



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

# **Algunas contribuciones a la modelación multinivel en la Investigación Escolar en España**

**Patricia Aurora Tapia Blásquez**

Universidad Politécnica de Valencia  
Departamento de Estadística y Optimización  
Valencia, España

2016



**Algunas contribuciones a  
la modelación multinivel en la  
Investigación Escolar en España**

**Patricia Aurora Tapia Blásquez**

Tesis presentada como requisito parcial para optar al título de:

**Doctor en Estadística y Optimización**

Director (a):

Dr. José Miguel Carot Sierra

Codirector (a):

Dr. Mario Miguel Ojeda Ramírez

Línea de Investigación:

Modelación Estadística en Investigación Educativa

Universidad Politécnica de Valencia

Departamento de Estadística y Optimización

Valencia, España

2016

*“Me imagino perfectamente que, cuando el Señor creó el Mundo y a las personas que habitan en él-obra que según la ciencia moderna llevó mucho tiempo- razonó consigo mismo de la manera siguiente: “Si lo hago todo predecible, estos seres humanos, a los que he dotado de cerebros bastante buenos, indudablemente aprenderán a predecirlo todo y, por lo tanto no tendrán aliciente alguno para hacer absolutamente nada, pues reconocerán que el futuro está totalmente determinado y que en él no puede influir ninguna acción humana. Por otra parte, si todo lo hago impredecible, gradualmente descubrirán que no existe base racional para tomar ninguna decisión y, por lo tanto, como en el primer caso, no tendrán motivo alguno para hacer nada. Ninguno de estos dos proyectos tiene sentido. Crearé, por tanto, una mezcla de los dos. Que unas cosas sean predecibles y otras impredecibles. Así, entre otras muchas cosas, su importante tarea será la de saber cuál es cuál”.*

*E.F. Shumacher.*

*De lo pequeño es hermoso*



## **Agradecimientos**

Quiero dedicar la realización de esta tesis a todas aquéllas personas con las que he tenido la oportunidad de coincidir en el camino y que me alentaron a continuar hasta la finalización de esta etapa de mi vida académica. En primer lugar agradezco a mis padres y a mi familia por su apoyo económico, moral y, sobre todo, por su amor incondicional.

Expreso mi gratitud a mi director de tesis, al Dr. José Miguel Carot Sierra, por haberme invitado a participar en el proyecto del Centro de Gestión de la Calidad y del Cambio (CQ) de la Universidad Politécnica de Valencia, donde surgió el tema de esta tesis, por su confianza en mis capacidades y por compartir sus conocimientos y experiencia profesional. A mi codirector de tesis en México, el Dr. Mario Miguel Ojeda Ramírez, quien me contagió de su pasión por la estadística y me introdujo en el mundo de la investigación. Le agradezco que haya creído en mí y que hasta la fecha continúe compartiendo sus experiencias.

De manera muy especial, quiero agradecer al Dr. José Ginés Mora (Pepe), a Amparo, a todos los amigos del CQ y de Valencia, por ser una segunda familia para mí mientras realicé mi estancia y con quienes comprobé las experiencias de vida y de amistad que la hacen significativa. Finalmente, agradezco al Consejo Nacional de Ciencia y Tecnología (CONACyT) en México, por el financiamiento para la realización de mi doctorado.

## Resumen

La modelación multinivel se ha convertido en una poderosa herramienta de análisis dentro del campo de la investigación escolar, especialmente en la identificación de factores que contribuyen a explicar el rendimiento de los estudiantes en una pluralidad de contextos. Recientemente, en España se ha avanzado en esta línea pero con un limitado nivel en el desarrollo metodológico. Falta explorar nuevas posibilidades para el análisis de datos anidados, pues la mayoría de investigaciones españolas en eficacia escolar que se han hecho, se han limitado a la aplicación de modelos multinivel univariados. El propósito de este trabajo fue mostrar las nuevas aportaciones que tienen los modelos multinivel multivariados, al ajustar simultáneamente dos o más respuestas, utilizando datos de la Evaluación General de Diagnóstico 2009, aplicada por el Instituto de Evaluación Educativa en España y que tiene por objetivo medir el rendimiento en las competencias básicas del alumnado de cuarto grado de educación básica. Se trabajó con una muestra de 28.708 alumnos, 884 profesores, 898 directores, 25.741 padres de familia y 900 centros escolares de toda España, de la que se obtuvo la información sobre el rendimiento en la competencia lingüística y matemáticas de los alumnos, así como de las variables explicativas para cada nivel, a través de las encuestas que respondieron los padres, profesores y directores de las escuelas. Se especificaron 3 modelos de tres niveles (alumno, aula, escuela), uno para cada competencia por separado (lingüística y matemática) y un modelo multivariado de cuatro niveles que incluye ambas respuestas simultáneamente. Se compararon los resultados de los estimadores, de los errores estándar, los gráficos de los residuos y la matriz de varianzas y covarianzas, así como los efectos asociados a cada uno de los niveles. Se concluye que los modelos multivariantes multinivel ofrecen pruebas más potentes y mayor información en su matriz de varianzas para comprender el fenómeno educativo.

**Palabras clave:** Modelos multinivel multivariado, investigación escolar, análisis estadístico, eficacia escolar, Evaluación General de Diagnóstico.

## **Abstract**

Multilevel modelling has become a powerful tool in school effectiveness research, particularly, finding those factors that influence student achievement in diverse contexts. Recently, Spain has registered an improvement in research on Education, but limited in the developed methodologies. There are a lot of issues on multilevel models to explore, because most of the models applied in Spanish educational research have been univariate. The aim of this work was to determine the new contributions of multivariate multilevel models adjusting simultaneously two responses. The data was collected from the General Evaluation Diagnostic applied by the Institute of Evaluation on Education in Spain at 2009, which has the purpose to examine the results in Language and Mathematics competences from students coursing fourth grade of elementary education. The sample used was integrated by 28.708 students, 884 teachers, 898 principals, 27.541 parents and 900 schools centers. We had the achievement in each competence by each student, as well as the variables gathered from the survey applied to teachers, parents and directors. The models specified include three levels (Student, classroom and school center), one model for each competence, and one multivariate model in order to compare the results and finding new approaches of using this methodology in the understanding of school effectiveness. We compared the results of the estimators, their standard errors, residuals and variance and covariance matrix. The main conclusion is that there are still sides to take advantage from multilevel multivariate models, which result more integrated models and more powerful to provide information about scholar effectiveness.

**Keywords:** Multilevel Multivariate Modelling, school effectiveness, educative research, General Evaluation of Diagnostic.

## Resum

La modelització multinivell s'ha convertit en una poderosa eina d'anàlisi dins del camp de la investigació escolar, especialment en la identificació de factors que contribueixen a explicar el rendiment dels estudiants en una pluralitat de contextos. Recentment, en Espanya se ha avançat en aquesta línia però amb un limitat nivell en el desenvolupament metodològic. Falta explorar noves possibilitats per a l'anàlisi de dades niades, doncs la majoria de recerques espanyoles en eficàcia escolar que s'han fet, s'han limitat a l'aplicació de models multinivell univariats. El propòsit d'aquest treball va ser mostrar les noves aportacions que tenen els models multinivell multivariats, en ajustar simultàniament dues o mes respostes, utilitzant dades de l'Avaluació General de Diagnòstic 2009, aplicada per l'Institut d'Avaluació Educativa en Espanya i que té per objectiu mesurar el rendiment en les competències bàsiques de l'alumnat de quart grau d'educació bàsica. Es va treballar amb una mostra de 28.708 alumnes, 884 professors, 898 directors, 25.741 pares de família i 900 centres escolars de tota Espanya, de la qual es va obtenir la informació sobre el rendiment en la competència lingüística i matemàtiques dels alumnes, així com de les variables explicatives per a cada nivell, a traves de les enquestes que van respondre els pares, professors i directors de les escoles. Es van especificar 3 models de tres nivells (alumne, aula, escola), un per a cada competència per separat (lingüística i matemàtiques) i un model multivariat de quatre nivells que inclou ambdues respostes simultàniament. Es van comparar els resultats dels estimadors, dels errors estàndard, els gràfics dels residus i la matriu de variàncies i covariàncies, així com els efectes associats a cadascun dels nivells. Es conclou que els models multivariants multinivell ofereixen proves mes potents i major informació en la seua matriu de variàncies per a comprendre el fenomen educatiu.

**Paraules clau:** Models multinivell multivariat, investigació escolar, anàlisi estadística, eficàcia escolar, avaluació general de diagnòstic

# Índice General

<b>INTRODUCCIÓN</b>	<b>1</b>
<b>OBJETIVOS</b>	<b>5</b>
GENERAL	5
ESPECÍFICOS	5
<b>CAPÍTULO 1: PANORAMA DE LA INVESTIGACIÓN ESCOLAR.</b>	<b>8</b>
1.1 EL MOVIMIENTO DE ESCUELA EFICAZ.	12
1.2 LA INVESTIGACIÓN EN ENSEÑANZA EFICAZ.	14
1.3 LA INVESTIGACIÓN EN EFICACIA EDUCATIVA.	15
1.4 LA INVESTIGACIÓN EDUCATIVA EN ESPAÑA.	19
1.4.1 Organismos para la investigación educativa.	19
1.4.2 Estado del Arte de la Investigación Educativa en España.	21
1.4.3 Investigaciones actuales en eficacia escolar en España.	27
<b>CAPÍTULO 2: MODELIZACIÓN MULTINIVEL EN INVESTIGACIÓN ESCOLAR.</b>	<b>33</b>
2.1 LOS MODELOS MULTINIVEL	33
2.2 COMPARACIÓN DE LOS MODELOS MULTINIVEL FRENTE A OTRAS TÉCNICAS DE ANÁLISIS.	34
2.3 CONTRIBUCIONES DE LOS MODELOS MULTINIVEL A LA INVESTIGACIÓN ESCOLAR.	36
2.4 ESTRUCTURA Y NOTACIÓN DEL MODELO MULTINIVEL.	37
2.5 PRINCIPIOS DE ESTIMACIÓN EN LOS MODELOS MULTINIVEL.	41
2.5.1 Método de Máxima Verosimilitud Completa.	42
2.5.2 Máxima Verosimilitud Restringida.	42
2.5.3 Estimación Bayesiana.	43
2.6 PRUEBA DE HIPÓTESIS EN LOS MODELOS MULTINIVEL.	43
<b>CAPÍTULO 3: INVESTIGACIÓN ESCOLAR EN ESPAÑA: MÉTODOS Y MATERIALES</b>	<b>46</b>
3.1 LA EVALUACIÓN GENERAL DE DIAGNÓSTICO.	46
3.2 LA COMPETENCIA EN LINGÜÍSTICA Y MATEMÁTICAS.	47
3.3. FACTORES ANALIZADOS.	48
3.4 NIVELES CONSIDERADOS EN EL MODELO.	50
3.5 MODELOS MULTINIVEL ESPECIFICADOS.	52
3.5.1. Modelo univariado de 3 niveles para la variable respuesta competencia lingüística.	52
3.5.2. Modelo univariado de 3 niveles para la variable respuesta competencia en matemáticas.	53
3.5.3 Modelo multinivel multivariado para las variables respuesta competencia en lingüística y competencia en matemáticas.	53
3.6 MODELO MULTIVARIADO DE CUATRO NIVELES.	54
<b>CAPÍTULO 4: RESULTADOS Y APORTACIONES</b>	<b>58</b>
4.1 ANÁLISIS DESCRIPTIVO DE LOS DATOS.	58
4.2 COMPARACIÓN DE MODELOS NULOS.	62
4.2.1. Factores en el primer nivel.	65
4.3 MODELO MULTIVARIADO DE INTERCEPTO Y PENDIENTE ALEATORIA DE UNA VARIABLE EXPLICATIVA.	67
4.4. COMPARACIÓN DE LOS MODELOS UNIVARIADOS Y LOS MODELOS MULTIVARIADOS MULTINIVEL.	69

4.5 ANÁLISIS DE RESIDUOS.	73
<b>CONCLUSIONES.</b>	<b>78</b>
<b>LÍNEAS FUTURAS DE INVESTIGACIÓN.</b>	<b>82</b>
<b>REFERENCIAS BIBLIOGRÁFICAS.</b>	<b>85</b>
<b>ANEXOS</b>	<b>93</b>
ANEXO 1: CUESTIONARIOS.	93
ANEXO 2: RESULTADOS DE LOS PARÁMETROS ESTIMADOS EN EL MODELO MULTIVARIADO MULTINIVEL (COMPETENCIA LINGÜÍSTICA Y COMPETENCIA EN MATEMÁTICAS).	97
ANEXO 3: RESULTADOS DE LOS PARÁMETROS ESTIMADOS EN EL MODELO UNIVARIADO MULTINIVEL PARA LA COMPETENCIA LINGÜÍSTICA (MODELO 1) Y PARA LA COMPETENCIA MATEMÁTICAS (MODELO 2).	98

## Lista de tablas

Tabla 1-1: Movimientos en la investigación escolar (1960 – a la fecha). .....	11
Tabla 1-2: Factores clave de Enseñanza Eficaz detectados por Brophy y Good (1984). .....	14
Tabla 1-3: Aportación de Investigaciones destacadas en Eficacia Educativa por país. ....	17
Tabla 1-4: Principales trabajos de investigación en eficacia escolar en España. ....	24
Tabla 2-1: Ventajas y desventajas de las técnicas convencionales de regresión o de un solo nivel. ....	35
Tabla 3-1: Distribución de la muestra .....	46
Tabla 3-2: Dimensiones de la práctica educativa del profesorado. ....	50
Tabla 3-3: Factores explicativos introducidos en el modelo. ....	51
Tabla 4-1: Resultados por alumno en Competencia Lingüística y en Competencia en Matemáticas .....	60
Tabla 4-2: Resultados de la EGD 2009 por comunidad autónoma. ....	61
Tabla 4-3: Resultados del modelo multinivel nulo para rendimiento en Lingüística .....	62
Tabla 4-4: Resultados del modelo multinivel nulo para rendimiento en Matemáticas. ....	63
Tabla 4-5: Resultados del modelo multinivel multivariado nulo para rendimiento .....	64
en Lingüística y Matemáticas. ....	64
Tabla 4-6: Componentes de la varianza modelo con factores en el primer nivel. ....	65
Tabla 4-7: Correlaciones entre los residuos de los diferentes niveles del modelo multivariado. ....	66
Tabla 4-8: Resultados del modelo multivariado de intercepto y pendiente aleatoria. ....	67
Tabla 4-9: Resultados Comparativos para los componentes de la varianza de los modelos univariados en competencia en Lingüística y Matemáticas. ....	70
Tabla 4-8: Estimación efectos aleatorios modelo multivariado multinivel. ....	73

## Lista de Figuras

Figura 1-1: Modelo de los Cinco Factores de Edmonds (1979) .....	13
Figura 1-2: Modelo Propuesto por Schereens y Creemers (1989) .....	16
Figura 2-2: Errores a nivel individual y grupal en un modelo de dos niveles. ....	38
Figura 3-1: Diagrama de unidad de los niveles del modelo. ....	51
Figura 4-3: Residuos a nivel centro escolar .....	68
Figura 4-4: Residuos del Modelo Multivariado Multinivel. ....	74
Figura 4-5: Residuos del Modelo Multivariado Multinivel (aula escolar). ....	74
Figura 4-6: Gráfico caterpillar modelo multivariado. ....	75
Figura 4-7: Diagnóstico de residuos del nivel centro escolar .....	76

## Introducción

En los últimos años, los modelos multinivel o modelos lineales jerárquicos, se han convertido en la herramienta de aplicación más utilizada en el campo de la investigación escolar, logrando avances vertiginosos en sus resultados y contribuyendo a profundizar en el conocimiento de los fenómenos educativos. Una de las líneas en las que más ha contribuido es en la Investigación sobre Eficacia Escolar, cuyo objetivo es ahondar en la comprensión de los factores que se asocian a mejorar el rendimiento de los estudiantes y explicar por qué unas escuelas obtienen mejores resultados que otras. Sus inicios se relacionan con la publicación del Informe Coleman en 1966, cuya principal conclusión fue que las escuelas no tenían ningún efecto en el aprendizaje de los estudiantes. Esto desencadenaría el desarrollo de una serie de investigaciones que buscarían demostrar la invalidez de tal resultado.

Sin embargo, las metodologías que utilizan los primeros trabajos cuantitativos no demostraron ser las adecuadas, hasta que Aitkin y Longford (1986) exponen por primera vez que la modelación multinivel sería la metodología apropiada para validar la comparación entre escuelas, al contraponer los resultados con la aplicación de modelos convencionales (modelos de regresión múltiple) y observar las diferencias, pues la agregación sobre las observaciones individuales puede conducir a resultados engañosos. Igualmente, Kreft (1993) menciona que estos modelos representan un marco sólido y más aproximado que cualquier otro método, para analizar datos con una estructura jerárquica, como es el caso de los datos escolares, en el que los estudiantes se encuentran anidados en escuelas, en su agrupación más simple. Las primeras investigaciones que aplicaron esta metodología permitieron comprobar el efecto que tiene la escuela en el desempeño de los estudiantes, en refutación a las conclusiones del Informe Coleman.

A partir de estos resultados, la comunidad científica adoptará la modelación lineal multinivel o modelación lineal jerárquica (HLM, por sus siglas en inglés), como metodología para análisis de datos en los que se desee medir los efectos del aula, centro, país o sistema educativo en el rendimiento de los estudiantes. Esto provocó un acelerado desarrollo teórico-metodológico que permitió su aplicación y divulgación, y a pesar de que las investigaciones pioneras fueron en el ámbito anglosajón, pronto se extendería al resto de la comunidad internacional.



La modelación multinivel continúa demostrando ser la herramienta más consolidada para la investigación en eficacia escolar, aunque recientemente han sido desarrolladas otras técnicas de análisis en investigación escolar, como la técnica de árboles estadísticos de decisión del área de minería de datos (Lizasoain y Joaristi, 2012; Castro y Lazasoain, 2012) y los modelos de ecuaciones estructurales (SEM, Structural Equations Models) (Raykov y Marcoulides, 2006; Skrandall y Rabe-Hesketh, 2004). Es importante mencionar, que esta situación también se ha visto favorecida por el rápido y variado desarrollo de software estadístico para la aplicación de esta metodología. Actualmente se cuenta con una gran cantidad de paquetería, entre la que destaca Ml win (Multilevel Modelling for Windows) Stata (Modelos mixtos) Proc Mixed de SAS (Statistical Analysis System), HLM de SSI (Hierarchical Linear and Nonlinear Modelling de Scientific Software International), ML3, VARCL (Kreft y de Leeuw, 1994).

Los modelos multinivel se enfocan al análisis de datos, donde las observaciones están anidadas dentro de grupos. Por ello, su aplicación en el ámbito educativo ha resultado muy pertinente, pues las unidades objeto de estudio, el alumnado, se agrupan en aulas que a su vez, se congregan en escuelas y las escuelas en ciudades o distritos y estas en países. Debido a la inclusión de estos niveles, se puede identificar más fácilmente dónde se encuentra la mayor fuente de variación en los resultados educativos, si en el aula, en las escuelas o en las comunidades. Las primeras investigaciones en eficacia escolar que aplicaron modelación multinivel, se centraron en modelos de dos niveles, considerando el alumno como primer nivel y la escuela como segundo nivel. Poco a poco, se fue incrementando su complejidad, hasta la especificación de modelos de cuatro niveles en los que se considera el contexto del país para explicar la variación en el rendimiento del alumnado, lo cual ha provocado que esta línea de investigación escolar sea de las más prolíficas y existan una gran diversidad de publicaciones al respecto.

Sin embargo, a pesar del consenso en la comunidad científica sobre las ventajas del uso de los modelos multinivel para el análisis de datos en educación, en España existe una “escasa tradición en investigación escolar y un nivel bajo en desarrollo metodológico” (Murillo, 2008, p. 344). Los trabajos de investigación que se habían realizado en España hasta el año 2000, habían estado desfasados con respecto a investigaciones similares en otros lugares, es decir, la aplicación de metodologías novedosa utilizando datos de la realidad educativa española, se llevaban a cabo años después de haber sido propuestas en otros países.

Por otro lado, aunque muchas investigaciones utilizaban una gran cantidad de información o recopilación de variables relativas al alumnado, al profesorado, a los directivos y a los centros, no incorporaban alguna prueba estandarizada para medir el rendimiento de los estudiantes como variable respuesta a explicar. Es a partir de la participación del CIDE en el proyecto internacional de la UNESCO para medir el nivel de calidad educativa internacional, que se utilizan los resultados de los alumnos españoles en la prueba PISA. (Choi de Mendizabal y Calero-Martínez, 2009). Estas investigaciones utilizaron análisis causal o técnicas cualitativas como el estudio de casos, salvo el caso de Castejón (1996), considerado uno de los más representativos en investigación educativa. y que aplica análisis de regresión múltiple y elaboración de índices.

Los estudios multinivel se inician en España a principios del año 2000, con investigaciones como las llevadas a cabo por Murillo (2001) y Calero y Escardíbul (2007), quienes para analizar los factores de contexto español que puede estar asociados al rendimiento, utilizaron modelos univariados multinivel. Blanco et al., (2014) proponen continuar la línea de investigación sobre eficacia escolar, aplicando modelos multivariados multinivel para analizar datos de los resultados en la prueba PISA para España.

Por ello, es importante continuar en la búsqueda de nuevas contribuciones de modelos más integrados que permitan incorporar variables que no han sido estudiadas como factores asociados al rendimiento escolar y particularmente, identificar nuevos aportes de esta metodología a la comprensión de este fenómeno educativo. Esto impulsa a seguir el camino ya iniciado desde hace 40 años, analizando a profundidad esta herramienta de análisis, identificando nuevos aportes que puede hacer su aplicación para la mejora en la calidad de la educación.

Las preguntas que motivan la realización de esta tesis son: ¿Qué nuevas contribuciones metodológicas en los modelos multinivel de respuesta multivariada aportan información adicional, con respecto a los modelos univariados multinivel para la comprensión del rendimiento escolar?, ¿A qué nuevas preguntas se puede dar respuesta con los resultados que arrojan estos modelos integrados?

Para llevar a cabo la investigación, se utilizaron datos de la Evaluación General de Diagnóstico, que son resultados de pruebas elaboradas para los alumnos españoles.

# Objetivos

## Objetivos

La realización de esta investigación tiene como objetivo principal, estudiar novedosas contribuciones que surgen de la aplicación de la modelación multinivel para la comprensión de los resultados educativos en la Educación Básica en España. Es decir, a través del uso de esta metodología estadística, conocer nuevas aristas en el desarrollo de las competencias en estudiantes de educación primaria.

Esta herramienta de análisis ha sido validada para realizar investigación en el ámbito educativo, por la estructura jerárquica que presentan los datos a analizar (estudiantes que pertenecen a diferentes centros escolares), en la cual se puede explicar la variabilidad de los resultados, a través de la integración simultánea de los factores explicativos en cada nivel y que permite determinar si esa variación se atribuye al estudiante o a las características del centro escolar al que asiste. Aunque existen copiosas aplicaciones de la modelación multinivel, aún hay líneas por explorar, tales como la especificación de un modelo de respuesta multivariada, en el que se analice simultáneamente el efecto de los factores en dos competencias básicas de los estudiantes, considerando el centro escolar. Asimismo, se realiza la comparación de los resultados, para identificar la eficiencia del uso de un modelo más complejo con respecto a los modelos más utilizados de respuesta univariada y la información adicional que se puede obtener para la comprensión del fenómeno educativo.

Los objetivos que guían la realización del trabajo son:

### General

Explorar nuevos ángulos y aportes de la modelación multinivel de respuesta multivariada, como herramienta para el análisis de datos en investigación escolar en España.

### Específicos

- Aplicar modelos integrados de modelación multinivel que proporcionen más información para la comprensión del fenómeno del rendimiento escolar.
- Comparar las aportaciones de los modelos integrados frente a los modelos de respuesta univariada.

- 
- Estudiar la correlación tanto entre los niveles del modelo como entre las variables respuesta en un modelo integrado.
  - Analizar los residuos de un modelo multivariado multinivel para identificar casos de éxito o fracaso en los centros escolares de Educación básica en España 2009.
  - Analizar si el efecto de los factores en las diferentes respuestas analizadas es mayor, igual o menor.

Para alcanzar los objetivos planteados anteriormente, en primer lugar se hizo una revisión del contexto de la investigación escolar, así como la inclusión de los antecedentes que permitieron identificar cuándo se aplicaron los modelos multinivel, con qué propósito y justificación académica, así como los resultados que se han obtenido hasta el momento. Esta información se encuentra descrita en el primer capítulo. Posteriormente, en el segundo capítulo, se presenta el análisis de la modelación multinivel, sus características, su formulación matemática y sus aportes.

En el tercer capítulo, se describen los datos que fueron analizados para la especificación de los modelos y lograr su comparación. Se utiliza y contextualiza los resultados de la Evaluación General de Diagnóstico, utilizada en España, como un instrumento para medir el nivel de adquisición de competencias de los estudiantes de cuarto grado de Educación Primaria.

Los análisis de los resultados obtenidos, la comparación y las aportaciones se desglosan en el cuarto capítulo, en el que al inicio se presenta un descriptivo de los datos utilizados. En un quinto capítulo se discuten los alcances y limitaciones de los resultados y finalmente se presentan las conclusiones medulares de la investigación.

## Capítulo 1

# Panorama de la Investigación Escolar

## Capítulo 1: Panorama de la Investigación Escolar.

Las temáticas que la investigación escolar ha abarcado, van desde el diseño de políticas educativas y análisis de sistemas de educación, hasta la elaboración de planes y programas de estudio y contenido de la enseñanza, métodos de enseñanza y estrategias pedagógicas, formación docente y ejercicio de la docencia, educación especial e integración escolar, problemas de aprendizaje, costos y financiamiento de la educación, gestión educativa y autonomía escolar, oportunidades educativas y discriminación, hasta Educación Superior y acreditación universitaria, por mencionar algunas (Fabara y Hernández, 2003).

