
- Edmonds, R. R. (1979). *Search for effective school: the identification and analysis of city school that are instructionally effective for poor children*. East Lansing: Michigan State University.
- Gaviria, J. L., Martínez-Arias, R. y Castro, M. (2004). Un estudio multinivel sobre factores de eficacia escolar en países en desarrollo: el caso de los recursos en Brasil. *Education Policy Analysis Archives*, 12(20).
- Goldstein, H. (1997). Methods in school effectiveness research. *School Effectiveness and School Improvement*, 8(4), 369-395.
- Goldstein, H. (2011). *Multilevel statistical models* (4th ed.). John Wiley & Sons.
- Hernández-Castilla, R., Murillo, F. J. y Martínez-Garrido, C. (2013). Factores de ineficacia escolar. *REICE*, 12(1), 103-118.
- Huffman, J. B., Pankake, A. y Muñoz, A. (2006). The tri-level model in action: site, district, and state plans for school accountability in increasing student success. *Journal of School Leadership*, 16(5), 569-582.
- INECSE (1998). *Diagnóstico general del sistema educativo. Avance de resultados*. Ministerio de Educación y Ciencia. <https://acortar.link/dEMBc>
- Instituto de Evaluación (2007). *Educación primaria 2007. Evaluación general del sistema educativo*. Ministerio de Educación. <https://acortar.link/oj7ma>
- Instituto de Evaluación (2009). *Evaluación general de diagnóstico 2009. Marco de la evaluación*. Ministerio de Educación. <https://acortar.link/MuNlx>
- Joaristi, L., Lizasoain, L. y Azpillaga, V. (2014). Detección y caracterización de los centros escolares de alta eficacia de la comunidad autónoma del País Vasco mediante modelos transversales contextualizados y modelos jerárquicos lineales. *Estudios sobre Educación*, 27, 37-61.
- Lizasoain, L. y Angulo, A. (2014). Buenas prácticas de escuelas eficaces del País Vasco. Metodología y primeros resultados. *Participación Educativa*, 3(4), 17-28.
- LOE (2006). Ley Orgánica 2/2006. *Boletín Oficial del Estado*, 106, de 4 de mayo de 2006, 17158-17207.

- López, E., Navarro, E., Ordoñez, X. G. y Romero, S. J. (2009). Estudio de variables determinantes de eficiencia a través de los modelos jerárquicos lineales en la evaluación PISA 2006: el caso de España. *Archivos Analíticos de Políticas Educativas*, 17(16).
- Madrid, D. (2010). El contexto social del alumnado y su relación con el rendimiento en lengua extranjera. En M. Falces, E. Hidalgo, J. Santana y S. Valera (eds.). *Para, por y sobre Luis Quereda* (pp. 519-533). Universidad de Granada.
- Marchesi, A. y Martín, E. (2002). *Evaluación de la educación secundaria: fotografía de una etapa polémica*. Fundación Santa María.
- Meinck, S. y Brese, F. (2019). Trends in gender gaps: using 20 years of evidence from TIMSS. *Large-Scale Assessment in Education*, 7. <https://doi.org/10.1186/s40536-019-0076-3>
- Murillo, F. J. (2005). *La investigación sobre eficacia escolar*. Octaedro.
- Newton, X., Darling-Hammond, L., Haertel, E. y Thomas, E. (2010). Value-added modeling of teacher effectiveness: an exploration of stability across models and contexts. *Educational Policy Analysis Archives*, 18(23).
- OCDE (2011). *La medición del aprendizaje de los alumnos: mejores prácticas para evaluar el valor agregado de las escuelas*. OCDE.
- Rascón, D. J. y Bretones, C. M. (2018). Socioeconomic status and its impact on language and content attainment in CLIL contexts. *Porta Linguarum: Revista Internacional de Didáctica de las Lenguas Extranjeras*, 28, 115-135.
- Raudenbush, S. W. y Bryk, A. S. (2002). *Hierarchical linear models: applications and data analysis methods* (2nd ed.). Sage.
- Reilly, D., Neumann, D. L. y Andrews, G. (2019). Gender differences in reading and writing achievement: evidence from the National Assessment of Educational Progress (NAEP). *American Psychologist*, 74(4), 445-458.
- Reynolds, D., Sammons, P., De Fraine, B., Van Damme, J., Townsend, T., Teddlie, C. y Stringfield, S. (2014). Educational Effectiveness Research (EER): a state-of-the-art review. *School Effectiveness and School Improvement*, 25(2), 197-230, <https://doi.org/10.1080/09243453.2014.885450>
- Scheerens, J. (2017). *Educational effectiveness and ineffectiveness. A critical review of the knowledge base*. Springer.
- Snijders, T. A. B. y Bosker, R. J. (2012). *Multilevel analysis: an introduction to basic and advanced multilevel modeling* (2nd ed.). Sage.
- Townsend, T. y Avalos, B. (2007). *International handbook of school effectiveness and improvement*. Springer.
- Van De Grift, W. J. C. M. y Houtveen, A. A. M. (2006). Underperformance in primary schools. *School Effectiveness and School Improvement*, 17(3), 255-273.
- Van De Grift, W. J. C. M. y Houtveen, A. A. M. (2007). Weaknesses in underperforming schools. *International Perspectives in At-Risk Education Research*, 12(4), 383-403.

Abstract

A Study of school effectiveness in Primary Schools using Hierarchical Linear Models

INTRODUCTION. A census study based on a secondary analysis of a previous diagnostic evaluation (DE) is presented on school effectiveness in Primary Education Schools in the Autonomous Community of La Rioja. The impact of contextual variables on the performance of the subjects evaluated was assessed, not using the gross mean scores obtained in the DE, but the adjusted

mean scores once the effect of contextual variables were controlled. **METHOD.** The control was carried out by multilevel statistical modelling with Hierarchical Linear Models at two levels: students and schools. The residues of the models were used as criteria for school effectiveness. **RESULTS.** A descriptive analysis by levels of the contextual variables and of the performances is shown, as well as a detailed analysis of the hierarchical models by subjects. From the residues of the models, an arrangement of the schools according to their effectiveness is achieved. **DISCUSSION.** The Hierarchical Linear Models proved to be an adequate tool to know the impact of the contextual variables in school learning, beyond the performance of the schools previously provided in DE. The models made it possible to know the schools with the best and worst results of controlled efficacy and the effect of the context. The hierarchical strategy detected that in the relation of schools there are schools from both networks and from very diverse socio-economic levels, as well as that schools that in one subject obtain a very high (or very low) residue, do not have the same in others, information that is impossible to obtain with more conventional analysis tools.

Keywords: *School effectiveness, Educational assessment, Contextual effects, Hierarchical linear models.*

Résumé

Étude de l'efficacité scolaire dans les écoles grâce à l'utilisation de modèles hiérarchiques linéaires

INTRODUCTION. Une étude de recensement sur l'efficacité scolaire dans les écoles de la Communauté Autonome de La Rioja est présentée, une analyse secondaire d'une précédente évaluation diagnostique (ED) dans cette communauté. L'impact des variables contextuelles sur la performance des sujets évalués a été évalué, pas en utilisant les scores moyens bruts obtenus dans l'ED, mais plutôt les scores moyens ajustés après contrôle de l'effet des variables contextuelles. **MÉTHODE.** Le contrôle a été réalisé par modélisation statistique multi-niveaux avec des modèles hiérarchiques linéaires à deux niveaux : élèves et centres. Les résidus des modèles ont été utilisés comme critère d'efficacité des écoles. **RÉSULTATS.** Une analyse descriptive par niveaux des variables contextuelles et des performances a été réalisée, ainsi qu'une analyse détaillée des modèles hiérarchiques par sujets. A partir des résidus des modèles, les écoles ont été classés en fonction de leur efficacité. **DISCUSSION.** Les modèles hiérarchiques linéaires se sont révélés comme un outil adéquat pour connaître l'impact des variables contextuelles sur l'apprentissage scolaire, au-delà de la performance des écoles, avant fournis à l'ED. Les modèles ont permis de connaître les écoles avec les meilleurs et les pires résultats d'efficacité, en contrôlant l'effet du contexte. La stratégie hiérarchique a détecté que dans la relation des écoles, il y a des écoles appartenant à tous les deux réseaux et avec des niveaux socio-économiques très différents. Les centres qui obtiennent un résidu très élevé (ou très faible) dans une matière, ne l'obtiennent pas dans une autre: des informations impossibles d'obtenir à partir des outils d'analyse plus conventionnels.