Uno de los objetivos principales que ha motivado al desarrollo de copiosas investigaciones, ha sido el observar qué ocurre en las escuelas para poder mejorarlas. Esto dio paso a una línea muy importante, la de investigación escolar, en la cual se identifican cuatro vertientes: los estudios sobre eficacia escolar, los que buscan encontrar una relación entre determinados factores y el rendimiento, los de eficacia docente, en el que se centra en el papel del profesor y los de eficacia educativa, que incluyen las dos anteriores, pero aplicando modelación multinivel. (Murillo, 2003).

La investigación en Eficacia Escolar tuvo su origen en las reacciones a los resultados del Informe Coleman publicado en 1966. Este estudio fue llevado a cabo en Estados Unidos (US) por un grupo de investigadores de distintas universidades, pero mayoritariamente de la Oficina de Educación (US), coordinados por James S. Coleman de la Universidad John Hopkins, en respuesta a lo establecido en la Sección 402 del Acta de Derechos Civiles de 1964 que establece:

“El Comisionado deberá realizar un estudio y elaborar un informe al Presidente y al Congreso, dentro de los dos años siguientes a la entrada en vigor de este título, relativo a la falta de disponibilidad de la igualdad de oportunidades educativas para las personas por razón de raza, color, religión origen o nacional en las instituciones educativas públicas a todos los niveles en los Estados Unidos, sus territorios y posesiones, y el Distrito de Columbia” (Coleman, J. et al., 1966).

Como se señala en el informe, el objetivo fue documentar el grado de segregación de los grupos minoritarios en las escuelas públicas a las que asistían, su relación con el rendimiento

logrado y compararlo con los resultados de la raza blanca. Para llevarlo a cabo, obtuvieron una muestra de 60.000 profesores pertenecientes a 4.000 escuelas de todo el país y 600.000 estudiantes (Báez de la Fe, 1991). El aprendizaje de los alumnos fue medido a través de la aplicación de pruebas estandarizadas y como indicadores de las escuelas, se obtuvo información del número de laboratorios que tenían, libros de texto, bibliotecas y otras relacionadas con los programas académicos ofrecidos. Para la información de los profesores, recopilaron datos como su educación, experiencia, salario y habilidades. Los grupos raciales que consideraron en el estudio fueron 6 (Afroamericanos, Indios-Americanos, Mexicanos-Americanos, Asiáticos-Americanos, Portorriqueños y Blancos) y utilizaron como método el análisis de regresión.

El reporte fue titulado *Equality of Educational Opportunity* (Igualdad de Oportunidades en Educación) y contiene 575 páginas. Se dividió en 9 secciones: Resumen, el ambiente de la escuela, el logro de los estudiantes y la motivación, futuros maestros para grupos minoritarios, educación superior, estudiantes no matriculados, casos de estudio e integración escolar, estudios especiales y apéndices.

Los principales resultados que presentaron se resumen en mostrar que la escuela tiene un efecto nulo en el rendimiento de los estudiantes, es decir que no hay diferencia en los resultados del aprendizaje por asistir a determinado centro escolar. Esto indicaba que no importaba cuánto dinero se destinaba a las instituciones, el medio del estudiante (nivel socioeconómico, preparación de los padres) influía mucho más en sus logros educativos, que los elementos físicos de la escuela, la preparación de sus profesores o sus programas académicos. El trasfondo del resultado se traduce en la conclusión a la que conlleva: No resulta necesario invertir dinero en las escuelas si el responsable del aprendizaje es únicamente el estudiante y su contexto personal.

“Quizá el hallazgo más sorprendente de esta sección es el hecho de que no hay diferencias significativas entre los colegios a los que asisten los diferentes grupos étnicos en cuanto a la variedad objetiva de los factores que determinan la calidad de un centro educativo. De hecho, las características tienden a ser mayores entre los centros educativos de las diferentes regiones del país que entre los colegios a los que asiste cada grupo étnico. [...]El único factor escolar en el que las minorías étnicas se ven desfavorecidas realmente es en el de las características psicológicas, sociales y económicas de sus compañeros de clase. [...] El hándicap del alumno negro no es



únicamente que debe hacer frente a sus propios problemas socio-económicos de su hogar, sino que está rodeado de gente con las mismas desventajas, que más que animarlo hacia un mejor desempeño académico le hacen ser más pesimista acerca de su futuro” (Coleman et al., 1966).

El informe Coleman, es considerado como uno de los estudios más completos y representativos en materia educativa del siglo XX. Sin embargo, sus resultados no fueron bien recibidos por la comunidad científica ni académica y han sido objeto de mucha crítica. En primer lugar, por las implicaciones sociales que tuvo, pues los resultados generaron el movimiento conocido como “*the white flight*” (fuga de blancos), las familias blancas sacaron a sus hijos de las escuelas donde convivían con alumnos de grupos raciales minoritarios.

Asimismo, desencadenó dos posturas políticas, la neo-conservadora que apoyó los resultados y acuñó la frase “*Schools don't matter*” para la definición de políticas públicas y la reactiva, que generó el movimiento conocido como *Effective Schools* (Escuelas Eficaces) y que a la fecha se ha consolidado como una de las líneas en investigación escolar más prolíficas en aportaciones y metodologías empleadas y que ha mostrado el importante efecto que tiene el centro escolar en el rendimiento de los estudiantes. A pesar de las críticas y de la posterior justificación de los autores, al argumentar que se encontraron presionados por el estudio, el Informe Coleman constituye un parteaguas en la investigación escolar.

Paralelamente a este estudio, en Gran Bretaña se realizó el informe Plowden en 1967 que confirmaba los hallazgos del informe Coleman (Murillo, 2003; Penalva-Buitrago, 2007). Esta investigación se llevó a cabo en dos fases, una con estudiantes de educación primaria y cuatro años después con estudiantes de educación secundaria, para observar si la variación demostrada entonces era determinada en la etapa primaria. Encontraron que la proporción de variación en logros educativos debido a variables relacionadas con el hogar, era al menos cuatro veces mayor que la aportada por las variables escolares (Wall y Varma, 1975). Esto provocó que el análisis se enfocara a la atención en la cultura del hogar y en cómo la actitud de los padres puede influir en la educación de los hijos. Los resultados fueron relevantes, el trabajo delimita los supuestos de una educación progresista y centrada en las necesidades e intereses del alumnado (Tyler, 1991). Investigaciones similares se llevaron a cabo por Jencks et al. (1972), Mayeske (1972) y Smith (1972) con alguna variación en las variables analizadas o en

el método empleado, pero confirmando las conclusiones hasta entonces logradas, del nulo efecto de la escuela en el logro educativo del estudiante.

Derivado de esta situación, se inicia una nueva línea en la investigación escolar basada en estudios de proceso-producto (Martínez-Garrido, 2011). El desarrollo de estos trabajos formales, entablan las bases para comprender el rendimiento escolar analizado desde un enfoque meramente cuantitativo que se identificó como input-output (Ver tabla 1-1).

Tabla 1-1: Movimientos en la investigación escolar (1960 – a la fecha).

	Fundamento	Periodo	Metodología	Principales aportes
<b>Eficacia Escolar</b>	Estudian los factores de la escuela que afectan al desarrollo académico de los estudiantes	1960-1970	Input-output.  Análisis cuantitativo-regresión lineal.	Inician los estudios de los factores de la escuela que afectan al desarrollo académico. Encuentran que las escuelas no hacen diferencia en los resultados ( <i>Schools don't matter</i> )
		1970-	Estudios de profundidad.  Escuelas prototípicas (outliers).	Para poder entender qué es y en qué consiste la eficacia escolar, es necesario saber qué ocurre dentro del aula
		Finales 70	Grandes muestras de datos.  Input-Process-Output	Comienzan a considerar variables que tuvieron gran variación entre las escuelas.  <i>"Schools can make a difference"</i>
<b>Enseñanza Eficaz</b>	Identificar los aspectos propios del proceso de enseñanza y del docente que influyen en la consecución de logros del estudiante.	Años 80	Experimental Cuantitativo (Correlaciones)	
<b>Eficacia Educativa</b>	Nexo entre eficacia escolar y enseñanza eficaz. Se retoma en concepto de eficacia escolar, pero en un sentido más global.	Años 90	Modelos multinivel	<i>"Schools matter"</i>

Fuente: Elaboración propia a partir de (Martínez-Garrido, 2011).

### 1.1 El movimiento de Escuela Eficaz.

Los primeros estudios de escuelas eficaces pretendían echar abajo las conclusiones del Informe Coleman, al demostrar el efecto que tiene la escuela en el logro educativo de los alumnos. Una de las principales diferencias con los estudios previos fue que emplearon una metodología cualitativa de investigación, al enfocarse en el estudio de casos con “escuelas prototípicas” que fueran sobresalientes.

El trabajo pionero en esta línea fue el realizado por Weber (1971)<sup>1</sup>, quien comparó cuatro<sup>2</sup> escuelas primarias que tenían un índice alto de rendimiento en lectura. El principal resultado fue haber identificado características comunes a las escuelas: un fuerte liderazgo instructivo; buen clima escolar centrado en la tarea, tranquilo y ordenado; tareas centradas en la enseñanza de la lectura; altas expectativas, enseñanza individualizada; utilización de un método fónico y evaluación constante del progreso de los alumnos; además detectó tres factores que no estaban relacionados con el alto rendimiento, como el tamaño del aula, segregación de los alumnos por capacidades e instalaciones físicas. Este estudio sustenta las bases para el desarrollo del “Movimiento de Escuelas Ejemplares”, investigaciones de profundidad en escuelas *outliers* o prototípicas (Murillo, 2003).

En este periodo de investigación, también destacan los trabajos de Dunkin y Biddle (1974), Edmonds (1979) con el modelo de los “Cinco Factores” y Rutter et al. (1979) con su famoso estudio de “Quince Mil Horas”. La principal aportación del movimiento fue dejar atrás el enfoque proceso-producto, para iniciar con el agrupamiento de cuatro tipos de variables para ser estudiadas como factores que influyen en el logro educativo: 1) Variables Presagio (Características y experiencias del profesor), 2) Variables proceso (El aula, comportamiento alumno y profesor en el aula), 3) Variables producto (Efectos a corto plazo, efectos a largo plazo) y 4) Variables contexto (Características del alumno y contexto social y escolar).

El modelo de los “Cinco Factores” de Edmonds (1979) fue un estudio de relevancia internacional que se caracteriza por la forma simplificada en la que describe los cinco factores

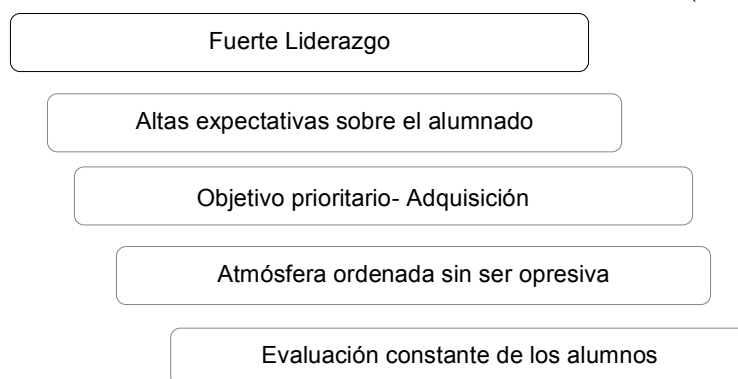
---

<sup>1</sup> Citado por Murillo (2007)

<sup>2</sup> Quedaron 4, después de haber seleccionado 17 que contaron con las condiciones para el estudio de una población de 95 escuelas propuestas.

clave en las escuelas eficaces (Véase Figura 1-1), los cuales coinciden con los que había identificado Weber (1971).

FIGURA 1-1: MODELO DE LOS CINCO FACTORES DE EDMONDS (1979)



Fuente: Elaboración propia a partir de Murillo (2003).

Rutter et al. (1979) elaboraron una investigación en Londres, con una amplia muestra de 2.000 estudiantes en 12 escuelas de educación secundaria. Este estudio fue valorado por la comunidad académica debido a la metodología que emplea, pues empieza a considerar variables que tuvieran gran variación entre los centros, a las que llamó “Características del Proceso Educativo”, con la finalidad de encontrar pruebas que refutarán los resultados del Informe Coleman. Las principales conclusiones que se extrajeron de esta investigación, fueron que las escuelas analizadas diferían considerablemente tanto en comportamiento de los alumnos como en rendimiento académico, que los centros con un mejor desempeño son aquéllos que tenían una mayor proporción de estudiantes con un coeficiente intelectual medio y que estas diferencias entre los centros, no puede ser atribuida a características físicas o administrativas de los mismos. La investigación fue titulada “Quince Mil Horas”, por ser el cantidad de tiempo que pasan los niños en la escuela (Murillo, 2003, p. 60).

El año 1979 sería clave en el desarrollo de la investigación escolar. En Estados Unidos, se lleva a cabo la primera macro investigación que se reconoce en eficacia escolar, *School Social Systems and student achievement: Schools can make a difference*, realizada por Brookover et al. (1979) cuya principal aportación es romper con el esquema de input-output, al considerar variables en el proceso, como el clima social dentro de la escuela (entendida como la opinión de alumnos,

profesores y directivos, sobre la percepción de la interacción dentro del mismo). También dentro de las respuestas (output), consideran la percepción de los estudiantes respecto a su capacidad académica y su autoconfianza.

## 1.2 La investigación en enseñanza eficaz.

Comenzando los años 80, Brophy y Good (1984) del Instituto de Investigación en Enseñanza de la Universidad Estatal de Michigan, llevaron a cabo un estudio para conocer la relación que existe entre el comportamiento del profesor y el logro de los estudiantes. Dieron seguimiento a 12 salones de clases en el periodo de 1973-1983 con énfasis en variables como contenidos cubiertos en los cursos, oportunidad de aprendizaje, expectativas de los profesores, tiempo que dedican los estudiantes, tamaño del grupo, instrucción activa del profesor y su forma de trabajo, la cual comprendía desde la presentación de la información (estructura, entusiasmo, secuencia y claridad) a la retroalimentación a los estudiantes y actividades en clase y fuera de clase (Brophy y Good, 1984):

Esta publicación inicia la etapa de Enseñanza Eficaz (ver Tabla 1-2) en la investigación escolar, pues el papel del profesor tomará relevancia como variable que explica el logro educativo de los estudiantes. Martínez-Garrido (2011) resume los principales resultados del trabajo en ocho factores:

Tabla 1-2: Factores clave de Enseñanza Eficaz detectados por Brophy y Good (1984).

Tiempo	Las escuelas más eficaces dedican más tiempo a la enseñanza directa. Mayor interacción maestro-alumno.
Fracaso Escolar	Los centros donde no se imparte una enseñanza no eficaz cuentan con mayores índices de fracaso escolar.
Expectativas de los docentes	Una enseñanza eficaz requiere unas altas expectativas sobre el progreso de los estudiantes.
Prácticas de refuerzo	A mayor reforzamiento de los contenidos, mayor será la eficacia de la enseñanza impartida.
Procedimientos de agrupación	La integración de grupos de estudiantes heterogéneos ayudará en la eficacia de la enseñanza.
Juegos Instructivos	Centros eficaces organizan frecuentemente actividades lúdicas que refuerzan los contenidos académicos.

El papel de la dirección.	Los directivos de los centros eficaces, se preocupan más por las cuestiones académicas y visitan las aulas asiduamente. En contraposición a los que se dedican a actividades burocráticas.
El compromiso del personal docente.	Mayor compromiso del personal, mayor rendimiento de los centros.

Fuente: Elaboración propia con datos de Martínez- Garrido (2011).

Slavin (1996) propuso un modelo de enseñanza eficaz basado en los elementos del “Modelo de Aprendizaje Escolar” de Carroll (1963), aquéllos que son controlables por los profesores y por las escuelas, al que se conoce como el modelo CATT (Calidad de la Instrucción, Adecuación a la Enseñanza, Incentivos y Tiempo). Estos elementos tienen una característica importante y es que todos deben estar presentes de manera adecuada para lograr una enseñanza efectiva, cada uno forma un eslabón en una cadena, si uno de ellos falla, el modelo no funcionará, pero si se trabaja con los 4 tendrá un efecto multiplicador (Slavin, 1996).

Una de las investigaciones en Enseñanza Eficaz considerada por Goldstein y Woodhouse (2000), como la primera que cumple con los requisitos mínimos necesarios para realizar inferencia válida, fue la que llevó a cabo Mortimore (2000), titulada *Inner London Educational Authority's Junior School Project (JSP)*, cuyo resultado final fue el libro *Schools Matters*. En esa investigación se buscaba analizar la variación en eficacia en los centros escolares y en las aulas, si se controlaba la entrada de los alumnos, así como ver si ciertas escuelas eran más efectivas para determinado grupo de estudiantes. También fue un referente como recomendación para que las autoridades educativas locales tomaran decisiones de mejora. Sus principales conclusiones fueron enfatizar en el desarrollo de un clima positivo dentro del aula, así como en los profesores utilizar más la recompensa que el castigo. (Murillo, 2003).

### 1.3 La investigación en eficacia educativa.

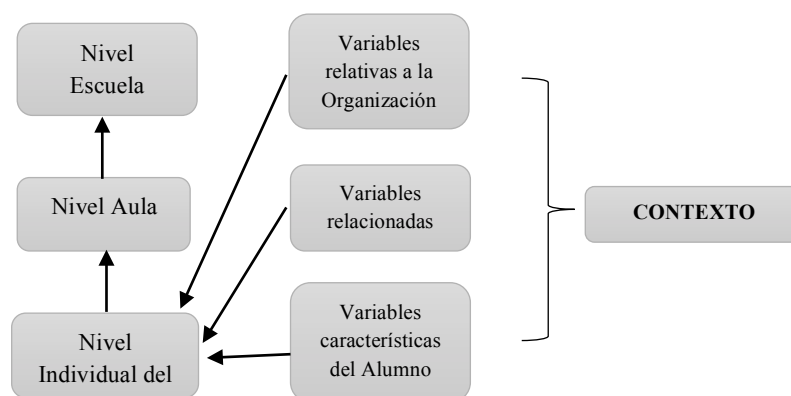
En el año 1986 fue publicado un estudio que marcaría el rumbo que actualmente ha tomado la investigación en eficacia escolar y en el que se fundamenta esta tesis. El artículo titulado *Statistical modelling issues in school effectiveness studies*, fue publicado por Aitkin y Longford (1986). Estos autores establecieron que el nivel mínimo requerido para validar la comparación entre escuelas, tendría que estar basado en un análisis a nivel individual y usando técnicas eficientes

de modelación multinivel como recurso metodológico para medir los efectos de los factores que están asociados al rendimiento de los estudiantes (Goldstein y Rasbash, 1993).

A partir de este momento, las investigaciones en el tema estarían definidas por la recopilación de grandes muestras de datos y por adentrarse en el contexto del aula y del centro escolar. También se verán influidas por el desarrollo que se le ha dado tanto nacional como internacionalmente a la evaluación de los sistemas educativos, a través de la aplicación de pruebas estandarizadas, como PISA y TIMSS. Adicionalmente, en 1991, se crea el Congreso Internacional de Eficacia y Mejora Escolar, ICSEI por sus siglas en inglés (*International Congress for School Effectiveness and School Improvement*), asociación que ha reunido a expertos sobre investigación en eficacia escolar para el intercambio académico de saberes, puntos de vista y encuentros a nivel internacional y que sienta las bases para la línea en la cual se enfocan las investigaciones actuales, la mejora de la eficacia escolar.

En este primer encuentro, surgiría otra investigación que desarrollaría un modelo que a la fecha es utilizado como referente en el planteamiento de investigaciones en Eficacia Escolar. Scheerens y Creemers (1989) plantean un marco teórico para la conceptualización de la eficacia escolar, que considera variables a nivel de escuela, el contexto de la escuela y el aula, cuyas relaciones se estructuran como se muestran en la figura 1-2.

FIGURA 1-2: MODELO PROPUESTO POR SCHEERENS Y CREEMERS (1989) PARA DEFINIR LAS VARIABLES DE CONTEXTO (NIVEL ESCUELA).



En la tabla 1-3 se resumen las principales aportaciones que ha habido en esta línea, por países, fecha y con sus principales resultados.

Tabla 1-3: Aportación de Investigaciones destacadas en Eficacia Educativa por país.

País	Publicación y fecha	Objetivo	Metodología	Resultados
Australia	The Victorian Quality of Schools Projects (Hill y Rowe, 1998).	Determinar las características de los centros en los que los alumnos mantienen un progreso más amplio y sostenido en Lengua Inglesa y Matemáticas.  Establecer las características de los centros en los cuales los alumnos presentan actitudes positivas.	Estudiaron un total de 13.909 alumnos de 59 primarias y 31 secundarias, sus familias y 903 profesores. Los datos de los alumnos se recogieron en los cursos de 1992, 2993 y 1994.  Utilizan un modelo de tres niveles.	La varianza entre las escuelas es del 18%.  La varianza entre aulas es muy alta (50%), siendo más baja para Lengua Inglesa y para Matemáticas.  La varianza entre centros escolares del 8%.  La influencia del aula es mucho mayor que la del centro.
Francia	Grisay (1997), citado por Scheerens y Bosker(1997).	Establecer las características de las escuelas y la familia relacionadas con el desarrollo cognitivo y socioafectivo de los alumnos.	Muestreo no probabilístico de 100 centros de Secundaria inferior pertenecientes a 6 distritos escolares diferentes.  Se aplicaron pruebas de rendimiento cognitivo y socioafectivo, así como cuestionarios a directivos y docentes.	El porcentaje de varianza explicado por el centro fue de 5,4% para Matemáticas y 2,6% para Lengua Francesa.  La varianza explicada por las aulas de 4,7% para Matemáticas y 1,9% para Lengua Francesa.
Reino Unido	Forging links: Effective schools and effective departments (Sammons, et al. 1997)	Analizar la eficacia en centros de secundaria y las diferencias de los resultados entre departamentos.	Se llevó a cabo en varias fases	Hay evidencia de la existencia de efectos entre departamentos, para Literatura, francés e Historia, la diferencia entre centros es mayor que con respecto al rendimiento en general.



Países Bajos	Kallestad (2000)	Determinar la existencia de diferencias entre los centros analizados y detectar cuáles son las características escolares que explican esa diferencias	Muestra de 252 centros de primaria	<p>La varianza explicada por el centro era de 11,6% del total en Aritmética, y un 8% en Lengua.</p> <p>Las variables del contexto de la escuela y de su organización, explican más varianza en lengua y en Aritmética son las variables en organización escolar las que tienen mayor impacto.</p>
Latinoamérica	<p>Laboratorio latinoamericano de Evaluación de la Calidad de la Educación Superior.</p> <p>(Abraham y Rojas, 1997; Murillo y Román, 2009).</p>	Determinar los factores asociados al logro educativo en 3° y 4° de Educación Básica en países de Latinoamérica	<p>Muestra de 48.688 alumnos, 48.088 padres o tutores, 3.675 maestros, 1.387 directivos y 1.509 establecimientos.</p> <p>Modelos Multinivel (No hay evidencia de que esta fue la muestra utilizada)</p> <p>(Murillo, 2003).</p>	<p>Escuelas más eficaces son aquellas que tienen altos niveles de recursos escolares. Aulas que no son multigrado y donde los estudiantes no están agrupados. Aulas en donde los alumnos son evaluados frecuentemente, Aulas con altos niveles de participación de los padres de los estudiantes. Aulas con clima positivo, especialmente relacionado con la disciplina.- Estudiantes mujeres tuvieron un rendimiento menor en matemáticas que los hombres (Moreira, 2009).</p>

Iberoamérica	La investigación sobre Eficacia Escolar en Iberoamérica IIEE Convenio Andrés Bello (, 2008).	Describir los efectos escolares, de aula y de país para educación primaria en Iberoamérica.  Encontrar los factores de eficacia escolar, de aula y de país para educación primaria en Iberoamérica	Participaron 9 países (Bolivia, Chile, Colombia, Cuba, Ecuador, España, Panamá, Perú y Venezuela) con una muestra p de 98 escuelas (10 por país en promedio, salvo Perú con 20), 248 aulas y 5.603 alumnos.  Modelo de cuatro niveles (alumno, aula, escuela, país).  Estudio de eficacia percibida por distintos miembros de la Comunidad Escolar.  Estudio de profundidad, escuelas prototípicas.	El porcentaje de varianza del logro académico cognitivo del alumno explicado por estudiar en una escuela u otra es del 10%. El efecto del aula y del país, es diferente en función de la asignatura. En Matemáticas, el efecto aula es del 22% y 15% el efecto país; en Lengua, el efecto aula es del 11% y el efecto país del 2%.
--------------	--	--	---	--

Fuente: Elaboración propia con datos de Peña Suárez et al. (2012)

#### 1.4 La investigación educativa en España.

##### 1.4.1 Organismos para la investigación educativa.

La investigación educativa en España se ha ido diversificando a la par de los cambios que ha tenido su sistema de educación. A partir de la promulgación de la Constitución española de 1978, se aprobaron tres grandes leyes que regulan el sistema educativo no universitario: La Ley Orgánica Reguladora del Derecho de Educación (LODE) en 1985, la Ley Orgánica de Ordenación General del Sistema Educativo (LOGSE), cinco años después, y la Ley Orgánica de la Participación, la Evaluación y el Gobierno de los centros docentes (LOPEG) en 1995 (Muñoz-Repiso y Murillo, 2000). Las características principales de estas reformas fueron: El modelo de descentralización adoptado, en el que se transfieren las competencias a las

Comunidades Autónomas; la autonomía de los centros docentes en cuestiones económicas, curriculares y de organización para gestionar sus procesos internos; la ampliación de la escolarización obligatoria hasta los 16 años; un nuevo modelo de dirección; y la formación del profesorado.