Mots-clés: *Efficacité scolaire, Évaluation pédagogique, Effets du contexte, Modèles hiérarchiques linéaires.*

Perfil profesional de los autores

Emelina López-González (autora de contacto)

Profesora titular de la Universidad de Valencia (UV). Forma parte del proyecto de investigación “Evaluación del riesgo de sesgo en revisiones sistemáticas y meta-análisis de estudios sobre intervenciones complejas (RoB-CI)”, de la Universidad Autónoma de Barcelona. Es miembro del grupo de investigación Evaluación y Análisis Secundarios del Sistema Educativo, de la UNIR. Sus líneas de investigación se centran en: análisis estadístico multivariante, modelos predictivos lineales y modelos lineales generalizados.

ORCID ID: <https://orcid.org/0000-0002-0357-3889>

Correo electrónico de contacto: emelina.lopez@uv.es

Dirección para la correspondencia: Dpto. Métodos de Investigación y Diagnóstico en Educación (MIDE). Facultad de Filosofía y CC. Educación. Avda. Blasco Ibáñez, 30. 46010 Valencia.

Enrique Navarro-Asencio

Profesor contratado doctor en la Universidad Complutense de Madrid (UCM) y director del Máster de Investigación en Educación de la UCM. Forma parte de los grupos de investigación Medida y Evaluación de Sistemas Educativos, de la UCM, y Evaluación y Análisis Secundarios del Sistema Educativo, de la UNIR. Es especialista en metodología de investigación educativa. Sus líneas de trabajo son: medición educativa y evaluación del rendimiento académico y factores asociados.

ORCID ID: <https://orcid.org/0000-0002-3052-146X>

Correo electrónico de contacto: enriquen@ucm.es

María José García-San Pedro

Coordinadora académica del Máster Universitario en Métodos de Investigación en Educación de la Universidad Internacional de La Rioja (UNIR). Doctora en Calidad y Procesos de Innovación Educativa por la Universidad Autónoma de Barcelona (UAB) y Licenciada en Filosofía por la Universidad del Salvador (Argentina). Investigadora posdoctoral en la Universidad de Alicante. Directora del área de Métodos de Investigación y Diagnóstico en Educación (MIDE), de la UNIR. Miembro del grupo de investigación Evaluación y Análisis Secundarios del Sistema Educativo, de la UNIR.

ORCID ID: <https://orcid.org/0000-0002-7361-3778>

Correo electrónico de contacto: mariajose.garciasanpedro@unir.net

Luis Lizasoain

Profesor titular en la Universidad del País Vasco (UPV). Sus líneas de investigación se centran en: evaluación de programas, centros y sistemas educativos con especial atención a las evaluaciones educativas a gran escala, a los estudios de eficacia y mejora escolar y a los procedimientos estadísticos de análisis multinivel. Coordinador de un grupo de investigación consolidado de la UPV/EHU e investigador principal de un proyecto sobre escuelas eficaces en el País Vasco.

ORCID ID: <https://orcid.org/0000-0002-0753-2806>

Correo electrónico de contacto: luis.lizasoain@ehu.eus

Javier Tourón

Catedrático de universidad en la Universidad Internacional de La Rioja (UNIR). Vicerrector de Innovación y Desarrollo Educativo en la Universidad Internacional de La Rioja (UNIR). Director de la Escuela UNIR de Formación de Profesores en Tecnología Educativa, Competencias Digitales y Desarrollo del Talento. Presidente del International Advisory Board (IAB) del Instituto de Investigación, Innovación y Tecnología Educativas (iTED) de la Universidad Internacional de La Rioja (UNIR).

ORCID ID: <https://orcid.org/0000-0001-8217-1556>

Correo electrónico de contacto: javier.touron@unir.net