Sin embargo, antes de todas estas reformas, España había recorrido un largo camino en investigación educativa. Tejedor (1988)<sup>3</sup>, menciona que muy probablemente, el Museo Pedagógico creado en 1882, fue la primera institución en realizar investigación en educación en España, teniendo como objetivo la pedagogía experimental. A partir de 1970, se comenzaría a consolidar la investigación en esta área a través de los planes de investigación. En ese año, se crearía el Centro Nacional de Investigaciones para el Desarrollo de la Educación (CENIDE), quien tendría como propósito la formación del profesorado. Posteriormente, se crearía en 1974, el Instituto Nacional de Ciencias en la Educación (INCIE), que sustituiría al CENIDE, ya como organismo autónomo con personalidad jurídica propia, pero que continuaría con el mismo objetivo, con la diferencia de que inicia investigaciones propias que no se habían realizado anteriormente, con un enfoque más sociológico, así como el otorgamiento de becas y la creación del Premio Nacional de Investigación Educativa, significando un incentivo para el desarrollo de investigaciones en el tema educativo en España. A finales de 1980, este organismo desaparece y sus funciones son absorbidas por la administración del estado, creándose en 1982 la Subdirección General de Investigación Educativa (SGIE), una de cuyas misiones era fomentar en todo el sistema educativo, la investigación educativa aplicada (Carmena et al., 2000).

Posteriormente, este instituto se convirtió en el Centro de Investigación y Documentación Educativa (CIDE), creado en 1983 y que se estructuró originalmente en tres funciones: la Investigación, la Documentación y la Evaluación. En la primera tenía el propósito de realizar investigación sobre los problemas educativos del sistema, que ayudara a la toma de decisiones en política educativa. La parte de Documentación, se encargaría de la recopilación de bases de datos y archivos del ámbito educativo. Finalmente, la función de Evaluación tendría la misión de valorar la implantación de las reformas en el sistema educativo. Con el paso del tiempo, sus

---

<sup>3</sup> Citado por Carmena et al. 2000

funciones se fueron diversificando y aumentando para incluso cambiar su nombre a Centro Nacional de Investigación e Innovación Educativa (CNIIE). Sus actividades más importantes en la actualidad, continúan siendo la realización de investigación propias en el ámbito educativo, incentivar con diferentes convocatorias y becas el desarrollo de la investigación, así como la elaboración de informes sobre el funcionamiento del sistema de Educación en España y la divulgación científica de la investigación a través de diferentes publicaciones, entre las que destacan las siguientes líneas: Informes sobre el sistema educativo, estudios sobre la transición de la educación secundaria a la superior, estudios sobre la eficacia y calidad del sistema educativo, estudios sobre la equidad del sistema educativo y estudios sobre la investigación educativa.

En el año de 1990, se crea el Instituto Nacional de Calidad y Evaluación, adscrito al Ministerio de Educación Cultura y Deporte, que en 2012 pasa a denominarse Instituto Nacional de Evaluación Educativa por el Real Decreto (257/2012). Este organismo es el responsable de la Evaluación del Sistema Educativo Español y ha coadyuvado en el desarrollo de investigaciones para la mejora del sistema educativo.

#### 1.4.2 Estado del Arte de la Investigación Educativa en España.

La reacción hacia los resultados arrojados por el Informe Coleman, fue lo que provocó el desarrollo de la línea en eficacia escolar a nivel internacional, se tenía el propósito de demostrar que la escuela sí tiene un papel activo en el aprendizaje de los estudiantes y qué factores escolares y de aula están relacionadas con el rendimiento de los alumnos. No obstante, previo a estas investigaciones ya se habían hecho trabajos al respecto en España.

Murillo (2003) menciona que el trabajo pionero en eficacia escolar fue la tesis del profesor de Psicología Manuel Millán de la Facultad de Filosofía y Letras de la Universidad de Valencia, titulada "*La eficacia de la educación escolar*" realizada en 1978, quien determinará las pautas que seguiría este proceso en los trabajos subsecuentes, al señalar que la investigación educativa estaba siendo abordada desde cinco perspectivas diferentes: input-output, psicológica, organizativa, pedagógica y escuelas alternativas, enfoques similares a los que marcó la investigación escolar internacional.

Es importante mencionar, que la mayoría de estas investigaciones fueron desarrolladas dentro de las Universidades, como trabajos de tesis doctoral, sin llegar a la divulgación científica en revistas internacionales, por lo que es difícil acceder a una revisión exhaustiva. Sin embargo, Murillo (2003) seleccionó 10 investigaciones importantes en el tema que se resumen en la tabla 1-4.

Como se puede observar en la tabla, después del trabajo de Millán (1978), Fuentes (1986)<sup>4</sup> llevó a cabo una investigación sobre los factores de eficacia escolar, en la que utilizó modelos causales, metodología validada en el momento para este tipo de estudios. Se enfoca en grupos de educación básica (EGB) y utiliza información de 91 escuelas y 615 profesores de segunda etapa EGB. Sus principales aportaciones son los factores encontrados que tienen un efecto en los resultados educativos, como el origen social del alumnado, el papel del profesor en el aula y el clima del centro escolar. Se corroboran los resultados de investigaciones internacionales como la de Rutter et al. (1979) y Brookover et al. (1979).

Después, Rodríguez Díez (1990), llevaría a cabo una investigación input- proceso- output, para evaluar los centros escolares, utilizando también modelos causales. Toma como variables producto, las calificaciones del alumnado, el número de repetidores, el número de alumnos con problemas de aprendizaje; como variables de entrada, el tipo de centro, características del alumno y del profesorado y variables de proceso, el clima del aula. Fueron varias las interacciones que encuentra entre ellas, pero los principales resultados hallados, es la relación positiva entre el clima del centro y el trabajo de los docentes en el aula, que a su vez, influye en el funcionamiento del equipo directivo, y cómo el grado de satisfacción del alumno depende de su nivel socioeconómico, siendo menor en los de estatus más alto.

García Durán (1991) realizó una investigación similar a la de Weber (1971) y Dunkin y Biddle (1974), en el que estudia escuelas prototípicas extremas para identificar casos de éxito o fracaso, al utilizar centros rurales y urbanos de alto y bajo rendimiento. Entre sus principales aportaciones destacan, encontrar que los centros escolares son más eficaces si tienen un plan de centro, funciona adecuadamente el equipo directivo, existe cooperación con las familias y producen sus propios recursos didácticos.

---

<sup>4</sup> Citados por Murillo (2003)

En el mismo año, Rodríguez Gómez (1991) llevó a cabo una investigación, similar en el objetivo, al evaluar centros escolares realizando un estudio input-proceso-output. La variable producto fue el rendimiento del alumnado, controlado por la variable clase social. Las variables analizadas fueron características del profesorado, el clima organizacional, el clima del centro, las relaciones con la comunidad y el funcionamiento de los órganos directivos. Dentro de sus principales resultados, se encuentra el concluir que las escuelas difieren en su nivel de eficacia independientemente del origen social de sus alumnos<sup>5</sup> (Murillo, 2003, p.327).

Siguiendo esta línea, Barrueco (1992) realizó su tesis, sobre calidad y eficacia de los centros escolares, usando un modelo causal explicativo de tres productos, tomando en cuenta las calificaciones escolares, la formación general del alumno y el nivel de satisfacción de los principales actores (padres, profesores y alumnos). Obtiene un modelo de cinco factores, similar al propuesto por Edmonds (1979) (véase Tabla 1-3).

Un estudio considerado como de los mejores sobre eficacia escolar, fue el de Castejón (1994). Su objetivo fue conocer los factores procesuales del centro educativo que contribuyen a explicar las diferencias del rendimiento entre escuelas. Los puntos por los cuales destaca la investigación son: 1) la cantidad de información que utiliza para el análisis, pues se recopiló una gran cantidad de variables de 1.924 alumnos y 24 centros de bachillerato, 2) la metodología empleada, al hacer uso de análisis de regresión múltiple y correlaciones y 3) los resultados, al obtener índices para medir la eficacia de los centros y encontrar los principales factores que influyen en el rendimiento, resaltando que los años de experiencia del docente, influyen de manera negativa.

Entre 1992 y 1995, el Centro de Investigación y Documentación Educativa (CIDE), participó en un estudio internacional dirigido por la OCDE (Organización para el Comercio y Desarrollo Económico) y que tenía como propósito medir la efectividad de los centros escolares y la administración de los recursos educativos. Para ello, utilizó información de 24 centros, tanto de escuelas primarias como de secundarias. Se realizaron entrevistas y encuestas a directivos, profesores, familias y alumnos, recopilando información de 542 variables. Los principales resultados se resumen en que se deben estudiar por separado los centros de

---

<sup>5</sup> Véase Tabla 1-3.

primaria y los de secundaria y los factores identificados en cada nivel, se detallan en la Tabla 1-4. Galán (2000), realizó un estudio para mostrar la importancia del clima escolar en los centros escolares, al tomar esta variable como producto y no como proceso. La metodología que emplea es un modelo causal de evaluación (De estructuras de covarianza). Utiliza variables de escala tipo Likert, con una muestra de 1.024 profesores y directivos de 79 centros escolares de nivel infantil, educación primaria y secundaria.

En cuanto a la línea de investigación en eficacia docente, se llevó a cabo un estudio cuyo objetivo fue determinar el perfil de cualidades necesarias del profesor de Educación Secundaria para el buen ejercicio de su tarea docente. Rodríguez Pérez (1984), realizó un estudio para evaluar la eficacia escolar, encontrando factores de ineficacia docente como el bajo estatus socioeconómico del profesorado, el exceso de mujeres desempeñando esta profesión (feminización), el escaso tiempo dedicado a enseñar y la resistencia al cambio, que influían de manera negativa en los resultados educativos.

Tabla 1-4: Principales trabajos de investigación en eficacia escolar en España.

Investigación y Autores	Metodología	Aportes
Fuentes (1986)  Procesos funcionales y eficacia de la escuela. Un modelo causal.	Modelos causales.  615 profesores de EGB  91 centros escolares	La eficacia de los centros escolares es el resultado del “funcionamiento global” de los mismos, más que de influencias aisladas de las variables y factores particulares.  Los resultados educativos dependen de:  a) Origen social del alumnado b) Papel del profesor en el aula c) Clima del centro escolar  “Los centros eficaces son aquéllos que han logrado “una cultura de centro” facilitadora del rendimiento individual y de la implicación de todos en una tarea común.
Rodríguez Díez (1990)  Modelo para la evaluación externa de los centros escolares de EGB.	Modelo causal.  Para la evaluación de centros escolares.	El clima del centro influye positivamente en el trabajo de los docentes en el aula e indirectamente en el rendimiento del alumnado.  El grado de insatisfacción de los alumnos con el centro, es menor en los de estatus

	<p>Variables:</p> <p>Entrada: Tipo de centro, características de alumnos y maestros.</p> <p>Proceso: Modelo organizativo del centro. Clima del aula.</p> <p>Producto: Calificaciones, No. de repetidores, No. alumnos c/problemas de aprendizaje, Satisfacción de las familias, Índice de absentismo escolar, solicitudes de ingreso que recibe el centro.</p>	<p>socioeconómico más alto.</p> <p>La estabilidad docente influye positivamente en el funcionamiento del equipo directivo.</p>
<p>García Durán (1991).</p> <p>Investigación evaluativa sobre las variables pedagógicas que discriminan entre los centros EGB de alto y bajo rendimiento en el medio rural de la provincia de Cádiz.</p>	<p>Casos prototípicos extremos.</p> <p>Centros rurales y urbanos de alto y bajo rendimiento.</p>	<p>Los centros eficaces son los que hacen una revisión del plan del centro y consejo escolar, funcionamiento del equipo directivo, orientación escolar, cooperación con las familias, producción de recursos didácticos.</p>
<p>Rodríguez Gómez (1991).</p> <p>Investigación evaluativa en torno a los factores de eficacia escolar de los centros públicos de EGB.</p>	<p>Variables: Características del profesorado, clima organizacional, clima del centro, funcionamiento de los órganos directivos y rendimiento de los alumnos controlados por la variable clase social.</p>	<p>Las escuelas difieren en su nivel de eficacia independientemente del origen social de sus alumnos.</p> <p>Escuelas eficaces aquéllas con:</p> <ol style="list-style-type: none"> <li>a) Profesores satisfechos.</li> <li>b) Participación de padres. Clima escolar positivo.</li> <li>c) Trabajo en equipo y liderazgo educativo.</li> </ol>
<p>Barrueco (1992)</p> <p>Calidad y eficacia de los centros educativos: análisis de las aportaciones de un modelo explicativo.</p>	<p>Modelo causal de tres productos.</p> <ol style="list-style-type: none"> <li>1) Calificaciones escolares</li> <li>2) Formación general del alumno.</li> <li>3) Nivel de satisfacción principales actores (padres, profesores y alumnos).</li> </ol>	<p>Cinco factores de eficacia escolar</p> <ol style="list-style-type: none"> <li>1. Naturaleza de la comunidad educativa</li> <li>2. Autocontrol participativo</li> <li>3. Gestión técnica de la escuela</li> <li>4. Cualificación y preparación docente del profesorado.</li> <li>5. Planificación didáctica.</li> </ol>
<p>Castejón (1996)</p> <p>La eficacia de centros de</p>	<p>Modelo de regresión múltiple y correlación.</p> <p>Factores procesuales del centro educativo que</p>	<p>Elaboración de índices para medir la eficacia de los centros.</p>



enseñanza secundaria: Un modelo de identificación y funcionamiento	<p>contribuyen a explicar las diferencias de rendimiento entre escuelas.</p> <p>24 Centros de Bachillerato.</p> <p>1.924 alumnos.</p> <p>Recopilación de información de gran cantidad de variables relativas a docentes, alumnos y directivos a través de la aplicación de cuestionarios.</p>	<p>Factores que influyen en el rendimiento:</p> <ul style="list-style-type: none"> <li>a) Percepción de los alumnos sobre el centro (positiva)</li> <li>b) Los años de experiencia del docente Influyen de manera negativa en el rendimiento.</li> <li>c) Satisfacción de los profesores y su relación con el director (positiva).</li> </ul> <p>“La influencia de los procesos educativos generados en los centros tiene un efecto moderado sobre el rendimiento pero significativo”.</p>
CIDE (1992)	<p>Estudio de carácter ex post ipso.</p> <p>24 centros escolares de primaria y secundaria.</p> <p>Entrevista y cuestionario a directivos, profesores, familias y alumnos.</p> <p>542 variables de contexto/ entrada, proceso y producto (Satisfacción, repetición y tasas de éxito)</p>	<p>Diferencias entre nivel primaria y secundaria, sin posibilidades de elaborar un modelo único.</p> <p>Primaria: El clima del centro educativo es el factor clave y representa el mejor predictor de las variables consideradas como producto.</p> <p>Secundaria: No encontraron una única dimensión que represente el grado de eficacia de un centro. (Rendimiento asociado al contexto del centro y Satisfacción de los miembros asociado a la calidad de las relaciones).</p>
González Galán (2000).  Calidad, eficacia y clima en centros educativos. Modelos de evaluación y relaciones causales.	<p>Estudio del clima escolar como característica fundamental de la calidad de los centros.</p> <p>Modelo causal de evaluación de la calidad centrado en el clima de trabajo.</p> <p>1.024 profesores y directivos</p> <p>78 centros escolares de infantil, primaria y secundaria.</p>	<p>Considera al clima escolar como variable de producto y no de proceso.</p> <p>Demuestra la importancia del clima de trabajo en los centros escolares</p>

Fuente: Elaboración propia a partir de Murillo (2003).

### 1.4.3 Investigaciones actuales en eficacia escolar en España.

Resulta una tarea compleja, delimitar los alcances de los enfoques de las investigaciones escolares. Sin embargo, como señala Martínez-Garrido (2011), a partir del año 2000, el término eficacia escolar englobaría los tres ejes principales que había tomado la investigación: estudios input-output, estudios de enseñanza eficaz y los estudios propiamente de eficacia escolar.

En España, a partir de esta fecha, se comienza a adoptar la modelación multinivel como estrategia de análisis para la identificación de factores asociados al rendimiento escolar. De las primeras investigaciones realizadas, se encuentra la de Murillo (2000), quién ha sido uno de los especialistas en el tema y el que más contribuciones tiene a la fecha al respecto. Murillo (2000), en su publicación: Los modelos jerárquicos lineales aplicados a la investigación sobre eficacia escolar, hace un ejercicio en el que utiliza los mismos datos que Castejón (1996), pero aplica una Modelación Multinivel para comparar las ventajas y desventajas de los resultados, frente a la técnica de Regresión Múltiple. Su principal conclusión es que los modelos multinivel son la estrategia más adecuada de análisis para este tipo de datos, pues se ajusta mejor al explicar en mayor medida la parte aleatoria del modelo. También, menciona que se tiene una mejor comprensión de esta variabilidad al conocer cuál es la fuente principal de variación, si el alumno, el aula o el centro escolar. En sus resultados, Castejón (1996) muestra que el centro tiene poco efecto y el del aula es un poco mayor. Posteriormente, en el año 2003, realizará una reflexión y una revisión exhaustiva sobre lo que se había realizado hasta entonces en materia de investigación escolar en España, resultados que presenta en su publicación La investigación sobre eficacia escolar en España, en el que sostiene como tesis principal la necesidad de continuar esta línea de investigación y mejorar en el aspecto metodológico. (Murillo, 2003).

En el año 2002, publicaría en la revista *Educational Research and Evaluation*, el artículo Good Effective School Improvement Practices in Spain, una investigación de estudios de caso, en el que seleccionó cinco centros escolares para evaluar la efectividad de los programas de mejora de la Escuela. De sus principales resultados, identifica cinco características del sistema de educación español y que tienen un efecto en el desarrollo de los programas de mejora de la

escuela: la dirección escolar, la cultura de equidad, el liderazgo en los centros escolares españoles, la participación e implicación de las familias y la cultura del centro. (Murillo, 2002).

Continuaría sus aportes en el tema, con su participación como coordinador de la Red Iberoamericana de Investigación sobre Eficacia Escolar y Mejora de la Escuela (RIEME) y el desarrollo de la Investigación Iberoamericana sobre Eficacia Escolar (IIEE), ambas iniciativas del Centro de Investigación y Documentación Educativa (CIDE) del Ministerio de Educación de España. De este trabajo conjunto saldría una primera publicación titulada *La investigación sobre eficacia escolar en Iberoamérica*, que estaría financiada por el Convenio Andrés Bello. (Murillo, 2003) y en la que se presentaría principalmente una revisión del estado del arte del tema de los países participantes: Argentina, Bolivia, Brasil, Chile, Colombia, México, Cuba, Ecuador, Panamá, España y Venezuela. Continuando con el trabajo de la red, cuatro años más tarde, se llevaría a cabo una investigación sobre eficacia escolar, aplicando un modelo multinivel de cuatro niveles, en el que se incluiría un cuarto nivel, el país. Recopilaron información de 98 escuelas, 248 aulas y 5.603 alumnos de los 9 países participantes. En promedio, cada país recopiló información de 10 escuelas, 20-30 aulas y 622 alumnos (excepto Perú quien participó con 20 escuelas, 48 aulas y 1,566 alumnos). Para homogeneizar los datos, la información de las aulas correspondió al tercer grado de las escuelas de educación primaria. Fue un estudio muy completo que incluyó no solo modelación estadística, sino también entrevistas, estudios de caso y análisis cualitativo de datos. Para fines de este trabajo, sólo se mencionarán las aportaciones respecto a la modelación y los factores encontrados<sup>6</sup>.

Las variables a explicar consideradas fueron tres: El rendimiento del alumno en lingüística, en matemáticas y un test de autoconcepto. Las primeras fueron desarrolladas por el Laboratorio Latinoamericano de Evaluación de Calidad de la Educación (LLECE), pues ya eran resultados validados, y para el test de autoconcepto utilizaron la prueba de SDQ-I (Self description questionnaire). Para cada variable respuesta, ajustaron un modelo multinivel. Los factores considerados como variables explicativas, fueron tomados de los cuestionarios aplicados a los alumnos, a los profesores y a los directivos de los centros escolares seleccionados (Murillo, 2008).

---

<sup>6</sup> Véase Murillo (2008) para profundizar en los resultados del estudio.

Los principales resultados de esta investigación fueron los coeficientes de variación por competencia. En el caso del rendimiento en matemáticas, el efecto del aula fue del 22% y para lengua del 11%; el porcentaje de la varianza explicada por estudiar en una escuela u otra fue de 10% y por el país al que pertenece la escuela, fue mayor en el rendimiento en matemáticas que en lengua (15% y 1% respectivamente).

Los factores que resultaron significativos fueron:

- El género: Los alumnos varones obtienen mejor rendimiento en matemáticas, mientras que sus compañeras mujeres, sacan mejores puntuaciones en lengua, con una diferencia muy similar. Las niñas obtienen 3,16 puntos más en lengua y 3,22 puntos menos en matemáticas que sus compañeros varones (Murillo, 2008, p.122).
- Mayor nivel socioeconómico y cultural de las familias del alumno, mayor será su rendimiento en ambas competencias.
- Los alumnos nacidos fuera del país donde estudian obtuvieron resultados más bajos en matemáticas, que los compañeros nativos.

Por su parte, Calero y Escardíbul (2007) realizaron una investigación para evaluar el rendimiento de los centros educativos en España, de acuerdo a su titularidad, es decir comparando los resultados de los centros privados y los centros públicos. Utilizaron datos de la evaluación de PISA 2003 y aplicaron un modelo de dos niveles, considerando a los estudiantes en el primer nivel y a las escuelas en el segundo nivel, con una muestra de 10.791 alumnos y 383 centros evaluados. La variable respuesta que consideraron, fue el puntaje obtenido por el alumnado de educación secundaria en la prueba de matemáticas. Su principal conclusión es que “las diferencias de puntuaciones que se observan a favor de los centros privados no vienen explicadas por la titularidad del centro, sino por otras variables referidas a los usuarios (individuales y familiares) y al propio centro” (Calero y Escardíbul, 2007, p. 24).

Choi de Mendizábal y Calero-Martínez (2013) publicaron los resultados de un trabajo de investigación sobre los determinantes del fracaso escolar en España usando los resultados de la prueba PISA 2009. Aplicaron un modelo de regresión logística de dos niveles (alumno y

escuela), para determinar la probabilidad de riesgo de que el alumno obtenga resultados desfavorables de acuerdo a un conjunto de variables. La muestra la conformaron 433.224 jóvenes de 15 años y 889 centros escolares. Una de las principales aportaciones de esta investigación fueron las variables que resultaron significativas para explicar el fracaso del estudiante, dentro de las propias de los alumnos se encontró el género, si había sido repetidor y si tenía o no previa educación infantil. Respecto a las de contexto, las variables relativas al hogar fueron las más significativas. En cuanto a su contexto familiar, las variables determinantes son el origen de los padres, la situación de inmigrante y los recursos educativos del hogar, es más probable que el alumno fracase si no inicia a una edad temprana la escuela o si ha repetido un curso, Por otro lado, las probabilidades de que termine la ESO son muy bajas, en comparación con aquéllos estudiantes que realizan una educación infantil, así como los alumnos de nacionalidad distinta a la española.

(2012) realizó su tesis doctoral sobre los factores de eficacia escolar en el programa PISA, utilizando modelación multinivel y planteando un modelo para cada los resultados del alumno en competencias en ciencias, matemáticas y lectura. Sus resultados confirman el efecto del centro en el desarrollo de estas competencias, así como de características individuales como el sexo, procedencia, el ser repetidor y el contexto familiar, como el estatus socioeconómico y principalmente, la motivación del alumno.

En resumen, se puede afirmar que existe un gran debate entre las expectativas sobre los resultados de las investigaciones en el ámbito educativo en España. Carmena et al. (2000, p. 20) describen esta situación de manera muy puntual “La investigación educativa no puede dar respuesta a problemas de carácter normativo, y tampoco puede producir conocimientos que respondan a las necesidades de los responsables políticos y educadores. Lo que sí puede hacer es generar información que contribuya a identificar problemas educativos y aumentar la base de conocimientos en la que políticos, administradores y educadores fundamentan sus decisiones y acciones”.

---

Aunque el boom de la investigación escolar en España, comenzó hace aproximadamente 15 años, los conocimientos que se han generado sobre el rendimiento educativo y los factores que influyen en él han sido vastos. Asimismo se ha visto en los últimos años una mejora en la metodología aplicada que permite sustentar mejor los resultados.

## Capítulo 2

# Modelización Multinivel en investigación escolar.

## Capítulo 2: Modelización multinivel en investigación escolar.

### 2.1 Los modelos multinivel

Los modelos multinivel tienen el propósito de analizar los datos considerando la estructura jerárquica de los mismos, al modelar la realidad con la existencia de diferentes niveles de variación (Rasbash et al., 2009). Por dicha razón, se aplican principalmente a datos que presentan una estructura jerárquica, es decir que se encuentran estructurados en un cierto número de niveles o clasificaciones. Estas condiciones permiten tener una mejor comprensión de la variabilidad de los datos, pues se logra conocer la varianza entre las unidades de un mismo grupo o nivel y la covarianza entre los grupos o niveles. También son conocidos como Modelos de Componentes de la Varianza (Dempster et al., 1981; Aitkin y Longford 1986), Modelos de Coeficientes Aleatorios (Kreft y de Leeuw, 1994), Modelo Lineal Jerárquico, o Hierarchical Linear Model (HLM) (Bryk y Raudenbush, 1992), Modelos Multinivel (Goldstein, 2007) y Modelos de Efectos Mixtos (Laird y Ware, 1983). La modelación multinivel ha adquirido especial atención desde finales de la década de los ochenta, aunque sus orígenes se remontan varios años atrás. Estos modelos fueron diseñados para analizar un fenómeno a partir de una o varias variables respuesta, considerando variables explicativas de diferentes niveles simultáneamente, para lo que se plantea y ajusta un modelo estadístico que apropiadamente incluye las diversas dependencias en los diferentes niveles (Bryk y Raudenbush, 1992).

El interés que provocó el desarrollo de la modelación multinivel en la comunidad científica, ha acelerado su aplicación en diferentes disciplinas, tales como la Sociología, donde se introdujo el concepto de efecto contextual en este campo (Hox, 2002), en la Medicina con el Meta-análisis, los estudios de medidas repetidas y curvas de crecimiento en las Ciencias del comportamiento (Laird y Ware, 1983), entre otras. Como señala Bryk y Raudenbush (1992), cuando se combinan con la gran cantidad de software disponible, esta expansión en la modelación ha inspirado toda una serie de nuevas aplicaciones y sin duda, en el ámbito educativo han sido muy recurridos.



## 2.2 Comparación de los modelos multinivel frente a otras técnicas de análisis.

La importancia que tomó la modelación multinivel en la investigación escolar, fue la demostración de su efectividad por Aitkin y Longford (1986) frente a otras técnicas de análisis utilizadas hasta el momento, pues las primeras metodologías empleadas no consideraban esta estructura jerárquica en los datos analizados.

Los primeros trabajos de carácter cuantitativo, aplicaron la metodología como Input-Output (Brookover et al., 1979) y posteriormente, como Entrada-Proceso-Producto (Schereens y Creemers, 1989), esta técnica de análisis fue desarrollada por Leontief (1936), y se trata de una matriz que analiza las relaciones de interdependencia entre un conjunto de variables de entrada, mediante una serie de ecuaciones lineales. Es una técnica nacida en la disciplina económica, que fue aplicada al ámbito educativo. Su objetivo era relacionar los <<inputs>> al sistema educativo, como lo son los recursos, con los <<outputs>> o los logros educativos del alumno. Como señala Báez de la Fe (1991), “la crítica esencial a este enfoque puede resumirse en el cuestionamiento del modelo economicista implícito: no se trata tanto de si la escuela logra lo que debería alcanzar, cuanto de cómo se relacionan sus recursos con sus resultados”. Además, el que comparar la escuela con una entidad económica, queda en completa desventaja, pues sus propósitos son distintos.

También se aplicaron modelos de regresión múltiple, pero estudiando las unidades sin considerar su estructura de anidamiento. Sin embargo, con esta omisión se incurría en un problema conceptual y metodológico. Hox (2002) identificó dos problemas por ignorar la estructura jerárquica de los datos: pérdida de información y análisis menos robustos. Las pruebas estadísticas ordinarias, tratan los valores de los datos desagregados como valores independientes de la muestra, lo que origina que los errores estándar sean pequeños y esto a su vez conduce a pruebas de hipótesis significativas, cuando realmente no lo son.

Los efectos que el grupo puede tener en las unidades han sido considerados por muchos investigadores, sin embargo, lo han llevado a cabo de tal manera que pueden cometer un error. Como señala Hox (2002) al interpretar los resultados de grupo como si se aplicara a los individuos se puede incurrir en una “falacia ecológica”. En otras palabras, se comete esta falacia al elaborar conclusiones a nivel individual considerando información agregada. De igual

forma, cuando se infieren relaciones a nivel grupal de relaciones que existen a nivel individual, se comete “falacia atomística” (Alker, 1969)<sup>7</sup>. En algunos casos, se incluía la variable grupo como variable dummy, en un modelo de regresión de un solo nivel, sin embargo, cuando se deseaba estudiar numerosos grupos (como en el caso de las escuelas), la cantidad de variables dummy (n-1) que se debían introducir al modelo eran excesivas, rompiéndose con esto el principio de parsimonia de los modelos (Ver Tabla 1-1.)

Por otro lado, los efectos de las variables explicativas a nivel de grupo no pueden ser estimados simultáneamente utilizando los residuos del agrupamiento (Rasbash et al., 2009), ni es posible calcular un solo parámetro que refleje esta información. Las técnicas usuales no están diseñadas para dividir la variación de esta manera y sólo estiman un término para explicar esta diferencia, al que se le denomina error  $\varepsilon_i$ . En la modelación multinivel esta variación presenta una estructura relevante susceptible de ser analizada y que aporta mucha información al problema, lo cual se considera una de las contribuciones más significativas de este tipo de metodología, porque a diferencia de la regresión ordinaria, donde solo se puede estimar la media de la variable respuesta dado un conjunto de variables explicativas, en los modelos multinivel se modela simultáneamente la varianza, por lo que se llega a conocer la variabilidad dentro de las unidades de estudio de un mismo nivel y entre grupos.

Tabla 2-1: Ventajas y desventajas de las técnicas convencionales de regresión o de un solo nivel.

TÉCNICA ESTADÍSTICA EMPLEADA	VENTAJAS	DESVENTAJAS
1. Modelos convencionales de regresión aplicados a observaciones en un primer nivel ignorando la estructura jerárquica.	Una de las técnicas más aplicadas (Familiaridad de los investigadores)	Se ignora la potencialidad de los efectos que ocurren en los niveles más altos. La estimación de los errores estándar puede ser incorrecta. Puede conducir a la falacia atomística (Alker, 1969).
2. Modelos convencionales de regresión aplicados a datos agregados en el nivel superior.	Se elimina el problema de la no independencia de las observaciones	Se pierde información. No hay posibilidad de evaluar los efectos que ocurren en los niveles más bajos. Se puede cometer falacia ecológica.
3. Uso de variables dummy para representar unidades de nivel más alto.	Permite una evaluación limitada de los efectos que ocurren en nivel superior.	¿Cuántas variables dummy se deben agregar?

Fuente: Rasbash et al. (2009)

<sup>7</sup> Citado por Hox(2002)

El uso de una técnica de análisis de regresión que no considera los efectos del grupo, fue lo más criticado al informe Coleman y lo que se presume pudo haber provocado que la escuela no resultara tener un efecto significativo. Es decir, se incurrió en un problema metodológico.

Otra técnica utilizada en esta etapa de la investigación escolar, fue el análisis de Covarianza (Rutter et al., 1979), sin embargo, como señala Murillo (2000), una de las limitantes de esta metodología en este campo, es la exigencia de que las unidades de análisis estén asignadas aleatoriamente a los tratamientos y en algunos casos el tener el mismo número de unidades para cada uno. En el caso de los modelos multinivel, este requisito no es necesario, se puede trabajar con grupos o escuelas, cuyo número de alumnos varíe, es decir con datos no balanceados y los estimadores continuarían siendo robustos (Maas-Cora y Hox, 2004).

Por otro lado, una gran cantidad de trabajos recurrió al análisis causal, que tuvo como especial controversia en su aplicación, la selección de variables educativas como causa y como efecto, así como el problema de la asignación aleatoria de los casos.

### 2.3 Contribuciones de los modelos multinivel a la investigación escolar.

En primer lugar, se puede decir que 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. Esta situación permitió que las primeras investigaciones realizadas con esta metodología demostraran el gran efecto que tiene la escuela o el aula, en los logros de los resultados de los alumnos. No obstante, la magnitud de los efectos escolares estimados por los modelos multinivel es un tema aún de controversia (Murillo, 2008, p.35), pues las diferentes investigaciones aplicadas, han obtenido diferentes efecto. Se puede decir que, la mayoría de los trabajos han arrojado que el efecto del grupo oscila entre un 10-15% (Rasbash et al., 2009).

Un segundo punto es que al utilizar la modelación multinivel, se deja de lado el problema de decidir quién será la unidad de estudio, si el alumno o la escuela, pues permite analizar sus efectos simultáneamente, en comparación frente a las técnicas clásicas de regresión múltiple, en las que la decisión implicaba la inclusión del factor grupo o individuo, con problemas de carácter metodológico.

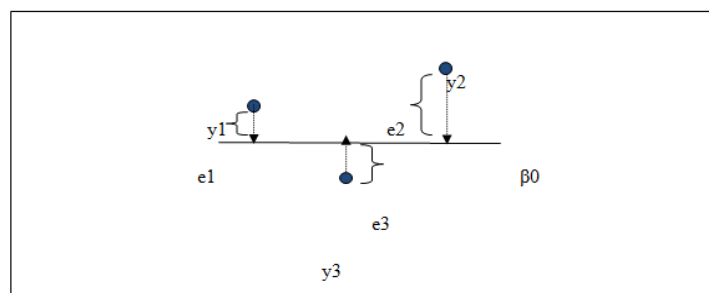
## 2.4 Estructura y notación del modelo multinivel.

El modelo multinivel busca estimar los parámetros desconocidos (intercepto y pendiente), pero además la varianza dentro de un grupo  $\sigma_e^2$  y la varianza entre los grupos  $\sigma_{u0}^2$ . La estimación de los coeficientes puede realizarse a través de diferentes enfoques como el de Máxima verosimilitud o Estimación bayesiana, y utilizando diversos algoritmos como el de Mínimos Cuadrados Generalizados Iterativos (MCGI) (Goldstein, 1996), el de Fisher-Scoring y el algoritmo EM (Longford, 1987). Actualmente, existen diversos paquetes estadísticos para el cálculo de los coeficientes. Antes de presentar el modelo multinivel, se parte del modelo de regresión ordinaria más simple. En un modelo de regresión ordinaria para un solo nivel, sin considerar variables explicativas, la ecuación es:

$$y_i = \beta_0 + \varepsilon_i \quad (1)$$

Donde  $y_i$  es el valor que toma la variable respuesta para la  $i$ -ésima observación ( $i = 1, 2, \dots, n$ ), el intercepto o  $\beta_0$  representa el promedio de  $y$  (*variable respuesta*) en la población, y  $\varepsilon_i$  es el “error” para la  $i$ -ésima observación; esto es la diferencia entre el valor observado de  $y$  con respecto a la media poblacional (véase Figura 2-1). Uno de los supuestos básicos de este modelo es que los residuos se distribuyen como una normal de media cero y varianza constante  $\varepsilon_i \sim N(0, \sigma^2)$ . La varianza resume la variabilidad alrededor de la media. Entre más grande sea este valor, la diferencia con respecto a la media se incrementa.

Figura 2-1: Residuos para tres puntos de un modelo de un solo nivel respecto a la media.



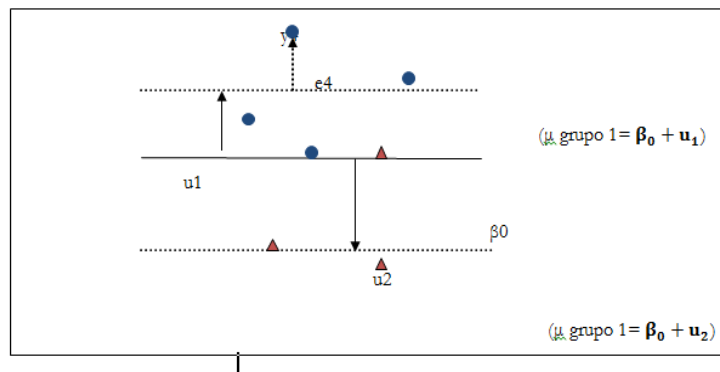
Ahora se introducirá el segundo nivel al modelo anterior. Supóngase que se tiene un conjunto de individuos en el nivel 1, anidados en grupos en el nivel 2. Para expresar algebraicamente esta relación, se añade el subíndice  $j$  a la respuesta de  $y_i$ , de esta manera  $y_{ij}$  representa el valor de  $y$  para el  $i$ -ésimo individuo en el  $j$ -ésimo grupo. El modelo multinivel permite estimar la

variabilidad entre los individuos de un mismo grupo y la variabilidad entre los grupos. Por lo tanto, el error se dividirá en dos componentes<sup>8</sup>, correspondiente a estas dos variaciones. Los errores entre los grupos se denotan como  $u_j$  y entre los individuos como  $\epsilon_{ij}$ . Integrando estos elementos al modelo (1), se origina la siguiente expresión:

$$y_{ij} = \beta_0 + u_j + \epsilon_{ij} \tag{2}$$

donde  $\beta_0$  ahora representa la media general de  $y$  para todos los grupos,  $u_j$  es la diferencia entre la media del grupo  $j$  y la media global (Véase Figura 2-2). En este caso, la media del grupo  $j$  es  $\beta_0 + u_j$ . Para los errores en el nivel 1,  $\epsilon_{ij}$  representa la diferencia entre los valores de  $y$  para el  $i$ -ésimo individuo con respecto a la media de su grupo,  $\epsilon_{ij} = y_{ij} - (\beta_0 + u_j)$ . Tal como en los modelos de regresión, se asume que ambos errores se distribuyen como una normal con media cero y varianza constante; es decir,  $u_j \sim N(0, \sigma_u^2)$  y  $\epsilon_{ij} \sim N(0, \sigma_e^2)$ .

Figura 2-2: Errores a nivel individual y grupal en un modelo de dos niveles.



El modelo también puede ser expresado de la siguiente forma:

Nivel 1  $y_{ij} = \beta_{0j} + \epsilon_{ij}$

Nivel 2  $\beta_{0j} = \beta_0 + u_j$

Combinado  $y_{ij} = \beta_0 + u_j + \epsilon_{ij}$

$$u_j \sim N(0, \sigma_u^2)$$

<sup>8</sup> Por esta razón, los modelos multinivel también son conocidos ampliamente como Modelos de componentes de la varianza.

$$\varepsilon_{ij} \sim N(0, \sigma_{\varepsilon}^2)$$

Los parámetros a estimar son  $\beta_0, \sigma_u^2, \sigma_{\varepsilon}^2$ .

### 2.3. Modelo multinivel de intercepto aleatorio

Hasta el momento, se ha estructurado el modelo sólo con el intercepto para ver el comportamiento de la variable respuesta debido sólo a la variabilidad entre los grupos o niveles y dentro de cada grupo. Supóngase que se tiene una variable continua explicativa en el nivel 1 denotada por  $X_{ij}$ . El subíndice  $ij$  en  $X$ , indica que los valores de  $X$  cambian de observación a observación dentro de un grupo. El modelo queda especificado de la siguiente forma:

$$y_{ij} = \underbrace{\beta_0 + \beta_1 X_{ij}}_{\text{parte fija}} + \underbrace{u_j + \varepsilon_{ij}}_{\text{parte aleatoria}} \quad (3)$$

En la expresión (3) la relación global entre  $x$  e  $y$  está representada por una línea recta en la que  $\beta_0$  muestra el intercepto o la altura de esta línea para el valor esperado de la variable respuesta dada una variable explicativa, y la pendiente o  $\beta_1$  constituye el cambio de la media de la variable respuesta para un cambio unitario de la variable explicativa. Se debe tener presente lo que se observa en la Figura 2-3, el intercepto para un grupo dado  $j$  está definido por la relación  $\beta_0 + u_j$ .

De esta forma, se tiene que como un tipo de modelo estadístico, el modelo multinivel está compuesto por dos partes: una fija y otra aleatoria como se observa en la expresión (3). La parte fija muestra la relación entre la media de  $Y$  y la variable explicativa y el componente aleatorio contiene los residuos del nivel 1 y del nivel 2. Usualmente, este modelo se conoce como modelo de intercepto aleatorio, porque el intercepto de la línea de regresión puede variar entre los grupos, pero la pendiente se asume fija para cada grupo. Gráficamente esto significa que se tendrán líneas de regresión para cada grupo paralelas entre sí, tal como se observa en la Figura 2-3. Por esta razón, también se puede especificar de la siguiente manera:

$$\text{Nivel 1} \quad y_{ij} = \beta_{0j} + \beta_1 x_{ij} + \varepsilon_{ij}$$

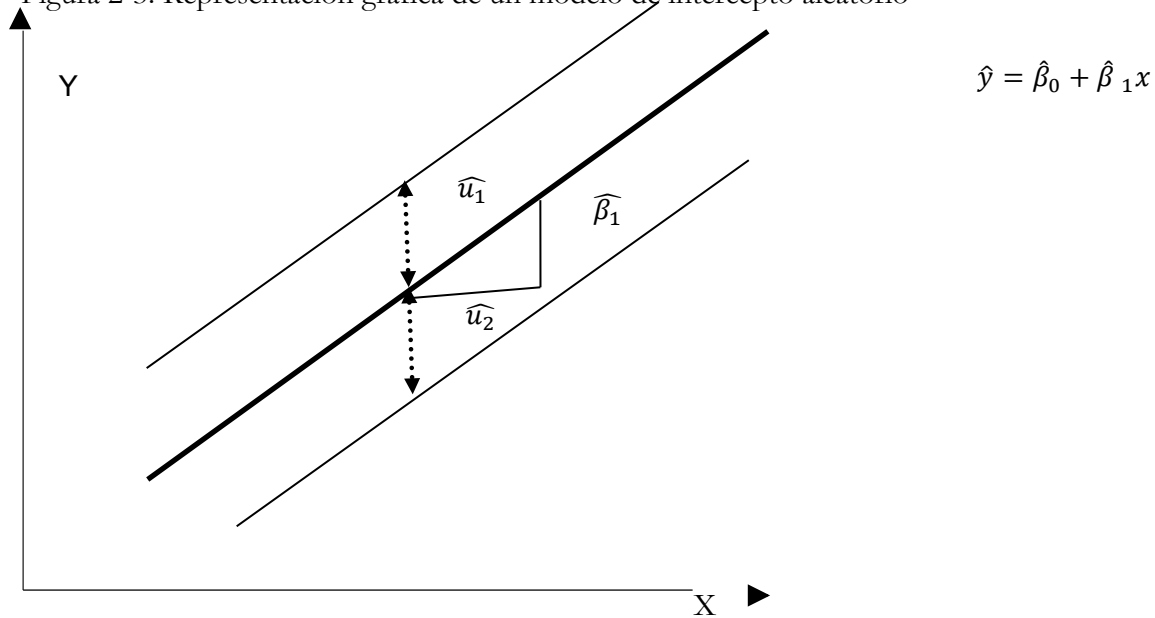
$$\text{Nivel 2} \quad \beta_{0j} = \beta_0 + u_j$$

Combinado  $y_{ij} = \beta_0 + \beta_1 x_{ij} + u_j + \varepsilon_{ij}$  (4)

$$u_j \sim N(0, \sigma_u^2)$$

$$\varepsilon_{ij} \sim N(0, \sigma_e^2)$$

Figura 2-3: Representación gráfica de un modelo de intercepto aleatorio



2.4 Modelo multinivel de pendiente e intercepto aleatorios.

En el modelo anterior, la pendiente  $\beta_1$  se mantenía fija para todos los grupos, pero supóngase que ésta varía aleatoriamente entre los grupos, lo que nos conduce a un modelo de pendiente aleatoria:

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + u_{0j} + u_{1j} x_{ij} + \varepsilon_{ij} \quad (5)$$

que también puede ser escrito como:

$$y_{ij} = \beta_{0j} + \beta_{1j} x_{ij} + \varepsilon_{ij}$$

$$\beta_{0j} = \beta_0 + u_{0j}$$

$$\beta_{1j} = \beta_1 + u_{1j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim N(0, \Omega_u): \Omega_u = \begin{bmatrix} \sigma_{u0}^2 & \\ \sigma_{u01}^2 & \sigma_{u1}^2 \end{bmatrix}$$

$$\varepsilon_{ij} \sim N(0, \sigma_e^2)$$

En la expresión (5), se han agregado nuevos términos a la ecuación, dentro del componente aleatorio. Ahora se tiene  $u_{1j}x_{ij}$ , y se añadió el subíndice 0 al término  $u_j$ . Asimismo, los supuestos se han modificado, pues ahora se asume que los errores  $u_{0j}$  y  $u_{1j}$ , se distribuyen como una normal bivariada con media cero y varianzas  $\sigma_{u_0}^2$ ,  $\sigma_{u_1}^2$ , y covarianza  $\sigma_{u_0u_1}$ , que es la covarianza entre los interceptos de grupo y las pendientes. Ahora la pendiente de la línea de regresión global es  $\beta_1$  y la pendiente para cada grupo  $j$  es  $\beta_1 + u_{1j}$ , por lo que la interpretación de los coeficientes cambia (Ver figura 2-4).

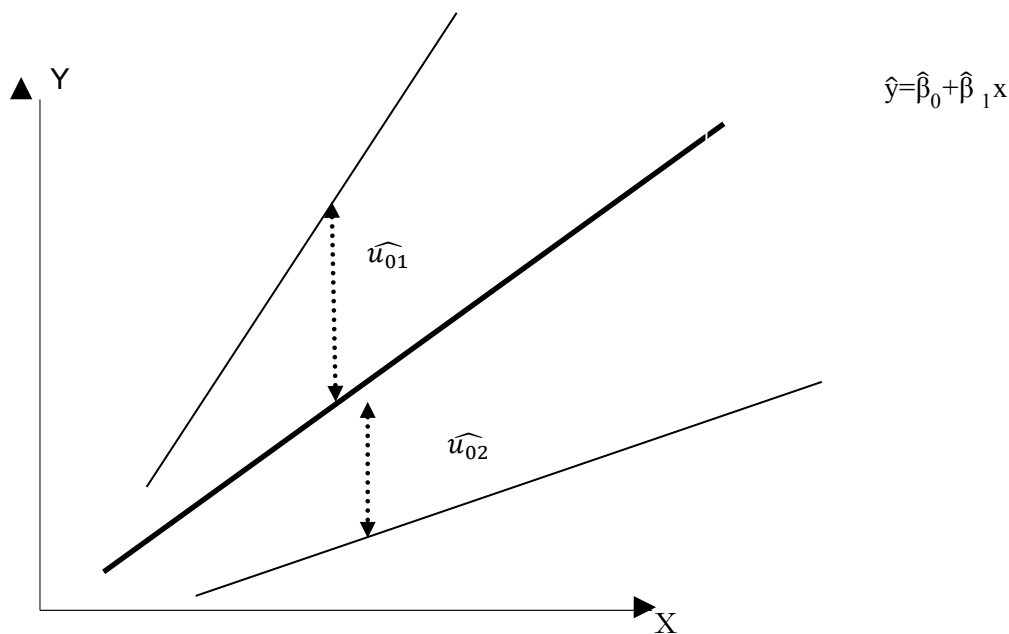


Figura 2-4: Representación gráfica de un modelo con pendiente aleatoria de dos niveles.

## 2.5 Principios de estimación en los modelos multinivel.

Una vez que se ha mostrado el modelo multinivel y sus variantes, se identifica que son tres los parámetros que deben estimarse en este modelo: Los efectos fijos, los coeficientes aleatorios del primer nivel y los componentes de la varianza y la covarianza. De acuerdo con Bryk y Raudenbush (1992), los métodos que se utilizan para la estimación son el de Máxima Verosimilitud Completo (Goldstein et al., 2002), Máxima Verosimilitud Restringida y la estimación Bayesiana. Los algoritmos utilizados para ello son: Algoritmo EM, Algoritmo de



Fisher – Scoring, propuesto por Longford (1987) y el de Mínimos Cuadrados Generalizados Iterativos (MCGI). A continuación se describe brevemente cada uno.

### 2.5.1 Método de Máxima Verosimilitud Completa.

El método de Máxima Verosimilitud completa o Full Maximum Likelihood (MLF) por sus siglas en inglés, busca principalmente estimar los valores de  $\beta$ ,  $T$  y  $\sigma_2$ , para los cuales la probabilidad de observar el valor verdadero de  $Y$  es máximo. Los estimadores que se obtienen a través de este enfoque cumplen con ciertas propiedades deseables. Si la variable aleatoria  $Y$  tiene función de probabilidad puntual  $f(y, \theta)$ , la verosimilitud  $L(\theta, y)$ , es simplemente  $f(y, \theta)$  mirada como función de  $\theta$  con  $y$  fijo. La función de probabilidad será definida en base al soporte  $y \in Y$ , mientras que la verosimilitud es definida sobre un espacio paramétrico  $\theta$ . Si se tiene una muestra aleatoria  $Y_1, \dots, Y_n$  con distribución  $f(y, \theta)$ , de manera que la verosimilitud será:

$$L(\theta) = \prod_{i=1}^n f(y_i, \theta)$$

Las propiedades de los estimadores a través de este método es que son consistentes, es decir que la probabilidad de obtener un valor muy cercano al verdadero parámetro será muy alta si se cuenta con suficientes datos, y eficientes, lo que significa que con una muestra grande de datos, los estimadores de máxima verosimilitud son insesgados y si la muestra se incrementa, su distribución muestral se aproxima a la normal con una varianza que puede ser fácilmente estimada (propiedad asintótica).

### 2.5.2 Máxima Verosimilitud Restringida.

La principal diferencia entre el método de estimación de Máxima Verosimilitud Restringida (MLR) y Máxima Verosimilitud Completa (MLF) estriba primordialmente en el número de unidades que se tengan en el nivel 2. Es decir, cuando las unidades en  $J$  sean grandes, ambos métodos originarán resultados similares en las estimaciones (Bryk y Raudenbush, 1992). No

obstante, si el número de unidades en el segundo nivel es pequeño, la varianza estimada por el método completo (MLF), será más pequeña que el método restringido (MLR) por un factor aproximado  $(J - F)/J$ , donde  $F$  es el número total de elementos en el vector de efectos fijos  $\mathbf{Y}$ .

### 2.5.3. Estimación Bayesiana.

Los dos métodos de estimación presentados anteriormente realizan inferencias sobre los efectos fijos  $\mathbf{Y}$ , en función de la precisión de los valores estimados de los parámetros de varianza y covarianza. Sin embargo, los estimadores bayesianos permiten al investigador hacer inferencias sobre  $\mathbf{Y}$  basado solo en la distribución posterior teniendo solamente los datos

(Bryk y Raudenbush, 1992). “La distribución posterior se calcula al incorporar  $\sigma^2$  y  $T$  de la distribución posterior conjunta de  $\mathbf{Y}$ ,  $\sigma^2$  y  $T$ . Para cualquier conjunto de valores específicos de ( $\sigma^2$  y  $T$ ), existe un estimador de  $\mathbf{Y}$  y sus errores estándar. El método Bayesiano toma esencialmente el peso promedio de todos los estimadores posibles de este tipo, donde el peso acordado para cualquier estimación es proporcional a la probabilidad posterior de ( $\sigma^2$  y  $T$ ).

La ventaja comparativa del método bayesiano está descrita en Lindley y Smith (1972), sin embargo la conceptualización bayesiana ha sido implementada recientemente solo para modelos lineales jerárquicos simples debido a que la integración numérica requerida para modelos más complejos ha resultado difícil de programar.

## 2.6 Prueba de hipótesis en los modelos multinivel.

Uno de los objetivos de los métodos de regresión es conocer si existe una relación lineal entre las variables bajo estudio a través del análisis de los datos recabados para poder estimar la media del valor de los parámetros que se desconocen de la población. En otras palabras, se busca encontrar que la relación entre las variables en cuestión sea estadísticamente significativa, de lo contrario no existiría ninguna relación causal entre las variables. Para ello, se recurre a elaborar una prueba de hipótesis. La prueba de hipótesis comienza con una suposición, denominada hipótesis, que se hace en torno a un parámetro del modelo estadístico que se asume describe a la población. Posteriormente, se reúnen los datos muestrales, se calculan las estadísticas de la muestra y en base a estos valores, con cierto grado de probabilidad, se dice si

la estimación del parámetro supuesto de la población es razonablemente aproximada a su verdadero valor. La diferencia entre el parámetro supuesto de la población y el estadístico muestral no suele ser ni tan grande que automáticamente se rechace la hipótesis, ni tan pequeño que de inmediato se acepte. Formalmente en una prueba de hipótesis se tienen los siguientes cuatro elementos:

Hipótesis Nula ( $H_0$ ): Esta es la hipótesis a probar, generalmente es una aseveración en el sentido de que un parámetro tiene un valor específico.

Hipótesis alternativa ( $H_1$ ): Esta hipótesis, sobre la cual se enfoca la atención, es una aseveración sobre el mismo parámetro poblacional que se utiliza en la hipótesis nula. Generalmente, se especifica que el parámetro poblacional tiene un valor diferente, de alguna manera al establecido en la hipótesis nula.

Estadístico de prueba: Es una función de las mediciones muestrales (variable aleatoria) en el cual se fundamenta la decisión “no se rechaza”  $H_0$ , o bien “se rechaza”  $H_0$ .

Región de rechazo: Especifica los valores del estadístico de prueba para los cuales se rechaza la hipótesis nula.

Además de los elementos anteriores, para llevar a cabo una prueba de hipótesis se deben definir el error estándar (ES) y los intervalos de confianza. El primero se refiere a la desviación estándar de la distribución muestral del estadístico y constituye una medida de imprecisión, en el que si se obtienen grandes valores indica una mayor incertidumbre sobre el verdadero valor poblacional. Los intervalos de confianza, son un rango de valores que contendrán el valor verdadero del parámetro poblacional, con una probabilidad asignada. Así, se tiene que el intervalo será el parámetro estimado  $\widehat{\beta}_1 \pm$  una medida de error.

En el caso de la modelación multinivel también se llevan a cabo las pruebas de hipótesis, no obstante, como la estructura de los datos es de anidamiento, pueden ser de un solo parámetro o pruebas multiparamétricas.

Capítulo 3

Investigación escolar en España:

Métodos y Materiales

## Capítulo 3: Investigación escolar en España: Métodos y Materiales

### 3.1 La Evaluación General de Diagnóstico.

Los datos utilizados para efectuar el presente análisis, fueron tomados del Instituto Nacional de Evaluación Educativa (INEE), que coordina el Ministerio de Educación Cultura y Deporte de España, que se encuentran disponibles en su página oficial de Internet. La base de datos corresponde a los resultados de la Evaluación General de Diagnóstico (EGD) del año 2009 y contiene información sobre el puntaje obtenido por el alumno en el examen nacional en la adquisición de las competencias básicas en el cuarto grado de Educación Primaria. Adicionalmente, el Instituto aplica una encuesta a una muestra de alumnos evaluados, los padres de familia, el profesor del grupo y los directores de los centros escolares. De acuerdo al Informe, la selección de los alumnos y centros participantes fue aleatoria y para garantizar la representatividad de la muestra, se aplicó un diseño de muestreo estratificado bietápico y por conglomerados, donde cada Comunidad Autónoma constituyó un estrato, incluyendo Ceuta y Melilla, y en donde los conglomerados fueron los centros escolares. También se consideró la titularidad<sup>9</sup> de los centros en una segunda etapa (INEE, 2010, p.16).

Se fijó un tamaño de muestra de 50 centros por comunidad autónoma y de 20 alumnos en promedio por grupo. Del tamaño de muestra proyectado, se tuvo una participación del 89% de los profesores, 92% de las familias y 99% de los directores de los centros escolares. La distribución final de los datos se muestra en la siguiente Tabla 3-1:

Tabla 3-1: Distribución de la muestra

Muestra	Datos
Alumnado	28.708
Padres de familia	25.741
Profesores	884
Directores	898
Centros escolares	900

Fuente: Informe de resultados EGD 2009.

<sup>9</sup>Centro público, privado y privado no concertado

Para responder a las pruebas los alumnos disponían de 200 minutos (50 minutos por competencia) más 45 minutos para cumplimentar el cuestionario, en dos días y en dos sesiones cada día.

### 3.2 La Competencia en Lingüística y Matemáticas.

En la prueba aplicada en el año 2009, fueron evaluadas cuatro de las ocho competencias básicas del currículo de cuarto grado de educación básica (fijado por el RD 1513/2006 de 7 de diciembre): La competencia básica en comunicación lingüística, matemáticas, en el conocimiento y la interacción con el mundo físico y la competencia básica social y ciudadana. Para fines de este estudio, sólo se utilizará la información correspondiente a las notas en competencia en lengua y en matemáticas, por ser también los principales aspectos que evalúan pruebas internacionales como PISA y TIMMS. Se elaboró una escala de rendimiento para cada una de las competencias evaluadas calculada por el INEE, basada en la dificultad de las preguntas. El índice de fiabilidad de las pruebas en las cuatro competencias evaluadas: competencia en comunicación lingüística (0,861) y competencia en matemáticas(0,800).

La OCDE define la competencia lectora o lingüística como “la capacidad de comprender, utilizar y analizar textos escritos para alcanzar los objetivos del lector, desarrollar sus conocimientos y posibilidades de participar en la sociedad; y la competencia en matemática como “una capacidad del individuo para identificar y entender las función que desempeñan las matemáticas en el mundo, emitir juicios fundados y utilizar relacionarse con las matemáticas de forma que se puedan satisfacer las necesidades de la vida de los individuos como ciudadanos constructivos, comprometidos y reflexivos”. (OCDE, 2001).

En el informe de resultados de la Evaluación General de Diagnóstico publicado por el INEE (2010), se detallan los ítems seleccionados y las dimensiones evaluadas en cada competencia, por si se desea profundizar en el tema. El propósito de este trabajo es realizar investigación cuantitativa utilizando los resultados alcanzados por los estudiantes.

### 3.3. Factores analizados.

Además de la aplicación de las pruebas estandarizadas correspondientes para medir el nivel de logro en cada una de las competencias, se aplicó un cuestionario a los alumnos, padres de familia, profesores y directores de los centros escolares participantes, para conocer su contexto y su percepción sobre el proceso de enseñanza-aprendizaje.

El número total de variables que contiene la base de datos es de 450 (sin incluir la nota en las competencias, ni las variables identificadoras de cada cuestionario), 148 corresponden al cuestionario alumnos, 45 al de las familias, 114 al de directores de centros escolares y 143 al profesorado de cuarto grado de primaria. En el Anexo A, se presenta a detalle qué aspectos se midieron en cada cuestionario.

Para la selección de variables que se incluyeron en el modelo, como factores que pueden estar asociados al rendimiento escolar, se tomaron en cuenta los siguientes criterios.

- a) **Confirmatorias:** Aquéllas variables que han sido consideradas en otras investigaciones mencionadas previamente y que se ha comprobado su efecto, basado en las investigaciones en España, hechas por Murillo (2008); Choi de Mendizábal y Calero-Martínez (2009); Cervini et al. (2014).
- b) **Generadas:** Variables construidas que condensan una gran cantidad de información contenida en un grupo de variables.

Este grupo de variables generadas será un aporte a la forma en que los factores son incluidos en los modelos lineales jerárquicos. Un gran número de investigaciones recurren principalmente a establecer interacciones entre las variables (Murillo, 2008) y otros a técnicas multivariadas como el Análisis de Componentes Principales o al Análisis Factorial. Estas técnicas indudablemente contribuyen a la generación de nuevos factores que contienen la información relevante de las variables originales y evitan problemas de carácter metodológico propios de los modelos. En este caso, se generaron variables para medir los factores de nivel 3 que se dan en el Aula escolar. Es decir, aquéllos relacionados con las características del profesorado y el clima en el aula.

En este sentido, se generaron 3 variables que condensan información sobre el estilo del profesor, respecto a su práctica docente, los materiales que utiliza y su forma de evaluar. Las variables que se eligieron fueron los 24 ítems correspondientes a los incisos de las preguntas 23, 24 y 25 del cuestionario aplicado (Ver tabla 3-1). Se le pidió a cada profesor indicar la frecuencia con la que realizan determinada actividad en una escala de Likert de 4 puntos: “Nunca o casi nunca”, “Algunas veces”, “Casi siempre” y “Siempre”.

El método utilizado para generar las tres nuevas variables, fue la técnica de agrupación en dos pasos para datos numéricos y categóricos (The two step method for clustering mixed numerical and categorical data, TMCM). Como señala Timm (2000), este método define la similitud entre los ítems de atributos categóricos, basado en la idea de la coocurrencia, es decir, todos los valores categóricos serán convertidos a valores numéricos en función de la similitud de hacer que los atributos contengan sólo un valor numérico, lo que se traduce en el supuesto de que si dos ítems siempre aparecen juntos en un solo objeto, existirá una fuerte similitud entre ellos. Se utilizó el programa estadístico IBM-SPSS versión 19 que incluye este algoritmo en su análisis de clasificación. El número de clúster se determina automáticamente verificando el criterio de información de Akaike (AIC, por sus siglas en inglés).  $AIC = -2_{lk} + 2_{rk}$ . Se utiliza este criterio, porque minimiza la medida asimétrica de diferencias entre una función de distribución de probabilidades verdadera y otra estimada. La medida utilizada es una función dependiente de la función de máxima verosimilitud y penalizando por el número de regresores estimados (Yi y Judge, 2008).



Tabla 3-2: Dimensiones de la práctica educativa del profesorado.

Perfil Práctica Educativa		
Práctica docente	Recursos didácticos	Formas de evaluar
<ol style="list-style-type: none"> <li>1. Tiempo que utiliza explicando durante la clase</li> <li>2. Promueve debates en clase</li> <li>3. Los alumnos exponen temas o trabajos</li> <li>4. Los alumnos trabajan individualmente</li> <li>5. Los alumnos trabajan en pequeños grupos.</li> <li>6. Los alumnos realizan las actividades que propone</li> <li>7. Proporciona actividades distintas a alumnos con necesidades específicas de apoyo educativo</li> </ol>	<ol style="list-style-type: none"> <li>1. El libro de texto para estudiar y hacer ejercicios</li> <li>2. Cuadernillos de ejercicios</li> <li>3. Libros de biblioteca del colegio o del aula</li> <li>4. Materiales elaborados por el propio profesor</li> <li>5. Prensa o revistas especializadas</li> <li>6. Ordenadores e Internet</li> <li>7. Medios audiovisuales</li> </ol>	<ol style="list-style-type: none"> <li>1. Realización de pruebas de conocimiento antes de cada tema</li> <li>2. Lleva controles después de cada tema</li> <li>3. Controles escritos trimestrales.</li> <li>4. Corrige los deberes y cuadernos.</li> <li>5. Revisa las actividades que los alumnos hacen en clase.</li> <li>6. Efectúa exámenes orales.</li> <li>7. Valora la autoevaluación del alumnado.</li> <li>8. Valora el interés y la participación en clase de alumnas y alumnos</li> </ol>

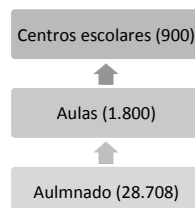
Fuente: Elaboración propia

### 3.4 Niveles considerados en el modelo.

El propósito de aplicar el modelo a los datos, es analizar qué información adicional se puede extraer de un modelo multivariado multinivel. Se utilizará la información contenida en la Evaluación General de Diagnóstico 2009 y se especificará un modelo de 3 niveles, dónde el primer nivel corresponde a la información del estudiante, el segundo nivel al salón de clases (Aula) y el tercer nivel, al centro escolar. No se utilizó la Comunidad Autónoma donde se encuentra ubicada la escuela, debido a que esta información no se proporciona en la base de datos publicada por el Instituto Nacional de Evaluación Educativa.

Los niveles en el modelo estarán definidos de la siguiente manera:

Figura 3-1: Diagrama de unidad de los niveles del modelo.



Fuente: Elaboración propia

En la siguiente tabla, se detalla cada una de las variables que se considerarán como factores en cada nivel.

Tabla 3-3: Factores explicativos introducidos en el modelo.

Nivel	Factores	Categorías
1- Alumnado (i)	Género	Chica Chico
	Edad inicia el colegio	<2 años 2-5 años* >6años
	Lugar de nacimiento	España* País de Europa América África Asia Otro país
	Lengua	Castellano * Catalán Gallego Valenciano Vasco
	Estudios madre	Sin estudios Incompletos Obligatorios Bachillerato y Formación P. Universitarios*
	Situación laboral madre	Laborando*C En paro Jubilada Labores del hogar
	Expectativas del alumno	Terminar ESO Ciclo de Formación P. Universitarios* No sabe
2. Aula (j) (Profesor y clima del aula)	Antigüedad del profesor	< 5 años 5-9 años 10-19 años 20-29 años >30 años
	Practica educativa profesor <sup>++</sup>	Innovador* (Promueve debates en

		clase) Tradicional (Explica la mayor parte del tiempo)
	Métodos didácticos que utiliza <sup>++</sup>	Innovador* (Utiliza medios audiovisuales) Tradicional (Se basa en el libro de texto)
	Forma de evaluar <sup>++</sup>	Innovador (Se apoya en la autoevaluación) Tradicional* (Realiza controles frecuentemente).
	Satisfacción	Poco Algo* Bastante Mucho
3. Centro escolar ( <i>k</i> )	Titularidad	Público* Privado concertado Privado no concertado
	Tamaño	< 200 estudiantes 201-500 estudiantes* >501 estudiantes
	%alumnos ext. misma lengua	Escala continua
	%alumnos ext. distinta lengua	Escala continua

\*Categorías de referencia, ++(variables generadas)

### 3.5 Modelos multinivel especificados.

Se especificaron 3 modelos multinivel, utilizando los mismos factores explicativos para cada uno de ellos. Uno para la competencia lingüística, otro para la competencia matemática y el modelo integrado de respuesta bivariada que incluye ambas competencias en el mismo modelo.

#### 3.5.1. Modelo univariado de 3 niveles para la variable respuesta competencia lingüística.

$$y_{1ijk} = \beta_{0jk} + e_{ijk}$$

$$\beta_{0jk} = \beta_0 + v_{0k} + u_{0jk}$$

$$v_{0k} \sim N(0, \sigma_{v_0}^2)$$

$$u_{0jk} \sim N(0, \sigma_{u_0}^2)$$

$$e_{ijk} \sim N(0, \sigma_e^2)$$

Donde  $y_1$  es el puntaje obtenido por el  $i$ -ésimo alumno en la competencia en lingüística de la  $j$ -ésima aula y del  $k$ -ésimo centro escolar;  $v_{0k}$ , las diferencias entre los centros escolares,  $u_{0jk}$ , los residuos a nivel aula y  $e_{ijk}$  las distancias entre los alumnos.

### 3.5.2. Modelo univariado de 3 niveles para la variable respuesta competencia en matemáticas.

$$y_{2ijk} = \beta_{0jk} + e_{ijk}$$

$$\beta_{0jk} = \beta_0 + v_{0k} + u_{0jk}$$

$$v_{0k} \sim N(0, \sigma_{v_0}^2)$$

$$u_{0jk} \sim N(0, \sigma_{u_0}^2)$$

$$e_{ijk} \sim N(0, \sigma_e^2)$$

Donde  $y_2$  es el puntaje obtenido por el  $i$ -ésimo alumno en la competencia en matemáticas de la  $j$ -ésima aula y del  $k$ -ésimo centro escolar;  $v_{0k}$ , las diferencias entre los centros escolares,  $u_{0jk}$ , los residuos a nivel aula y  $e_{ijk}$  las distancias entre los alumnos.

### 3.5.3 Modelo multinivel multivariado para las variables respuesta competencia en lingüística y competencia en matemáticas.

El modelo multinivel multivariado, consiste en especificar un conjunto de variables explicatorias ( $X$ ), para estimar su efecto ( $\beta$ ) en dos o más respuestas ( $y_1, y_2, \dots, y_n$ ) de las unidades ( $i$ ) que presentan anidamiento en dos o más niveles ( $j, k, l \dots p$ ). De esta manera, se integra un modelo en el cual además de estimar los parámetros, considerando la estructura jerárquica y observando la variabilidad intra e inter grupal, se tiene una matriz de varianzas y covarianzas ( $\Omega_u$ ) entre las variables dependientes en los mismos niveles de análisis. Esta matriz muestra la relación que existe entre dichas varianzas y la fuerza con la que estas dos variaciones están asociadas.

Sea un conjunto de unidades  $i = 1, 2, \dots, I_j$ , anidadas en un grupo que se denota por  $j = 1, \dots, J$ , para un modelo de dos niveles se tiene la siguiente estructura jerárquica multivariada (Goldstein et al., 2009):

$$y_{ij}^{(1)} = X_{1ij}\beta^{(1)} + Z_{1ij}u_j^{(1)} + e_{ij}^{(1)} \quad ( 1 )$$

$$y_j^{(2)} = X_{2j}\beta^{(2)} + Z_{2j}u_j^{(2)} \quad ( 2 )$$

$$e_{ij}^{(1)} \sim MVN(0, \Omega_1), u_j = (u_j^{(1)}, u_j^{(2)})^T \quad ( 3 )$$

$$u_j \sim MVN(0, \Omega_2) \quad ( 4 )$$

Los superíndices denotan el nivel en el cual la variable tiene una respuesta. Esta estructura representa el modelo general, donde se asume el mismo conjunto de predictores para cada respuesta de nivel 1 y de nivel 2 y los residuos de nivel 1, se calculan al restar el valor estimado del componente fijo del modelo a cada una de las respuestas de nivel 1.

$$y_{ij} = X_{1ij}\beta + Z_{ij}u_j + e_{ij}$$

$$e_{ij} \sim MVN(0, \Omega_1)$$

$$u_j \sim MVN(0, \Omega_u), \quad \Omega_u = \begin{bmatrix} \sigma_{u0}^2 & \\ \sigma_{u01}^2 & \sigma_{u1}^2 \end{bmatrix} \quad ( 5 )$$

La forma en la que opera el modelo multinivel de respuesta multivariada es como señala Rasbash et al. (2009) creando un nivel extra “por debajo” de las unidades del primer nivel. De esta manera, se tienen las respuestas de un individuo, que a su vez están agrupadas en un nivel superior. Esta especificación mide simultáneamente si una variable explicativa tiene un efecto mayor, menor o igual en cada una de las variables respuestas medidas para cada individuo.

Una de las ventajas de ajustar un modelo multivariado multinivel es que no se necesitan eliminar los casos, donde existan datos faltantes para una de las mediciones. Esto implica que se puede llevar a cabo el análisis, aunque haya ausencia de respuesta, siempre y cuando no sea de manera sistemática (Rasbash et al., 2009).

### 3.6 Modelo multivariado de cuatro niveles.

En este trabajo interesa identificar qué nuevo conocimiento aporta la especificación de un modelo bivariado multinivel para explicar los factores que se asocian al rendimiento en la competencia lingüística y la competencia en matemática de un alumno de cuarto grado de

Educación Básica, al modelarlo simultáneamente y no de manera separada como se ha realizado en varias investigaciones (Calero y Escardíbul, 2007; Murillo, 2008; Achilli, 2009; Cervini et al., 2014). El objetivo es analizar si los factores tienen un efecto similar o diferente en cada competencia y si existe una correlación entre los resultados de las mismas.

Por tanto, se considerarán cuatro niveles, en el primer nivel, se encuentran las 2 respuestas para cada individuo, en el segundo nivel los alumnos, en el tercer nivel las aulas o salones de clase, y en el cuarto nivel, los centros escolares; dos variables dependientes: el puntaje en competencia lingüística y el puntaje en competencia matemática y los factores descritos en la tabla 3-1. El modelo queda especificado de la siguiente manera:

$$y_{1jkl} \sim N(XB, \Omega)$$

$$y_{2jkl} \sim N(XB, \Omega)$$

$$y_{1jkl} = \beta_{0jkl} \text{cons}_{0ijkl}$$

$$\beta_{0jkl} = \beta_0 + f_{0l} + v_{0kl} + u_{0jkl}$$

$$y_{2ijkl} = \beta_{1jkl} \text{cons}_{1ijkl}$$

$$\beta_{1jkl} = \beta_1 + f_{1l} + v_{1kl} + u_{1jkl}$$

$$\begin{bmatrix} f_{0l} \\ f_{1l} \end{bmatrix} \sim N(0, \Omega_f): \Omega_f = \begin{bmatrix} \sigma_{f0}^2 & \\ \sigma_{f01}^2 & \sigma_{f1}^2 \end{bmatrix}$$

$$\begin{bmatrix} v_{0kl} \\ v_{1kl} \end{bmatrix} \sim N(0, \Omega_v): \Omega_v = \begin{bmatrix} \sigma_{v0}^2 & \\ \sigma_{v01}^2 & \sigma_{v1}^2 \end{bmatrix}$$

$$\begin{bmatrix} u_{0jkl} \\ u_{1jkl} \end{bmatrix} \sim N(0, \Omega_u): \Omega_u = \begin{bmatrix} \sigma_{u0}^2 & \\ \sigma_{u01}^2 & \sigma_{u1}^2 \end{bmatrix}$$

Donde  $y_{1jkl}$ , es el puntaje en competencia lingüística del alumno  $j$  y  $y_{2jkl}$ , es el puntaje en la competencia en matemáticas del alumno  $j$ .

No existe variación de primer nivel, porque el nivel 1 en estos modelos sirve para definir la estructura multivariada (las dos respuestas).  $\Omega_u$  representa la matriz de varianzas y covarianzas de nivel 2, es decir las diferencias (residuos) entre los estudiantes para cada respuesta

---

( $\sigma_{u0}^2$  y  $\sigma_{u1}^2$ ) y su correlación  $\sigma_{u01}^2$ , mientras que  $\Omega_v$  corresponde a la matriz de nivel 3, para el aula y  $\Omega_f$ , nivel 4 para los centros escolares.

Capítulo 4

Resultados y Aportaciones



## Capítulo 4: Resultados y Aportaciones

### 4.1 Análisis descriptivo de los datos.

Una aproximación previa de la información que arrojó la aplicación de la Evaluación General de Diagnóstico y las encuestas aplicadas a padres de familia, profesorado y directores de centros escolares, mostraron los siguientes resultados. Este análisis exploratorio permite tener un acercamiento para conocer las principales características de los niveles incluidos en los modelos.

La edad promedio de los alumnos evaluados es de 9 años, 6 de cada 10 asisten a colegios públicos, 3 a privados y 1 a privado no concertado y el 48,8% son niñas. El 6,4% nació en Andalucía, el 6,2% en Madrid y el 5,9% en Sudamérica, mientras que el 2% proviene de otros países Europeos y un 0,7% de Asia, Australia y África. Respecto a la lengua que utilizan cuando están en casa, en el colegio o con los amigos, el 77% respondió que el castellano, el 11% el catalán, el 5,4% el gallego, 4,8% vasco y un 1,4% valenciano.

Es importante mencionar, que muchas de las preguntas que se incluyen en el cuestionario de alumnos, de padres de familia y de profesores, son las mismas, esto se hizo como una manera de corroborar la información y de contraponer la percepción de cada uno respecto al mismo tema. Así, el 33,8%<sup>10</sup> de los alumnos respondió que inició el colegio a la edad de 3 años, sólo un 2,7% haberlo iniciado a los 6 o más años y un 6% declaró haber reprobado 2º o 4º.

En cuanto a su método y actividades de estudio, el 44,8% manifiesta dedicar a sus tareas escolares una hora o menos y sólo un 6,1% indicó que generalmente no tenía deberes. El 65,3% prefirió no responder a la pregunta si recibe ayuda de un profesor o de sus padres, después de clase. En esta línea, el 68,5% respondió que todos o casi todos los días, juega con los amigos y un 57,4% hace deporte, contra un 44% que nunca o casi nunca lee periódicos o revistas.

---

<sup>10</sup> Esta información no se pudo comparar con la respuesta del cuestionario de familias, porque la pregunta en este último tenía una proporción muy alta de no respuesta.

Sobre sus recursos para el estudio, el 30% menciona que en su casa hay de 26 a 100 libros, 42,9% tienen más de 3 televisores y un 46,5% tiene un ordenador, 14,3% no tiene acceso a Internet. Los que sí tienen, el 39,3% lo usa con mucha frecuencia, el 41% busca información para sus tareas una o dos veces a la semana y el 47% menciona nunca o casi nunca usarlo para comunicarse con otras personas.

En relación a su contexto familiar, el 68,7% expresó tener siempre una buena relación con su madre y el 67,7% con su padre, por el contrario un 1,9% indicó nunca o casi nunca tenerla con la madre y un 2,9% con el padre.

En cuanto al profesorado, llama principalmente la atención que el 72,2% son mujeres y la edad en la que se encuentran, pues el 44,9% tiene entre 50-59 años, y el 34,4% declaró tener más de 30 años de experiencia. Se trata pues de una edad promedio más alta con respecto a la media de los países de la OCDE. El 79,3% de los profesores, cuenta con titulación de diplomado, técnico o equivalente, y 16% de licenciado. En este nivel educativo, el nivel de profesionistas con doctorado es mínimo 0,3%. Cabe destacar, que más de la mitad (52%) tiene plaza de funcionario definitiva y el 93,3% se encuentra satisfecho con sus labores académicas.

Respecto a las variables generadas, se aplicó un análisis clúster en dos fases y se identificó la existencia de tres perfiles para cada dimensión: El perfil tradicionalista, el perfil innovador o progresista y un perfil intermedio o mixto (Ver tabla 3-2). El número de profesores agrupados no fue el mismo en cada dimensión, debido a los valores perdidos que fueron excluidos por ausencia de respuestas, sin embargo el porcentaje fue mínimo y no afectó la validez de los resultados. Así, para la dimensión de práctica docente, el número de profesores clasificados fue de 853 (94,8% del total), de los cuales, el 48% fue agrupado en un estilo Tradicionalista, respecto al uso de recursos didácticos de los 857 (95,22%) profesores agrupados, el 45% se identificó más con el estilo Innovador y, finalmente en cuanto a la forma de evaluar, de los 835 (92,8%) profesores en este grupo, el 49% se inclinó por un estilo Tradicional.

En un análisis inicial de las competencias evaluadas, se observa que el promedio de puntaje que obtuvieron los alumnos en competencia lingüística es casi igual a la competencia en matemáticas, 502,65 y 502,28, respectivamente, con varianzas similares y una desviación típica

de 100 puntos. El puntaje máximo que obtuvo un alumno en competencia lingüística fue de 853,21 y en matemáticas de 851,69 (Véase tabla 4-1).

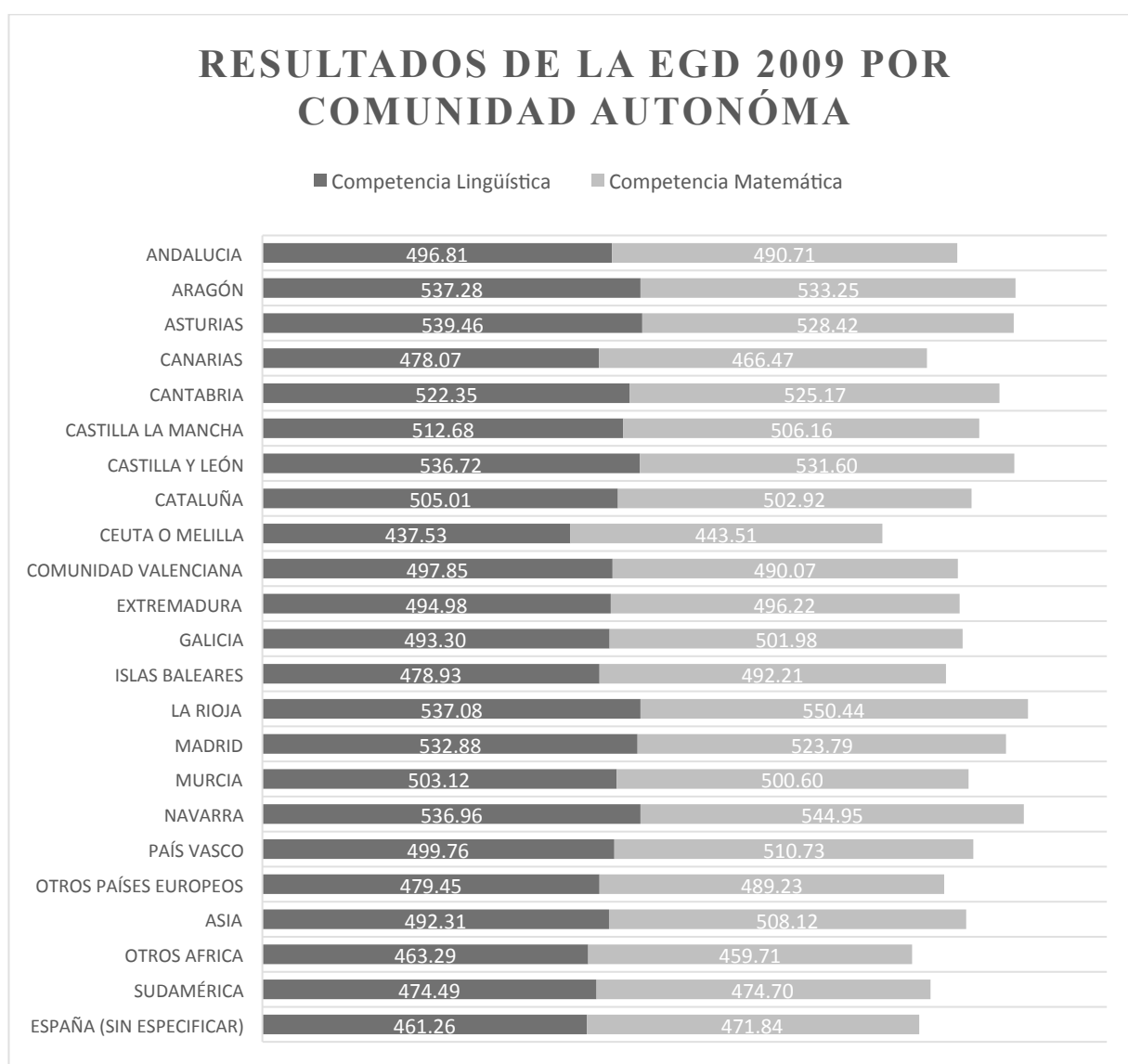
Tabla 4-1: Resultados por alumno en Competencia Lingüística y en Competencia en Matemáticas

Competencia	Media	Máximo	Mínimo	Varianza	Desviación Típica
	Aritmética				
Lingüística	502,65	853,21	136,80	9814,93	99,07
Matemática	502,28	851,69	136,95	9925,49	99,62

Fuente: Elaboración propia

Al comparar los resultados promedio por la comunidad autónoma en la que nació el alumno, se observa que sí hay variación entre las mismas, siendo de La Rioja, Navarra, Asturias y Aragón los alumnos que obtuvieron puntajes más altos, y de Canarias, Ceuta y Melilla y los provenientes de países africanos, los que lograron puntajes más bajos.

Tabla 4-2: Resultados de la EGD 2009 por comunidad autónoma.



Fuente: Elaboración propia

En la figura 4-1 y 4-2, se observa que los resultados obtenidos por los alumnos en cada una de las competencias, varía por la titularidad del centro escolar y por el género del alumno. En los centros privados es más alto el puntaje que obtienen los chicos en matemáticas que en lingüística. En los centros públicos, se observa que los chicos obtienen puntajes más bajos en lingüística que las chicas.

Figura 4-1: Resultado de la competencia Lingüística por titularidad del centro

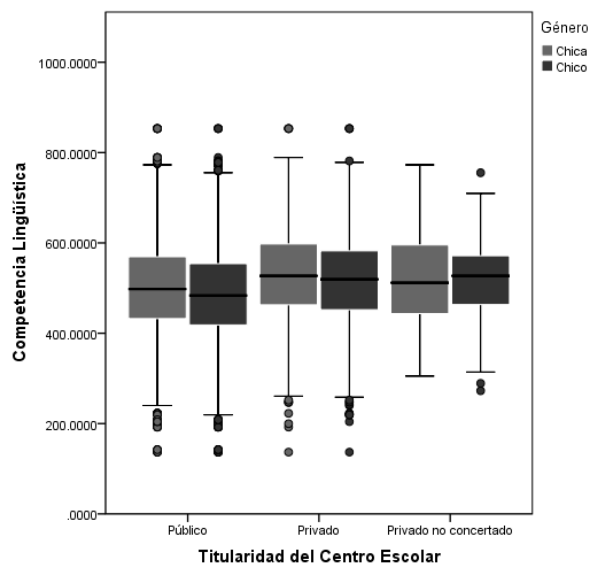
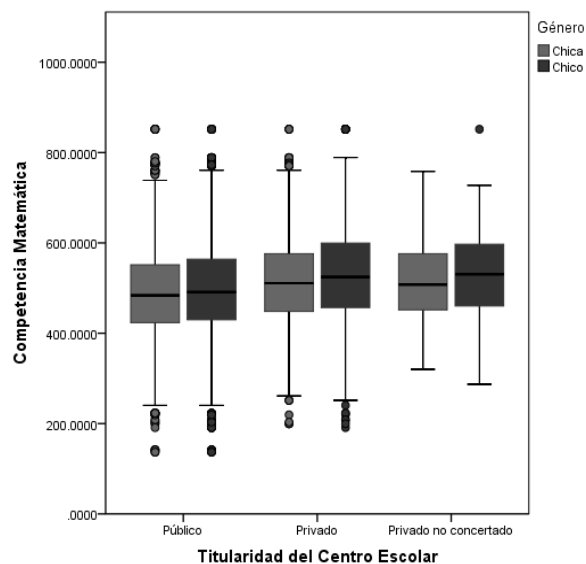


Figura 4-2: Resultados de la competencia en Matemática por titularidad del centro



#### 4.2 Comparación de modelos nulos.

El modelo nulo se establece como punto de partida para iniciar el proceso de modelado, no tiene variables exploratorias, pero da información de la variabilidad del fenómeno y la proporción que es explicada por los diferentes niveles de análisis. Únicamente se estima el intercepto  $\beta_0$ , y los componentes de la varianza.

Tabla 4-3: Resultados del modelo multinivel nulo para rendimiento en Lingüística

	Estimador	SE*
<i>Parte fija</i>		
$\beta_{0jk}$	499,57	(1,47)
<i>Parte aleatoria</i>		
$\sigma_{v0}^2$	1459,7	(95,1)
$\sigma_{u0}^2$	218,9	(32,4)
$\sigma_e^2$	8157,04	(70,9)

Tabla 4-4: Resultados del modelo multinivel nulo para rendimiento en Matemáticas.

	Estimador	SE*
<i>Parte fija</i>		
$\beta_{0jk}$	499,44	(1,52)
<i>Parte aleatoria</i>		
$\sigma_{v0}^2$	1313,76	(90,0)
$\sigma_{u0}^2$	250,8	(41,8)
$\sigma_e^2$	8347,04	(72,4)

Fuente: Elaboración propia

\*SE: Estándar Error (Error Estándar)

Al ejecutar los modelos univariados de forma separada, se aprecia que en el intercepto fue muy parecido en ambos, es decir, el promedio para todos los alumnos de los diferentes centros escolares tanto para la competencia en lingüística como para la competencia en matemáticas fue de 499,57, con un error típico de 1,47 y 1,52. No obstante, se observan distintas variabilidades en cada uno de los niveles. Se registra una mayor varianza entre el alumnado  $\sigma_e^2$  y entre las aulas  $\sigma_{u0}^2$  en la competencia en matemática que en la competencia lingüística, y mayor variación a nivel escuelas  $\sigma_{v0}^2$  (Ver tabla 4-2 y 4-3).

En el modelo nulo para rendimiento en lingüística (tabla 4-3), el porcentaje de varianza explicado por las diferencias entre los alumnos es de 83%, entre las aulas de 2% y entre las escuelas de 15%. Mientras que en el modelo de competencia en matemáticas la variabilidad es de 84%, 2% y 13% respectivamente. Este porcentaje se obtiene a través del Coeficiente de Correlación Intraclase, que resulta de dividir cada una de las varianzas entre la total ( $\sigma_{v0}^2 + \sigma_{u0}^2 + \sigma_e^2$ ).

El modelo multivariado multinivel nulo, en el que se incluyen rendimiento en lingüística y rendimiento en matemáticas simultáneamente, arroja resultados en los estimadores similares a los que se obtuvieron en los modelos univariados, tanto de la parte fija como aleatoria (ver tabla 4-3). Adicionalmente, se tiene información sobre la covarianza que existe en los residuos entre ambos rendimientos para los diferentes niveles ( $\Omega_u, \Omega_v, \Omega_f$ ). Es importante recordar, que

en el modelo multivariado, se tendrá un nivel adicional, pues se toma como primer nivel las observaciones de cada individuo para cada una de las respuestas, así  $\sigma_{u0}^2$ , corresponde a la variabilidad entre el alumnado (nivel 2),  $\sigma_{v0}^2$ , entre las aulas (nivel 3) y  $\sigma_{f0}^2$  entre los centros escolares (nivel 4).

Los resultados de las covarianzas de las respuestas en los distintos niveles  $\sigma_{u01}^2, \sigma_{v01}^2, \sigma_{f01}^2$  muestran que existe una dependencia directa positiva entre los resultados de los alumnos en la competencia lingüística y matemáticas. Esto indica que alumnos con altos puntajes en lingüística (residuo positivo  $u_0$ ), tenderán a tener un puntaje alto en matemáticas. También se observa que la covarianza es mayor en los centros escolares que en comparación con las aulas, lo que indica que hay una mucho mayor variación entre los puntajes que obtiene un alumno en ambas competencias entre centros escolares, que entre diferentes aulas de un mismo centro escolar.

Tabla 4-5: Resultados del modelo multinivel multivariado nulo para rendimiento en Lingüística y Matemáticas.

	Competencia Lingüística		Competencia Matemática	
	Estimador	SE	Estimador	SE
<i>Parte fija</i>				
$\beta_{0jkl}$	499,04	(1,46)	$\beta_1$	499,08 1.42
<i>Parte aleatoria</i>				
$\sigma_{f0}^2$	1499,2	(93,9)	$\sigma_f^2$	1334,5 (90,2)
$\sigma_{f01}^2$	1311,0	(84,5)		
$\sigma_{v0}^2$	216,0	(38,8)	$\sigma_v^2$	250,28
$\sigma_{v01}^2$	151,29	(32,1)		
$\sigma_{u0}^2$	8161,75	(70,9)	$\sigma_u^2$	8351,5 (72,4)
$\sigma_{u01}^2$	3816,82	(56,1)		
-2*Log likelihood 656213.96(55657 casos)*				

Fuente: Elaboración propia

\*El número de casos se duplica porque hay dos respuestas por individuo.

Desde el modelo nulo, se puede apreciar, la primera contribución de especificar un modelo integrado, para tener una nueva aproximación del rendimiento de los alumnos. El incluir ambas respuestas simultáneamente, además de reducir la complejidad de tener dos modelos, brinda más información respecto a la variación entre los diferentes niveles y las covarianzas para cada  $y$  incluida, permitiendo observar dónde se encuentra la mayor fuente de variación.

#### 4.2.1. Factores en el primer nivel.

La primer variable que se introdujo al modelo multivariado multinivel fue el género del estudiante, que ha resultado tener un efecto significativo en la mayoría de las investigaciones en España (Calero y Escaldibur, 2007) y en otros ha resultado nulo (Murillo, 2008). En estos modelos, se tomó como variable de referencia a las chicas y se generó una variable dummy para chico. Los resultados indican que el género del alumno sí resulta significativo para explicar las diferencias en los resultados.

Tabla 4-6: Componentes de la varianza modelo con factores en el primer nivel.

		Lingüística		Matemáticas		Deviance
		Estimador	SE	Estimador	SE	-
						2*Loglikelihood
Modelo univariado rendimiento	$\beta_{0jk}$	506,24	(1,57)			327619,02
	$\beta_1Chico_{ijk}$	-12,94	(1,10)			(27520)
Lingüística						
Modelo univariado rendimiento	$\beta_{0jk}$			494,38	(1,53)	323530,702
	$\beta_1Chico_{ijk}$			+11,61	(1,13)	(27128)
Matemáticas						
Modelo multivariado LyM	$\beta_{0jkl}$	506,07	(1,56)	494.05	(1,54)	6436601,10
	$\beta_1Chico_{ijkl}$	-13,04	(1,10)	+11.42	(1,12)	(54848)



Se confirma como el valor de los estimadores es el mismo para los tres modelos, así como el de los errores estándar. El efecto de la variable muestra que los chicos, tienden a tener 12,94 puntos menos en promedio en su rendimiento en lingüística con respecto al de las chicas (506,24), mientras que en matemáticas tienen más puntos en promedio (11,42), con respecto al de las chicas (494,05). En los modelos univariados, el porcentaje de la varianza explicado a nivel 1, se redujo a 80% para ambas competencias, sin embargo, en el modelo multivariado, se tiene una mayor información con la matriz de varianzas y covarianzas y las correlaciones entre las respuestas.

La matriz de covarianzas confirma la relación directa positiva que existe entre las respuestas, es decir a mayores puntajes de un alumno en lingüística, también tendrá mayor puntuación en matemáticas. Por su parte, la correlación entre los residuos a nivel alumno, aula y centro escolar, va aumentando (Tabla 4-6). A nivel alumno, considerando las diferencias de género es de 0,419 a nivel 2 (aulas o salones de clases) de 0,645 y a nivel centro escolar es 0,939. Esto significa que hay menos diferencias entre las chicas de una misma escuelas, que entre las chicas de distintas escuelas. Esta información, sólo se puede deducir de un modelo integrado de dos o más variables respuestas y no de un modelo univariado.

Tabla 4-7: Correlaciones entre los residuos de los diferentes niveles del modelo multivariado.

		Lingüística		Matemáticas	
<b>Nivel Centro Escolar (l)</b>	Lingüística	$\sigma_{f0}^2$	1449.293 Corr: 1.000		
	Matemáticas	$\sigma_{f01}$	1306.657 Corr: <b>0.939</b>	$\sigma_{f1}^2$	1335.561 Corr: 1.000
<b>Nivel Aula (k)</b>	Lingüística	$\sigma_{v0}^2$	215.424 Corr: 1.000		
	Matemáticas	$\sigma_{v01}$	151.244 Corr: <b>0.645</b>	$\sigma_{v1}^2$	255.234 Corr: 1.00
<b>Nivel alumno (j)</b>	Lingüística	$\sigma_{u0}^2$	8088.865 Corr: 1.000		
	Matemáticas	$\sigma_{u01}$	3834.332 Corr: <b>0.469</b>	$\sigma_{u1}^2$	8275.556 Corr: 1.000

Por otro lado, es importante resaltar que en los modelos multivariados la eficiencia de los estimadores no se ve afectada por la falta de respuesta, es decir no todos los individuos tienen registradas las dos medidas (Rowe y Hill, 1998).

#### 4.3 Modelo multivariado de intercepto y pendiente aleatoria de una variable explicativa.

Información adicional que se obtiene con los modelos multivariados, es una matriz más compleja de covarianzas, al hacer aleatorio la pendiente de una variable explicativa a nivel centro escolar. Por ejemplo, continuando sólo con la variable género, estimar el efecto que tiene ser chico o chica en su rendimiento en lingüística y matemáticas, haciéndolo variar para cada centro escolar.

Tabla 4-8: Resultados del modelo multivariado de intercepto y pendiente aleatoria.

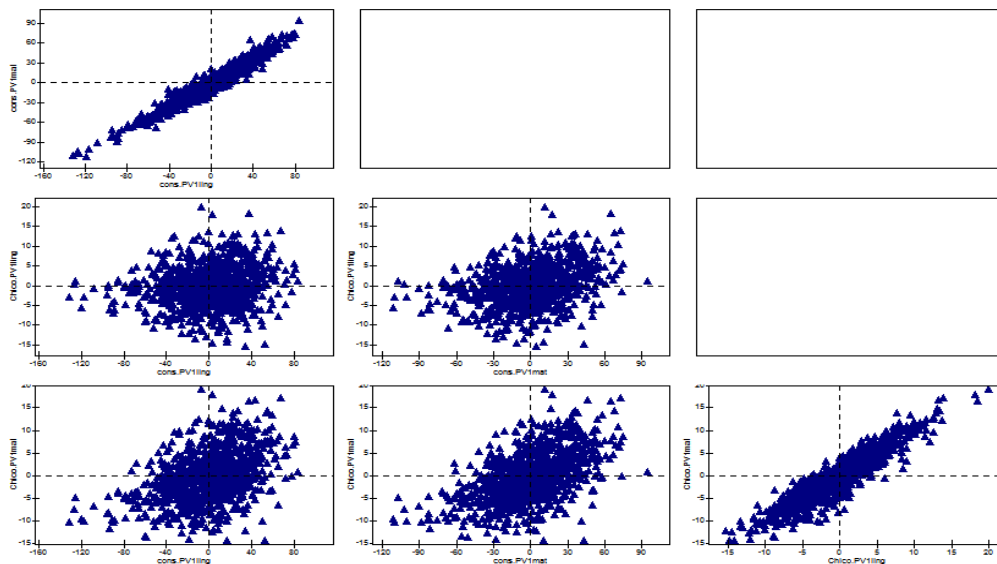
	Competencia Lingüística		Competencia Matemática		
	Estimador	SE	Estimador	SE	
<i>Parte fija</i>					
$\beta_{0jkl}$	506,13	(1,57)	$\beta_{1jkl}$	494,18	1,21
$\beta_{2l}Chico_{ijkl}$	-13,17	(1,18)	$\beta_{3l}Chico$	+11,13	1,21
<i>Parte aleatoria</i>					
$\sigma_{f0}^2$	1448,85	(106,3)	$\sigma_{f1}^2$	1228,9	(98,0)
$\sigma_{f2}^2$	136,15	(55,0)	$\sigma_{f3}^2$	162,67	(58,0)
$\sigma_{f01}$	1231,5	(91,09)	$\sigma_{f12}$	36,85	(42,1)
$\sigma_{f02}$	-34,27	(56,18)	$\sigma_{f13}$	64,03	(54,8)
$\sigma_{f03}$	42,05	(55,74)	$\sigma_{f23}$	134,18	(45,2)
$\sigma_{v0}^2$	215,8	(38,9)	$\sigma_{v1}^2$	254,82	(42,2)
$\sigma_{v01}$	15119	(32,5)			
$\sigma_{u0}^2$	8055,27	(71,1)	$\sigma_{u1}^2$	8235,51	(73,5)
$\sigma_{u01}$	3801,61	(56,6)			
-2*Loglikelihood 643594.036(54648 casos)					

Fuente: Elaboración propia.

Con un modelo multinivel multivariado de intercepto y pendiente aleatoria para una variable explicativa, se generan más parámetros: las varianzas en cada nivel y las covarianzas para ambas variables respuestas. Como se observa en la tabla 4-8.

De los resultados, destacan que hay una mayor variación entre los residuos a nivel escuela, para alumnos varones en la competencia en matemáticas ( $\sigma_{f3}^2 = 162,67$ ), que para la competencia en lingüística ( $\sigma_{f2}^2 = 136,15$ ). Es decir, que la escuela a la que asistan los chicos tendrá un efecto positivo para su rendimiento en matemáticas. Por el contrario, la covarianza de los residuos entre los centros escolares ( $\sigma_{f02}$ ) resultó negativa, lo que indica que centros escolares que tengan altos promedios para las chicas en lingüística, la diferencia de género tenderá a ser negativa. Esto es, que en esas escuelas, los chicos tendrán puntajes más bajos que las chicas. Aunque, es importante mencionar que la correlación a pesar de ser negativa, arrojó un valor bajo (-0,077). Es decir, que en este caso, será mínima la diferencia (5 puntos en promedio). Esta situación, se confirma al observar los gráficos de los residuos del cuarto nivel (centro escolar).

Figura 4-3: Residuos a nivel centro escolar.



Se observa la fuerte relación positiva que existe entre los resultados de los alumnos en su competencia lingüística y en su competencia en Matemáticas.

#### 4.4. Comparación de los modelos univariados y los modelos multivariados multinivel.

El mismo conjunto de factores para nivel alumno, aula y centro escolar, fue utilizado en cada uno de los modelos: el de competencia en lingüística, matemáticas y finalmente, el modelo multivariado que interesa analizar, en el que se incluyen ambas competencias.

Comparando los resultados de los modelos univariados, se aprecia como la mayoría de las variables resultó con el mismo efecto para ambas competencias (recuérdese que se especificaron dos modelos independientes), salvo para la variable género, que si tuvo un efecto contrario, es decir ser chico afecta más al rendimiento en lingüística que en matemáticas, las demás tuvieron el mismo efecto (aunque variaron en magnitud), además que las variables que no resultaron significativas para una competencia, no fueron las mismas que no tienen efecto en la otra.

Resaltando los principales resultados para los factores a nivel alumno, se observó en el modelo para Lingüística, que las variables que tienen un efecto negativo de gran magnitud, fueron : Iniciar el colegio después de los 6 años (-48,76), tener como lugar de origen de nacimiento África (-48,6), que los padres no tengan estudios (-84,15 y -50,93, para madre y padre, respectivamente), o los tengan incompletos (-35,77) y que nunca o casi nunca, el alumno tenga una buena relación con su madre (-46,38), así como las expectativas del propio estudiante sobre hasta qué nivel desea realizar sus estudios; si considera que sólo realizará los estudios obligatorios, su puntuación en lingüística será de (-46,79). Las variables que no resultaron significativas para esta competencia, fueron si el alumno habla Valenciano o Vasco la mayor parte del tiempo, si su relación con su padre no es buena o si su padre se encuentra en paro. Los efectos se interpretan individualmente, manteniendo controlados el resto de los factores.

Con respecto al modelo de la competencia en matemáticas, las variables de nivel alumno ( $\beta$ ) que tuvieron un mayor efecto negativo en los resultados de los alumnos fueron: que los padres no tengan estudios (-40,87 y -40,89, de la madre y del padre respectivamente), que el alumno haya nacido en otro país (-35,06) o en América (-22,8), que nunca o casi nunca, tenga una buena relación con su madre (-32,48) y que sus expectativas sean únicamente concluir sus estudios obligatorios (-44,7). Las variables que no tuvieron efecto en el rendimiento en matemáticas

fueron, tener como lugar de nacimiento África o Asia, hablar gallego, valenciano o vasco y mantener siempre una buena relación con la madre<sup>11</sup>.

Los factores relacionados con el alumnado, asociados a un mayor rendimiento en la competencia lingüística son ser niña, haber iniciado el colegio entre los 2 y 5 años de edad, haber nacido en España, usar el castellano como lengua principal, tener una buena relación con sus padres, que éstos tengan estudios universitarios, que se encuentren laborando y que sus expectativas sean continuar su formación académica hasta una formación universitaria. Los mismos factores se asociaron al rendimiento en competencia en matemáticas. Todos estos factores, se usaron como categoría de referencia e incrementaron el valor del intercepto, es decir el promedio del rendimiento pasó de 499,8 en el modelo nulo, a 578,75, considerando estas variables.

La estimación de los parámetros aleatorios muestra, como se aprecia en la Tabla 4-8, que la inclusión de estos factores redujo más la varianza entre alumnos y entre centros escolares, en la competencia lingüística que en la competencia en matemáticas. El coeficiente de correlación intraclase para este modelo fue de 4% para el cuarto nivel, proporción de la variación total de los residuos que se debe a diferencias entre las escuelas y de 2% entre las aulas.

Tabla 4-9: Resultados Comparativos para los componentes de la varianza de los modelos univariados en competencia en Lingüística y Matemáticas.

Parámetros aleatorios e intercepto	Modelo Competencia en Lingüística							
	Modelo nulo		Modelo con variables relacionadas con el alumno (Segundo nivel)		Modelo con variables relacionadas con el clima del aula (tercer nivel)		Modelo con variables relativas al centro escolar (Cuarto nivel)	
	Estimador	SE	Estimador	SE	Estimador	SE	Estimador	SE
$\beta_{0jk}$	499.57	(1.42)	578.75	(2.47)	566.4	(4.8)	563.7	(7.6)
$\sigma_{v0}^2$	1459.7	(95.1)	331.66	(47.6)	287.6	(53.6)	402.7	(107.2)
$\sigma_{u0}^2$	218.9	(35.4)	240.36	(43.9)	253.0	(52.0)	144.5	(90.4)
$\sigma_e^2$	8157.04	(70.9)	7080	(70.9)	7087.0	(77.9)	7264.8	(159.83)
-2 *Log-likelihood	330800.33		243536.511				51877.531	
Modelo en Competencia en Matemáticas								
	Estimador	SE	Estimador	SE	Estimador	SE	Estimador	SE

<sup>11</sup> La tabla con los resultados completos se encuentra en el ANEXO 2

$\beta_{0jk}$	499.44	(1.42)	562.13	(2.64)	560.47	(4.12)	549.26	(8.08)
$\sigma_{v0}^2$	1313.76	(90.0)	573.29	(59.58)	556.88	(65.45)	715.4	(136.04)
$\sigma_{u0}^2$	250.18	(41.8)	218.38	(45.33)	198.49	(51.5)	128.35	(90.62)
$\sigma_e^2$	8347.04	(72.4)	7592.33	(77.28)	7591	(83.8)	7292.82	(161.42)
<sup>-2</sup> *Loglikelihood	322801.36		243536.511				51385.208	

Fuente: Elaboración propia

Respecto a los factores relativos al tercer nivel o el clima del aula, se observó que algunas de las variables utilizadas no tuvieron un efecto significativo en el rendimiento de los estudiantes, es decir, no influye el género del profesor, su edad, su antigüedad ni las variables generadas para práctica educativa, materiales didácticos que utiliza el profesor o la forma de evaluar. El factor que sí tuvo un efecto fue el grado de satisfacción del profesor, es decir, alumnos en un aula donde el profesor se encuentra nada satisfecho o algo satisfecho, tendrán 11 y 15 puntos menos en promedio en Lingüística y Matemáticas respectivamente, manteniendo los demás factores invariables. Por lo anterior, se observa poca reducción en las varianzas en los distintos niveles (ver Tabla 4-9).

Esta situación refleja que la satisfacción del profesor es un determinante para propiciar un buen clima en el aula que favorezca el aprendizaje. Finalmente, al añadir las variables del cuarto nivel relativas al centro escolar, se observa que la titularidad del centro escolar sí tuvo un efecto en el modelo de competencia lingüística, siendo positivo. Es decir, los alumnos que asisten a colegios privados, tendrán un rendimiento 6 puntos por encima de aquéllos que asisten a colegios públicos. El tamaño del centro, ni la localidad en la que se ubica tuvieron un efecto significativo, como tampoco el porcentaje de alumnos extranjeros que se encuentran matriculados en dicho centro.

Los parámetros estimados para los componentes de la varianza, muestran como estos últimos factores reducen la variabilidad entre los residuos de nivel centro escolar, aula y alumnos. También se observa que la deviance o el estadístico de máxima verosimilitud (log-likelihood) para medir la eficiencia del modelo va disminuyendo, lo que indica que las variables consideradas en los distintos niveles constituyen un mejor modelo.

Los anteriores resultados se obtuvieron al comparar los valores de los parámetros estimados para cada modelo univariado. Ahora se analizan los resultados que se obtuvieron en el modelo

integrado, en el que se incluyen las variables respuestas (lingüística y matemáticas) simultáneamente. Lo primero que se observa es que para algunos estimadores disminuye la magnitud del efecto y también el tamaño de los errores estándar, lo que hace que el estimador sea más potente<sup>12</sup>.

Los factores relacionados con el alumno (segundo nivel) que resultaron tener un efecto opuesto para ambas competencias, fue el género, es decir, los niños tendrá un puntaje de 6 puntos por debajo del promedio de las niñas en lingüística, pero será 17 puntos mayor que el de las niñas, en la competencia en matemáticas. Si el alumno nació en Asia, su resultado en competencia matemática será 42,64 puntos por encima de los que nacieron en España, pero no tendrá un efecto en su competencia lingüística. Si el alumno nació en África, su rendimiento en competencia Lingüística será 30,53 puntos por debajo del promedio de los alumnos que nacieron en España, pero no tendrá un efecto en su competencia en matemáticas. La lengua que más utilizan para comunicarse, tienen un efecto negativo para la competencia lingüística pero no para matemáticas (-29,31 puntos por debajo del promedio si habla catalán, con respecto a los que hablan castellano, -19,47 si habla gallego y -29,89 si utiliza el vasco).

El iniciar entre los 2 y 5 años a asistir al colegio, tener una buena relación con la madre y con el padre, que los padres tengan estudios universitarios y que las expectativas del alumno sean continuar sus estudios hasta su formación universitaria, son factores que tienen un efecto positivo en ambas competencias (Anexo 3).

De esta manera, se confirma que el estimar un modelo integrado, se puede analizar y comparar con mayor facilidad, las variables que tienen un menor, mayor o igual efecto en ambas respuestas. Se observa en la tabla 7-8, con la estimación de los parámetros aleatorios, como se reduce la variabilidad entre los residuos a nivel alumno y a nivel centro escolar.

A nivel aula escolar, los factores relacionados con el estilo del profesor y que resultaron significativos en el modelo integrado fueron la utilización frecuente de los medios audiovisuales, para la competencia en matemáticas, resultó positiva generando un efecto de 44,76 puntos en promedio por encima, de los alumnos en cuyas aulas sus profesores lo utilizan

---

<sup>12</sup> Tabla completa en el Anexo B

con menor frecuencia y negativo el hecho de que los profesores promuevan debates en clases (-14,60), el resto de las variables no tuvo ningún efecto.

Tabla 4-8: Estimación efectos aleatorios modelo multivariado multinivel.

	Competencia Lingüística		Competencia Matemáticas	
	Estimador	SE	Estimador	SE
Parte fija				
$\beta_{0jkl}$	499,04	(1,46)	$\beta_{1jkl}$ 499,08	1,42
Parte aleatoria				
$\sigma_{f0}^2$	1499,2	(93,9)	$\sigma_{f1}^2$ 1334,5	(90,2)
$\sigma_{f01}^2$	1311,0	(84,5)		
$\sigma_{v0}^2$	216,0	(38,8)	$\sigma_{v1}^2$ 250,28	
$\sigma_{v01}^2$	151,29	(32,1)		
$\sigma_{u0}^2$	8161,75	(70,9)	$\sigma_{u1}^2$ 8351,5	(72,4)
$\sigma_{u01}^2$	3816,82	(56,1)		
-2*Log-likelihood 656213.96(55657 casos)				

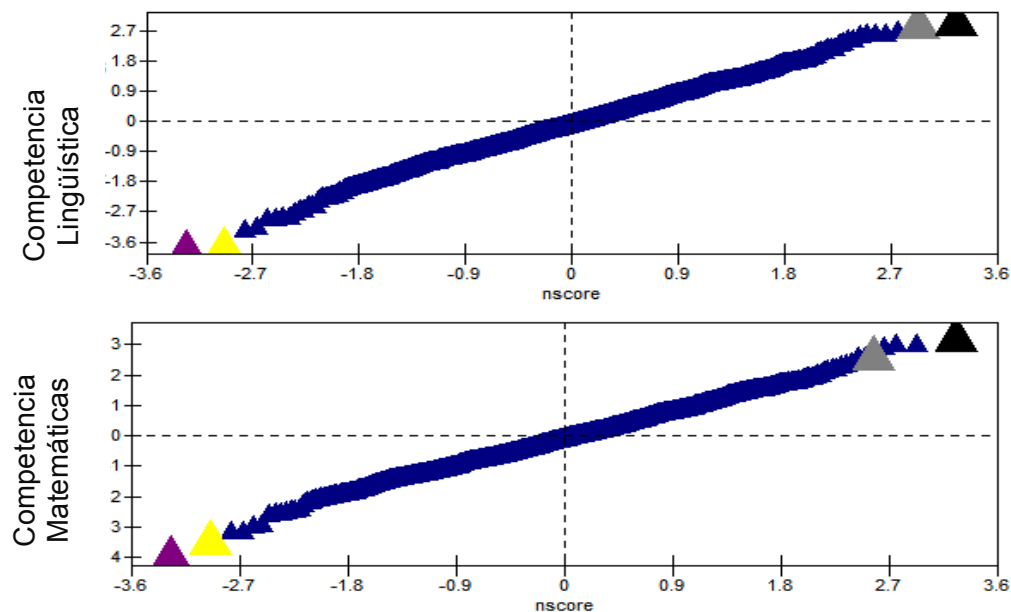
Del conjunto de variables sobre el centro escolar, la única que resultó significativa fue la titularidad del centro para la competencia lingüística, con 13,78 puntos en promedio, por encima de los alumnos que asisten a colegios públicos.

#### 4.5 Análisis de residuos.

Una de las principales aportaciones de los modelos multinivel (univariado o multivariado), con respecto a los modelos de regresión, es que los residuos, no solo sirven para comprobar los supuestos de normalidad del modelo o que únicamente se consideren errores; en este tipo de modelación, no se consideran como tal, sino que contiene una estructura de información interesante (Rasbash et al., 2009). En la siguiente figura (4-3), se observan los gráficos para el diagnóstico de los residuos estandarizados entre los distintos niveles y se comprueba el cumplimiento del supuesto de normalidad de los residuos. Además, se pueden identificar los centros escolares con un mejor desempeño, considerando todos los factores incluidos en el modelo.



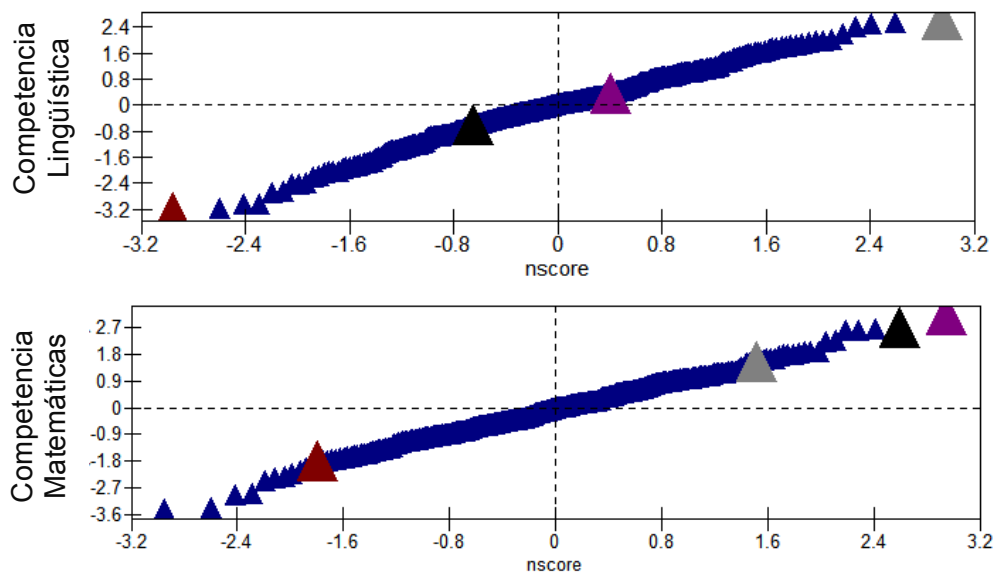
Figura 4-4: Residuos del modelo multivariado multinivel (nivel 4-Centros escolares)



Fuente: Elaboración propia

En la figura 4-4, cada uno de los puntos representan los 900 centros escolares analizados para cada una de las repuestas del modelo: Competencia lingüística y competencia matemática. Los que fueron resaltados y que se ubican en la parte superior, son aquéllos centros escolares con mejores resultados y aquéllos remarcados en la parte de abajo, son los centros con el desempeño más bajo. En la base de datos, los centros escolares sólo están identificados por número, no se tiene acceso a la información del nombre, pero sí es posible ver las variables asociadas con ese centro escolar. Por ejemplo, los dos centros escolares superiores, son de titularidad privada de tamaño pequeño, mientras que los de abajo son públicos con alto número de alumnos matriculados.

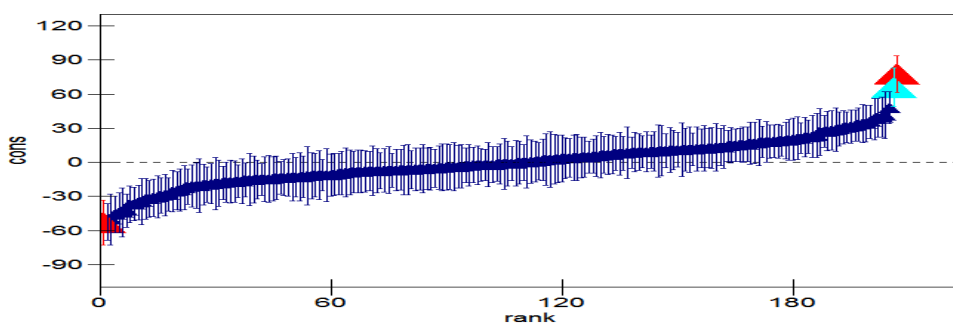
Figura 4-5: Residuos del modelo multivariado multinivel (nivel 3-Aulas)



En la figura 4-5, se muestran los residuos de nivel 3, correspondientes a las diferencias entre las aulas. Siguiendo el mismo proceso que con los residuos de nivel centro, se resaltaron las aulas escolares con el mejor desempeño (superiores) y con el desempeño más bajo (inferiores).

Otro gráfico que permite analizar los residuos, se muestra en la figura 5-4, en el que cada punto (representado por un triángulo), es uno de los 900 centros escolares analizados con un intervalo de confianza alrededor de él. Se puede observar que son pocos los centros escolares que no se traslapan con el cero y son las escuelas que difieren significativamente del promedio a un 5% de confianza. Aproximadamente son 50 centros escolares, el 5% de la muestra.

Figura 4-6: Gráfico Caterpillar (oruga) para residuos de nivel Centro Escolar.



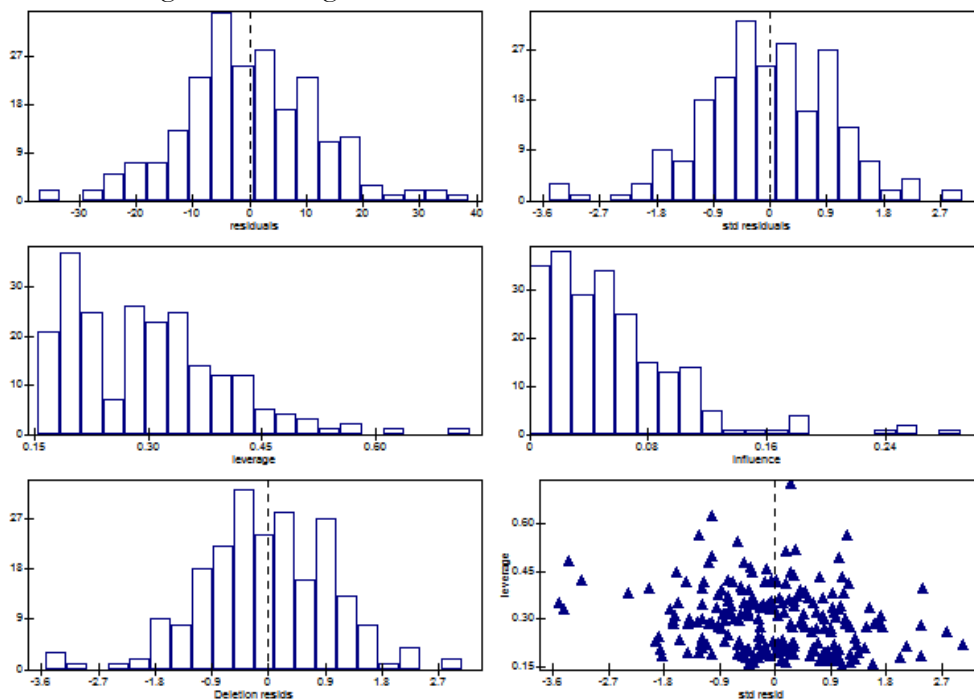
Fuente: Elaboración propia con el programa Mlwin

La Figura 4-6 muestra diferentes medidas de análisis de residuos a nivel centro escolar. En el primer gráfico de izquierda a derecha, se representa el histograma de los residuos nivel 4 y se confirma su comportamiento normalizado. El gráfico a su derecha, muestra los residuos estandarizados, que se obtienen al dividir el valor del residuo entre su error estándar (Goldstein, 2007). Este gráfico permite identificar las escuelas que estén a  $+2$  o  $-2$ , son las que difieren significativamente de la media a un 95% de confianza. Lo que se observa con los datos es que sí hay escuelas, aunque pocas, con estos valores.

Los gráficos de en medio, corresponden a histogramas de los valores de influencia (*Leverage values*), que se calculan usando la proyección de los valores ajustados del modelo. Un alto valor de influencia para una escuela determinada podría indicar que este dato está afectando las estimaciones del modelo.

Finalmente, el gráfico inferior de la izquierda, muestra el histograma de los residuos eliminados (deletion residual), que representan la desviación entre el intercepto para cada escuela y el intercepto medio para todas las escuelas (Rasbash et al., 2009). Como se está trabajando con una muestra grande de escuelas, el gráfico es similar al primero. El gráfico a la derecha relaciona los valores de influencia con los residuos estandarizados para identificar las escuelas con un comportamiento inusual.

Figura 4-7: Diagnóstico de residuos nivel Centro Escolar



# Conclusiones

## Conclusiones.

Desde la demostración de la modelación multinivel como herramienta metodológica más adecuada para el análisis de datos educativos, su aplicación ha permitido un desarrollo acelerado en la investigación escolar, particularmente en los estudios de eficacia que tienen por objeto la comprensión de los factores asociados al rendimiento. Gracias al uso de estos modelos, que consideran la estructura de anidamiento en los datos, se ha podido comprobar no sólo el efecto que tiene la escuela en el aprendizaje, sino de otros factores asociados a las distintas fuentes de variación que también están influyendo en los resultados educativos, como el clima del aula, la comunidad a la que pertenece la escuela o el país.

En España, el uso de esta herramienta comenzó en el año 2000, con diversos trabajos de Murillo (2000), sin embargo, aunque existe una gran cantidad de investigaciones aplicadas, el tema no se ha agotado. Como el mismo Murillo (2000) señala, apenas se sabe algo sobre los factores de eficacia escolar en el contexto educativo español. Si bien ya hay una mayor aproximación a las variables tanto del alumno, como de su contexto familiar, del clima del aula y del centro escolar que resultan significativos, los nuevos escenarios y configuraciones sociales que se están suscitando, requieren un estudio más profundo. La propuesta de este trabajo ha sido, explorar algunas contribuciones adicionales que tiene esta metodología para lograr más conocimiento en la complejidad de este fenómeno.

La mayoría de los trabajos han aplicado modelos de dos a tres niveles para cada una de las competencias que desean analizar, o bien se han centrado en el análisis de algunos factores en particular, como el género, la titularidad del centro o el origen del alumno. Blanco et al. (2014) inició con la línea de exploración de la modelación multivariada multinivel, es decir la elaboración de modelos integrados en los que se consideran dos variables respuestas simultáneamente y que permiten analizar las diferencias de variación en el rendimiento de los estudiantes en distintas competencias académicas, las relaciones entre ellas y observar si el efecto de los predictores es mayor, menor o igual en cada respuesta.

Los resultados que se presentan en este trabajo, permiten continuar esta línea para identificar algunas contribuciones adicionales de conocimientos que puede tener la aplicación de modelos

multivariados multinivel a la investigación escolar. En este caso, se trabajó con datos de la Evaluación General de Diagnóstico y no con datos PISA, por considerar que evalúan conocimientos más contextualizados. Se especificaron 3 modelos de tres niveles, uno para la competencia en Lingüística, otro para la competencia en matemáticas y el multivariado en el que se incluyen ambas competencias. Los niveles contemplados fueron el alumno, el aula y el centro escolar. Esto permitió comparar los resultados e identificar las nuevas aportaciones.

En primer lugar, se confirmó los factores que hasta ahora han encontrado autores como Murillo (2008), Calero y Escardíbul (2007), Peña Suárez et al. (2012) y Cervini et al. (2014). A nivel alumno, los más importantes son el género, al comprobar que los chicos obtienen mejores resultados en Matemáticas que las chicas y éstas en la competencia lingüística mejores que los chicos; la edad a la que inician el colegio, entre más temprana es, favorecerá su resultado. El tener una buena relación con sus padres influirá de manera positiva en sus logros y entre menos estudios tengan éstos, el efecto será negativo, como había anticipado Logan y Feiler (2006). Una variable de controversia considerada en este estudio fue el origen del alumno. Los resultados contraponen los hallazgos de Calero et al. (2012), que mencionan que la condición de inmigrante tiene un efecto negativo. En los resultados de las estimaciones, se observa que los alumnos cuyo origen es asiático, tienen mejores resultados en la competencia en matemáticas (42 puntos por arriba) que los alumnos de origen español. Los más afectados son los alumnos de origen africano en la competencia lingüística, situación que debe de estar explicada por la diferencia de idiomas. Esta variable muestra que lejos de representar una amenaza, la diversidad cultural debe significar una oportunidad.

Llama la atención, que las expectativas del alumno tienen un efecto significativo en sus resultados, entre más altas sean, mejor será su rendimiento. Lo que indica que se debe trabajar en motivar al alumno a formarse y no verlo como obligación. Respecto al aula escolar, las características del profesor ni su estilo de enseñanza tuvieron un efecto y en cuanto al centro, la titularidad tampoco resultó significativa, a diferencia de lo publicado por Calero et al. (2012), en la que demuestra que sí hay diferencias en los centros públicos y privados pero no tanto por la titularidad del centro, sino por los contextos de los alumnos que acuden a él.

No obstante, además de tener información sobre los factores, con el modelo multivariado multinivel, se confirmó la posibilidad de profundizar más en el tema del rendimiento. Al analizar la competencia en lingüística y matemática simultáneamente, se obtuvo la matriz de varianzas y covarianzas que permitió concluir que los centros escolares no favorecen el desarrollo de una u otra competencia, es decir, si el alumno obtiene alto puntaje en lingüística, hay una correlación positiva, de que obtendrá también puntajes altos en matemáticas.

Por tanto, las preguntas que permite responder la modelación multivariada multinivel, además de los factores asociados y la cantidad de variación explicada por cada nivel, son las siguientes, ¿Existe una correlación entre las competencias que adquiere un estudiante en el centro escolar?, ¿Alguna variable es favorecida por algún centro escolar? Por ejemplo, ¿Ser chica en un determinado centro tienen un efecto en el resultado, diferente del que tendría en otro centro escolar?, ¿Existe una relación entre la variación de los residuos en cada nivel analizado? Los residuos de estos modelos integrados, también mostraron proporcionar mayor información al permitir identificar centros escolares con resultados atípicos que pudiera estar afectando el valor de las estimaciones. Este tipo de análisis podría complementar los estudios de escuelas prototípicas, al tener una base más sólida que permita identificar estos centros escolares con resultados sobresalientes.

Se concluye que la profundización en la metodología multinivel, permite aportar nuevos conocimientos sobre investigación en eficacia escolar. Para futuras investigaciones, los trabajos que se consideren de nueva generación, deberán considerar modelos de mayor complejidad en el que se incluyan más de dos variables respuestas y con factores asociados a todos los niveles. Además, se debe replantear la variable que se desea explicar, es decir, ¿Hasta qué grado los resultados de pruebas estandarizadas representan los conocimientos y habilidades adquiridos en la escuela?

Líneas futuras de investigación.



## **Líneas futuras de investigación.**

Los resultados obtenidos en esta investigación, llevan a hacer algunos apuntes y consideraciones respecto a la misma, para identificar alcances y limitaciones que permitan el planteamiento de futuras líneas de investigación.

En primer lugar, es importante mencionar la complejidad del fenómeno analizado, particularmente con la utilización de los datos publicados por el Instituto Nacional de Evaluación Educativa que contienen resultados de la Evaluación General de Diagnóstico aplicada en el año 2009 a estudiantes del cuarto grado de Educación Primaria, de los diferentes centros escolares en todas las provincias de España, así como los resultados de las encuestas aplicadas a los padres de familia, profesorado y directores de las escuelas seleccionadas aleatoriamente en la muestra.

Respecto a los datos, insumo del modelo:

Se tiene una matriz  $\mathcal{M}_{n \times m}(\mathbb{R})$ , de  $\mathcal{M}_{483 \times 28.708}(\mathbb{R})$ . El tamaño de la matriz provoca que los datos presenten ruido que puede estar afectando las estimaciones del modelo. El porcentaje de no respuesta o datos faltantes no afecta, porque los modelos multinivel trabajan con datos no balanceados.

El 99% de las variables son categóricas, en las que su medición está expresada en más de 5 categorías. El excesivo número de variables dummy creadas para las estimaciones, podría estar generando estimadores imprecisos, aumentando el error estándar y afectando el requisito de parsimonia de los modelos.

Los factores utilizados como variables para explicar el nivel de competencia en los estudiantes, está basado en la información proporcionada por los propios alumnos, los padres de familia, el profesorado y los directivos. Sin embargo, se presentan discrepancias en las respuestas a una misma pregunta. Por lo que la calidad de la información presentada, podría no estar reflejando una correcta medición.

Respecto al modelo multinivel.

Representa una línea de investigación futura, medir la capacidad predictiva de los modelos multinivel. En esta investigación se obtuvo el coeficiente de correlación intraclase, que refleja el porcentaje explicado de la varianza total que se debe a las unidades de segundo nivel, pero es tema pendiente conocer la capacidad de acierto del modelo, así como trabajar en el cálculo de los errores estándar que permiten identificar la significancia de los efectos de los parámetros en las variables respuesta.

En la modelación multinivel, es el investigador quien decide qué variables explicativas se deben incluir en el modelo, aunque su selección debe estar basada y justificada teóricamente, puede representar un sesgo esta elección. Resulta necesario comparar los resultados que arrojan los modelos multinivel con respecto a otras técnicas, en las cuales no interviene el investigador en la selección de variable, sino el mismo modelo detecta el patrón de asociación en los datos, tales como las técnicas de minería de datos.

El modelo multinivel, ya sea de respuesta univariada o multivariada, asume el supuesto de linealidad en las variables, lo que significa suponer a la educación y sus factores como un fenómeno de causa-efecto y resulta un fenómeno mucho más complejo (Galán et al., 2014).

Referencias

Bibliográficas

## Referencias bibliográficas.

- Abraham, M. y Rojas, A. (1997). “La Investigación Educativa en Iberoamérica.” *Revista de Educación*, 312: 21–42.
- Achilli, E. L. (2009). “Investigación Educativa y Escuela Pública: Algunas Perspectivas y Debates. (spanish).” *Educational Research and the Public School: Some Perspectives and Debates*, 7(7): 29–38.
- Aitkin, M. A. y Longford, N. (1986). “Statistical Modeling in School effectiveness studies”. *Journal of the Royal Statistical Society. Series A*, 149: 1-43.
- Báez de la Fe, B.F. (1991). “El Movimiento de Escuelas Eficaces: Implicaciones Para La Innovación Educativa.” *Revista Iberoamericana de Educación*, 4: 93-116.
- Barrueco, A. (1992). *Calidad y eficacia en los centros educativos: análisis de las aportaciones de un modelo educativo*. Tesis doctoral inédita. Universidad de Salamanca.
- Blanco, A., López Martín, E. y Ruiz de Miguel, C. (2014). “Aportaciones de los modelos jerárquico-lineales multivariados a la investigación educativa sobre el rendimiento. Un ejemplo con datos del alumnado español en PISA 2009”. *Revista de Educación*, 365, DOI: 10.4438/1988-592X-RE-2014-365-267.
- Brookover, W. B., Beady, C., Flood, P., Schweitzer, J., y Wisenbaker, J. (1979). *School Social Systems and Student Achievement: Schools Can Make a Difference*. Nueva York: Praege.
- Brophy, J., and Good, T. (1984). *Teacher Behavior and Student Achievement*. The Institute for Research on Teaching.
- Bryk, A.S. y Raudenbush, S.W. (1992). *Hierarchical Linear Models*. Newbury Park, California. Sage.
- Calero, J. y Escardíbul, J. O. (2007). “Evaluación de servicios educativos: el rendimiento en los centros públicos y privados medido en PISA 2003”. *Hacienda Pública Española. Revista de Economía Pública*, 183 (4): 33-66.
- Calero, J., Escardíbul, J. O. y Choi de Mendizabal, A. (2012). “El fracaso escolar en la Europa mediterránea a través de PISA 2009: Radiografía de una realidad latente”. *Revista Española de Educación Comparada*. 19: 69-104.

- Carmena-López, G., Ariza-Cobos, A., Bujanda-Bujanda, M.E. y Muñoz-Repiso, M. (2000). *El sistema de investigación educativa en España*. Madrid: Ministerio de Educación y Cultura.
- Carroll, J.B. (1963). "A model for school learning". *Teachers College Record*, 64(8): 723-733.
- Castejón, J.L. (1994). "Estabilidad de diversos índices de eficacia de centros educativos". *Revista de Investigación Educativa*, 24: 45-60.
- Castejón, J.L. (1996). *Determinantes del rendimiento académico de los estudiantes y de los centros educativos: Modelos y factores*. Alicante, Editorial Club Universitario.
- Castro, M., y Lizasoain, L. (2012) "Las Técnicas de Modelización Estadística en la Investigación Educativa: Minería de Datos, Modelos de Ecuaciones Estructurales y Modelos Jerárquicos Lineales." *Revista Española de Pedagogía*, 251: 131-148.
- Cervini, R., Dari, N., y Quiroz, S. (2014). "Estructura Familiar y Rendimiento Académico En Países de América Latina: Los Datos del Segundo Estudio Regional Comparativo y Explicativo." *Revista Mexicana de Investigación Educativa*, 19(61): 569-597.
- Choi de Mendizabal, Á., y Calero-Martínez, J. (2009). "Determinantes del riesgo de fracaso escolar en España en PISA-2009 y propuestas de reforma." *Revista de Educación*, 362: 36-42.
- Coleman, J. S., Campbell, E. Q., Hobson, C., McPartland, J., Mood, A.M., Weifield, F., y York, R. L. (1966) *Equality for Educational Opportunity*. Washington, U.S. Government Printing Office.
- Dempster, A.P., Laird, M.N y Rubin, D.B. (1977). "Maximum Likelihood from incomplete data via the EM algorithm". *Journal of the Royal Statistical Society*, series B, 39: 1-38.
- Dunkin, M. y Biddle, B. (1974). *The study of teaching*. New York. Hult, Reinhart and Winston.
- Edmonds, R. (1979). *Effective schools for the urban poor*. Cambridge. Harvard University. Center for Urban Studies. Reprinted for CEMREL, Inc.
- Fabara, E., y Hernández, M.L. (2003). "La Educación y la Investigación Educativa en Iberoamérica." En *La Investigación Sobre Eficacia Escolar En Iberoamérica*, Murillo, F.A. (editor): 21-51.
- Galán, A., Ruiz Corbella, M., y Sánchez Melado, J.C. (2014). "Repensar la Investigación Educativa: de las Relaciones Lineales al Paradigma de la Complejidad." *Revista Española de Pedagogía*, 252(78): 281-298.

- García Durán, M. (1991). *Investigación evaluativa sobre las variables pedagógicas que discriminan entre los centros de EGB de alto y bajo rendimiento en el medio rural de la provincia de Cádiz*. Tesis doctoral inédita. UNED.
- Goldstein, H. (1996). “Consistent estimators for multilevel generalised linear models using estimated bootstrap”. *Multilevel Modelling Newsletter*, 8(1): 3-6.
- Goldstein, H. (2007). “Becoming Familiar with Multilevel Modeling.” *Significance*, 3 (4): 133–135.
- Goldstein, H., y Rasbash, J. (1993) “A Multilevel Analysis of School Examination Results.” *Oxford Review of Education*, 19(4): 22-37.
- Goldstein, H., y Woodhouse, G. (2000). “School Effectiveness Research and Educational Policy.” *Oxford Review of Education*, 26(3/4) : 353–363.
- Goldstein, H., Browne, W and Rasbash, J. (2002). “Partitioning Variation in Multilevel Models.” *Understanding Statistics*, 4(1): 223-231.
- Goldstein, H., Carpenter, J., Kenward, M y Levin, K. (2009) “Multilevel Models with Multivariate Mixed Response Types.” *Statistical Modelling: An International Journal* , 3: 173–197.
- González Galán, A. (2000). *Calidad, eficacia y clima en centros educativos. Modelos de evaluación y relaciones causales*. Tesis doctoral inédita. Universidad Complutense de Madrid.
- Herrero Olaizola, J., Rodríguez Díaz, F.J, y Musitu Ochoa, G. (2014) “Comparing Intergroup Contact Effects on Blatant and Subtle Prejudice in Adolescents: A Multivariate Multilevel Model.” *Psicotema*, 26 (1): 33–38.
- Hill, P.W., y Rowe, K.J. (1998). “Modelling Student Progress in Studies of Educational Effectiveness.” *School Effectiveness y School Improvement*, 3: 310–333.
- Hox, J.J y Kreft, I. (1994) “Multilevel Analysis Methods.” *Sociological Methods y Research*, 22 (3): 283–299.
- Hox, J.J. (2002). *Multilevel analysis: Techniques and Applications*. London, Lawrence Erlbaum Associates Publishers.
- INNE. (2010). *Evaluación General de Diagnóstico 2010. Informe de Resultados*. Madrid, Ministerio de Educación.

- Jencks, C.S., Smith, M., Acland, H., Bane, M.J, Cohen, D., Gintis, H., Heyns, B., y Michelson, S. (1972). "Inequality: a reassessment of the effect of family ad schooling in America". New York: Basic Books.
- Kallestad, J. (2000). "Teacher Emphases on General Educational Goals: An Approach to Teacher Socialisation in Norwegian Schools?" *Scandinavian Journal of Educational Research*, 44(2): 193–205.
- Kreft, I. (1993). "Using Multilevel Analysis to Assess School Effectiveness: A Study of Dutch Secondary Schools." *Sociology of Education*, 66 (2): 104–129.
- Kreft, I., de Leeuw, J. y Var der Leeden, R. (1994). "Review of Five Multilevel Analysis Programs: BMDP-5V, GENMOD, HLM, ML3, VARCL." *American Statistician*. 48 (4): 324-335.
- Laird, N.M y Ware, J.H. (1983). "Random effects models for longitudinal data. *Biometrics*, 38: 963-974.
- Leontief, W.(1936). "Quantitative Input and Output Relations in the Economic System of the United States". *The Review of Economic and Statistics*, 18: 105-125.
- Lindley, D.V. y Smith, A.F. (1972). "Bayes Estimates for the linear model". *Journal of Royal Statistical Society*, Series B, 34: 1-41.
- Lizasoain, L., y Joaristi, L. (2012). "Las Nuevas Tecnologías y La Investigación Educativa. El Análisis de Datos de Variables Catoriales". *Revista Española de Pedagogia*, 251(70): 111–130.
- Logan, E. y Feiler, A. (2006) "Forging Links between Parents and Schools: A New Role for Teaching Assistants?" *Support for Learning*, 21 (3): 115–120.
- Longford, N.T. (1987). "Hierarchical Linear Models: Applications and Data Analysis Methods (Book)." *Contemporary Sociology*, 22 (2): 293–295.
- Maas-Cora, J.M. y Hox, J. (2004). "Robustness Issues in Multilevel Regression Analysis." *Statistica Neerlandica*. 58(2): 127–137.
- Martínez-Garrido, C. (2011). "25 Investigaciones Clave En Eficacia Escolar." *Revista de Currículum y Formación del Profesorado* 15 (3): 149-174.
- Mayeske, G.W. (1972). *A study of our nation's schools*. Washington, D.C. US Department of Health, Education and Welfare.

- Moreira, E. (2009). “Factores Endógenos y Exógenos asociados al Rendimiento en Matemática: Un Análisis Multinivel”. *Revista Educación* 33 (2): 61–80.
- Mortimore, P. (2000). “Does Educational Research Matter?” *British Educational Research Journal*. 26(1): 5–24.
- Muñoz-Repiso, M., y Murillo, F.J. (2000). Effective School Improvement: Spanish case studies. En R. de Jong (Ed.), *Effective School Improvement Programmes. A description and evaluation of ESI programmes in eight European countries*. Groningen: GION: 379-443.
- Murillo, F.J.(2000). *Investigación iberoamericana sobre mejora de la eficacia escolar. Revisión Internacional del Estado del Arte*. Bogotá, Convenio Andrés Bello.
- Murillo, F.J. (2002). “Good Effective School Improvement Practices in Spain”. *Educational Research and Evaluation*, 8(4): 387-410.
- Murillo, F.J. (2003). “El Movimiento de Investigación en Eficacia Escolar.” En F.J. Murillo (ed.), *La Investigación sobre Eficacia Escolar en Iberoamérica*. Bogotá, Convenio Andrés Bello: 53–86.
- Murillo, F.J. (2008). “Los modelos multinivel como herramienta para la investigación educativa.” *Revista Internacional de Investigación en Educación*, 1 (1): 45–62.
- Murillo, F.J., y Román, M. (2009). “Mejorar el desempeño de los estudiantes de América Latina.” *Revista Mexicana de Investigación Educativa*, 41 (14): 451–484.
- OCDE. (2001). *Education Policy Analysis*. París.
- Penalva-Buitrago, J. (2007). “La Descentralización Educativa: Problemas de Aplicación.” *Revista Iberoamericana de Educación*, 42(5): 1-15.
- Peña Suárez, E., Campillo, A., Santarén, M. y Muñiz, J. (2012). “El papel de los centros escolares en la adquisición de competencias. *Revista Iberoamericana de Psicología y Salud*, 3(1): 75-87.
- Rasbash, J., Steel, F., Browne, W.J. y Goldstein, H. (2009). *A User's Guide to Mlwin. Version 2.10*. Center for Multilevel Modelling.
- Raykov, T. y Marcoulides, G. A. (2006). *A first course in structural equation modeling* (Segunda edición) Mahwah, NJ. Lawrence Erlbaum Associates.



- Rodríguez Díez, B. (1990). *Modelo para la evaluación externa de los centros escolares de EGB*. Tesis doctoral inédita. Universidad de Oviedo.
- Rodríguez Gómez, G. (1991). *Investigación evaluativa en torno a los factores de eficacia escolar de los centros públicos de EGB*. Tesis doctoral inédita. UNED.
- Rodríguez Pérez, A. (1984). *Algunos factores de ineficacia docente*. Tesis doctoral inédita. Universidad de Salamanca.
- Rowe, K.J. y Hill, P. (1998). "Modeling Educational Effectiveness in Classrooms: The Use of Multi-Level Structural Equations to Model Student's Progress." *Educational Research and Evaluation*, 4: 307-347.
- Rutter, M., Mortimore, O., Ouston, J., y Maughan, B. (1979). *Fifteen Thousand Hours*. Londres: Open Books.
- Sammons, P., Thomas, S. y Mortimore, P. (1997). *Forging links: effective schools and effective departments*. Londres: Paul Chapman Publishing.
- Scheerens, J., y Bosker, R.J. (1997). *The foundations of educational effectiveness*. Oxford: Pergamon.
- Scheerens, J., y Creemers, B.P.M. (1989). "Towards a More Comprehensive Conceptualization of School Effectiveness". En Creemers, B.P.M, T. Peters y D. Reynolds (Eds.), *School effectiveness and school improvement*, Lisse: Swets y Zeitlinger: 61-80.
- Skronvall, A. y Rabe-Hesketh, S. (2004). "Generalized latent variable modeling: Multilevel, Longitudinal and Structural Equation models". Londres, Chapman and Hall /CRC.
- Slavin, R.E. (1996). "Salas de Clase Efectivas, Escuelas Efectivas: Plataforma de Investigación para una Reforma Educativa en América Latina." *Programa de Promoción de la Reforma educativa en América Latina y el Caribe PREALC*. Documento de Trabajo N° 3.
- Smith, M. (1972). "Equality of Educational opportunity: the basic findings reconsidered". En F. Mosteller y D.P. Moynihan (Eds.) *On inequality educational opportunity*. New York. Vintage Books.
- Timm, N.H. (2000). *Applied Multivariate Analysis*. Springer.
- Tyler, W. (1991). *School Organization: A Sociological Perspective*. Croom Helm Ltd.
- Wall, W.D. y Varma, V.P. (1975). *Avances en Psicología de la Educación*. Editorial Morata.

Yi, G. y Judge, G. (1988). "Statistical Model Selection Criteria". *Economic Letters*, 28(1): 47-51.

# Anexos

## Anexos

### Anexo 1: Cuestionarios.

Nivel	Tipo perfil	Variables	Escala de medición
Alumnado	Socioeconómico	<p data-bbox="730 427 1054 461">Ítem 13 Cuestionario Alumnos</p> <p data-bbox="730 510 1011 544">13. ¿Qué tienes en tu casa?</p> <p data-bbox="730 593 1066 667">a) Una habitación para ti solo.</p> <p data-bbox="730 683 1066 757">b) Un sitio tranquilo para estudiar.</p> <p data-bbox="730 772 1023 806">c) Una mesa para estudiar</p> <p data-bbox="730 822 932 855">d) Un ordenador</p> <p data-bbox="730 871 995 904">e) Conexión a Internet.</p> <p data-bbox="730 920 959 954">f) Libros de lectura</p> <p data-bbox="730 969 959 1003">g) Libros de apoyo.</p> <p data-bbox="730 1019 943 1052">h) Un diccionario.</p> <p data-bbox="730 1068 938 1102">i) Un lavavajillas.</p> <p data-bbox="730 1205 1054 1238">Ítem 15. Cuestionario familias*</p> <p data-bbox="730 1288 1066 1321">¿Cuántos libros hay en su casa?</p> <p data-bbox="730 1370 884 1404">a) De 0 a 25</p> <p data-bbox="730 1420 895 1453">b) De 26 a 50</p> <p data-bbox="730 1469 911 1503">c) De 51 a 100</p> <p data-bbox="730 1518 922 1552">d) De 101 a 150</p> <p data-bbox="730 1568 900 1601">e) Más de 150</p>	<p data-bbox="1093 427 1278 461">Ítem 13: Variable</p> <p data-bbox="1093 477 1374 551">dicotómica de respuesta Sí o No.</p>

			Ítem 15: Variable ordinal de cinco categorías.
<b>Los padres de familia</b>	Nivel educativo y situación laboral	<p><b>Cuestionario Familias</b></p> <p>Datos de la madre</p> <p>Ítem 16. ¿Qué estudios realizó?</p> <p>a) Sin estudios</p> <p>b) No completó estudios obligatorios (EGB, ESO)</p> <p>c) Estudios obligatorios (EGB, ESO)</p> <p>d) Bachillerato</p> <p>e) Formación Profesional o equivalente</p> <p>f) Titulación universitaria media</p> <p>g) Titulación universitaria superior</p> <p>Ítem 17. ¿Cuál es su situación laboral?</p> <p>a) Trabaja fuera de casa por cuenta propia</p> <p>b) Trabaja fuera de casa por cuenta ajena</p> <p>c) Trabaja en casa de forma remunerada por cuenta propia</p> <p>d) Trabaja en casa de manera remunerada por cuenta ajena</p> <p>e) Está en paro</p> <p>f) Está jubilada</p>	<p>Ítem 16 y 20: Variable nominal de 7 categorías.</p> <p>Ítem 17 y 21: Variable nominal de 7 categorías que fue recodificada a cuatro, para conocer únicamente la información siguiente</p>

		g) Trabaja haciendo las tareas de casa	respecto a la situación laboral del padre y de la madre:
		Ítem 20 e Ítem 21, mismas preguntas para datos del padre.	
			a) Trabaja b) En paro c) Jubilación d) Tareas domésticas
			Para fines de este estudio, no se consideró la información de si el trabajo es autónomo.
El profesorado	1. Biográfico	Cuestionario profesores	Ítem 2: Variable ordinal de 5 categorías.
	2. Estilo de enseñanza.	1. Biográfico	
		Ítem 2: Edad	Ítem 4: Variable nominal de 7 categorías, recategorizada a 4.
		a) Menos de 30 años	
		b) 30 a 39 años	
		c) 40 a 49 años	
		d) 50 a 59 años	a) Menos de 9 años
		e) 60 años o más	b) De 10 a 19 años
			c) De 20 a 29 años
		Ítem 4: Antigüedad en la docencia.	d) Más de 30 años
		a) Menos de 5 años	
		b) De 5 a 9 años	Ítem 5: Variable nominal recategorizada a 4 categorías.
		c) De 10 a 14 años	
		d) De 15 a 19 años	
		e) De 20 a 24 años	
		f) De 25 a 29 años	
		g) Más de 30 años	
		Ítem 5: ¿Cuántos años, incluido	a) Menos de 9 años
			b) De 10 a 19 años

	<p>el presente, lleva en este centro?</p> <p>a) Menos de 5 años</p> <p>b) De 5 a 9 años</p> <p>c) De 10 a 14 años</p> <p>d) De 15 a 19 años</p> <p>e) De 20 a 24 años</p> <p>f) De 25 a 29 años</p> <p>g) 30 años o más.</p>	<p>c) De 20 a 29 años</p> <p>d) Más de 30 años</p>
	<p>Ítem 6: ¿Cuál es su situación administrativa o laboral?</p> <p>a) Funcionario/a con destino definitivo.</p> <p>b) Funcionario/a con destino provisional.</p> <p>c) Interino/a</p> <p>d) Contratado/a laboral indefinido.</p> <p>e) Contratado/a laboral temporal.</p> <p>f) Otra.</p>	<p>a) Destino definitivo</p> <p>b) Destino provisional</p> <p>c) Interino</p> <p>d) Otro</p>
	<p>2. Estilo de enseñanza.</p> <p>Ítem 23, 24 y 25</p>	
	<p>Cuestionario directores</p>	
El centro escolar	<p>Ítem 9: Localidad y % inmigración</p>	

Anexo 2: Resultados de los parámetros estimados en el modelo multivariado multinivel (Competencia lingüística y Competencia en matemáticas).

FACTORES ALUMNOS		Competencia Lingüística	Competencia Matemática
$\beta_{0jk}$			
Sexo	Chico	-6.87	2.46
Edad inicia	<2años	2.34	2.72
	>6años	-27.19	(7.66)
Lugar de nac.	País Europa	-34.55	(9.12)
	América	-29.06	(5.84)
	África	-30.53	(13.70)
	Asia	-7.64	(21.50)
	Otro país	-45.10	(19.39)
Lengua	Catalán	-29.31	(5.69)
	Gallego	-19.47	(6.25)
	Valenciano	-8.99	(15.59)
	Vasco	-29.89	(10.44)
Rel madre	Siempre	29.94	(4.96)
	Algunas veces	28.27	(5.43)
	Nunca o casi n	-8.96	(11.27)
Rel padre	Siempre	-1.60	(4.89)
	Algunas veces	10.26	(5.45)
	Nunca o casi n	-20.79	(8.86)
Est_madre	Sin estudios	-65.69	(10.17)
	Incompletos	-45.19	(5.36)
	Obligatorios	-28.81	(3.90)
	Byfp	-11.09	(3.42)
Est_padre	Sin estudios	-27.49	(10.24)
	Incompletos	-21.62	(5.12)
	Obligatorios	-17.54	(4.04)
	Byfp	-2.68	(3.64)
Sit Laboral	En paro		
	Jubilado		
	Lab. Hogar		
Satisfaccion	Nada		
	Bastante		
	Beta 0		



Anexo 3: Resultados de los parámetros estimados en el modelo univariado multinivel para la Competencia Lingüística (Modelo 1) y para la Competencia Matemáticas (Modelo 2).

FACTORES ALUMNOS			Modelo (Lengua) Modelo 1		Modelo (Matemáticas) Modelo 2	
$\beta_{0jk}$			499.5	1.4	499.4	1.4
	Sexo	Chico	-12.94	1.1	16.97	1.25
	Edad inicia	<2años	9.035	1.25	5.51	1.39
		>6años	-48.76	3.42	-26.83	4.22
	Lugar de nac.	Pais Europa	-22.64	3.88	-10.22	4.64
		America	-22.51	2.54	-22.81	2.45
		Africa	-48.66	6.250	-9.92	7.69
		Asia	-20.86	11.64	1.031	14.51
		Otro pais	-32.12	9.29	-35.06	11.62
	Lengua	Catalan	-18.31	4.25	-10.32	3.58
		Gallego	-18.26	6.05	-3.57	4.96
		Valenciano	-0.76	8.22	-2.83	7.77
		Vasco	-9.75	6.22	-6.10	5.69
	Rel madre	Siempre	11.83	1.37	-1.29	1.79
		Algunas veces	-24.9	2.43	-21.23	2.96
		Nunca o casi n	-46.38	4.37	-32.48	5.48
	Rel padre	Siempre	2.92	1.56	-11.46	1.86
		Algunas veces	1.66	2.57	-10.85	2.93
		Nunca o casi n	-14.85	3.71	-24.63	4.68
	Est_madre	Sin estudios	-84.15	4.60	-40.07	5.99
		Incompletos	-66.32	2.28	-36.94	2.80
		Obligatorios	-44.01	1.66	-25.50	1.99
		Byfp	-25.35	1.5	-15.42	1.72
	Est_padre	Sin estudios	-50.93	4.84	-41.89	5.51
		Incompletos	-35.77	2.46	-26.73	2.66
		Obligatorios	-25.82	1.93	-22.27	2.05
		Byfp	-12.64	1.71	-15.42	1.72
	Sit Laboral	En paro	-5.27	2.2	-4.42	4.51
		Jubilado	-8.84	4.38	-3.13	4.59
		Lab. hogar	-11.63	9.9	-14.40	10.54
	Expectativas	ESO	-46.79	2.05	-44.71	2.14
		Ciclo FP	-21.54	2.8	-21.49	2.91
		Bachillerato	-38.56	2.64	-33.90	2.71
		No sabe	-20.13	1.41	-19.31	1.47
	Ant profesor	<5 años	2.88	2.48	3.40	2.67
		20-29 años	0.00	0.00	0.00	0.00
		>30 años	0.00	0.00	0.00	0.00

	Procdida	Innovador	3.66	1.96	3.42	2.14
	Materiales	Innovador	1.25	2.08	2.95	2.97
	Evaluacion	Tradicional	-2.33	2.07	-5.97	2.27
	Satisfaccion	Nada	-11.50	5.76	-15.57	6.25
		Algo	-9.48	3.57	-9.98	3.26
		Bastante	-2.22	2.07	-4.03	2.25
		Beta 0	566.41	4.84	560.47	4.12
		V nivel 3	287.6	53.62	556.88	65.45
		V nivel 2	253.04	52.04	198.49	51.59
		V nivel 1	7087.05	77.91	7591.66	83.85
	FACTORES					
	CENTROS					
	ESCOLARES					
	Titularidad	Privado	6.22	2.27	3.148	5.53
		Privado no co	-11.41	9.32	-31.28	36.91
		%ALUml	-0.001	0.005	0.007	0.005
		%aludl	0.003	0.003	-0.003	0.003
	$\sigma_{v0}^2$		1459.7	95.1	1313.7	90.08
	$\sigma_{u0}^2$		218.03	39.2	250.1	41.8
	$\sigma_e^2$		8157.0	70.9	8347.4	72.4
	-2					
	*Loglikeli					
	hood					