

Archivos Analíticos de Políticas Educativas

Revista Académica evaluada por pares

Editor: Sherman Dorn

College of Education

University of South Florida

Editores Asociados para Español y Portugués

Gustavo Fischman
Arizona State University

Pablo Gentili
Laboratorio de Políticas Públicas
Universidade do Estado do Rio de Janeiro

Volumen 15

Número 8

Abril 2, 2007

ISSN 1068-2341

Efectos Contextuales del Nivel Socioeconómico sobre el Rendimiento Académico en la Educación Secundaria Obligatoria en la Comunidad Autónoma Vasca (España). Estudio Diferencial del Nivel Socioeconómico Familiar y el del Centro Escolar

Luis Lizasoain
Luis Joaristi
José Francisco Lukas
Karlos Santiago

Universidad del País Vasco, España

Citación: Lizasoain, L., Joaristi, L., Lukas, J. F. & Santiago, K. (2007). Efectos contextuales del nivel socioeconómico sobre el rendimiento académico en la educación secundaria obligatoria en la Comunidad Autónoma Vasca (España). Estudio diferencial del nivel socioeconómico familiar y el del centro escolar. *Archivos Analíticos de Políticas Educativas*, 15 (8). Recuperado [fecha] de <http://epaa.asu.edu/epaa/v15n8/>.

Resumen

El objetivo de este trabajo es analizar el efecto contextual del nivel socioeconómico (SES) sobre el rendimiento académico en Matemáticas y Lengua



Los lectores/as pueden copiar, mostrar, y distribuir este artículo, siempre y cuando se de crédito y atribución al autor/es y a Archivos Analíticos de Políticas Educativas, se distribuya con propósitos no-comerciales, no se altere o transforme el trabajo original. Mas detalles de la licencia de Creative Commons se encuentran en <http://creativecommons.org/licenses/by-nc-nd/2.5>. Cualquier otro uso debe ser aprobado en conjunto por el autor/es o AAPE/EPAA. AAPE/EPAA es publicada conjuntamente por el Mary Lou Fulton College of Education at Arizona State University y el College of Education at University of South Florida y, en su versión en español y portugués, por el Observatorio Latinoamericano de Políticas Educativas (OLPED) del Laboratorio de Políticas Públicas (LPP). Los artículos que aparecen en AAPE son indexados en el Directory of Open Access Journals <http://www.doaj.org> y por H.W. Wilson & Co. Contribuya con comentarios y sugerencias a Fischman@asu.edu y/o pablo@lpp-uerj.net

en la Educación Secundaria Obligatoria (ESO) en la Comunidad Autónoma Vasca (España), realizando un estudio diferencial del nivel socioeconómico familiar (SESF) y el del centro escolar (SESC). Para ello, en primer lugar, se verifica dicho efecto contextual mediante modelos jerárquicos lineales y se acepta la hipótesis de doble riesgo en el sentido de que los estudiantes de nivel socioeconómico familiar bajo obtienen un rendimiento académico aún más bajo cuando estudian en centros escolares de nivel socioeconómico bajo. Para estudiar con más profundidad el efecto diferencial de ambos indicadores del nivel socioeconómico hemos generado una variable que, para cada estudiante, combina los valores de ambos SES mediante el producto cruzado del SESF y SESC previamente categorizados. Mediante el empleo de técnicas de segmentación estadística (árboles de clasificación y regresión, CART), hemos hallado que los estudiantes de nivel socioeconómico familiar bajo que acuden a centros escolares de nivel alto son los que obtienen mejores resultados académicos, sólo superados por los estudiantes de nivel socioeconómico familiar alto que acuden a centros escolares también de nivel socioeconómico alto, pero por encima del resto de los subgrupos de la muestra, incluso de aquellos con mejores expectativas. El trabajo concluye con algunas hipótesis explicativas de este hecho y con recomendaciones para investigaciones posteriores.

Palabras clave: Comunidad Autónoma Vasca; España; educación secundaria obligatoria; nivel socioeconómico (SES); SES de la familia; SES del centro escolar; rendimiento académico; efecto contextual; Hipótesis de Doble Riesgo; Modelos Multinivel; HLM; Árboles de Clasificación y Regresión (CART).

Contextual effects of socioeconomic level on academic achievement in compulsory secondary education in the Basque Autonomous Community (Spain). Differential study about socioeconomic level of families and school centers

Abstract

The aim of this paper is to analyze the contextual effect of the socioeconomic status (SES) on the academic achievement in Mathematics and Language in Compulsory Secondary Education at the Basque Autonomous Community (Spain). We have carried out a differential study taking into account family SES and school SES in a multi-level study context. First, via tested hierarchical models, the hypothesis of the contextual effects (i.e., double jeopardy) is accepted, showing that the academic achievement of students from low SES families tend to worsen when they attend low SES schools. In order to illustrate the different effect of both SES, a new variable is generated so that, for each student, it combines the values of the previously categorized family and school SES. Using statistical segmentation techniques (regression and classification trees, CART), the present study has found that low family SES students attending high SES schools obtain the best academic achievement results, only outperformed by high family SES students who are studying at high SES schools, and also, even better than the sample subgroups who were expected to get much better scores. The study ends with some explanatory hypotheses about the findings and with some suggestions for further research.

Keywords: Basque Autonomous Community; Spain; compulsory secondary education; socioeconomic status (SES); family SES; school SES; academic

achievement; contextual effect; Double Jeopardy Hypothesis; Multilevel Models; HLM; Classification and Regression Trees (CART).

Antecedentes y Revisión de la Literatura¹

La relación causal que existe entre el nivel socioeconómico (SES)² y un amplio conjunto de resultados sociales está sólidamente apoyada en los resultados de la investigación. Y esto es así en el campo de la educación, en las ciencias sociales en general, y también en las ciencias biomédicas.

Willms (2003) afirma que esta relación es evidente desde el nacimiento y que persiste a lo largo de todo el ciclo vital. De forma que, a edad temprana, los niños de familias de SES bajo tienen más riesgo de padecer problemas de conducta y muestran un desarrollo cognitivo más bajo. Durante el periodo escolar, estos sujetos obtienen malos resultados académicos y tienden en mayor grado a abandonar prematuramente la escuela. En la adolescencia, tienen un mayor riesgo de padecer obesidad, de tener adicciones a las drogas, o de mantener conductas sexuales de riesgo. En la edad adulta, experimentan más problemas mentales y de salud y tienden a morir a edad más temprana.

Si nos centramos en el campo educativo, se acaba de cumplir el 40 aniversario de la publicación del informe Coleman (1966) en el que se analizaba el posible papel de la educación como elemento corrector de las desigualdades sociales de cara a lograr una real igualdad de oportunidades. En el mismo se concluía que ni la escuela, el currículo o los recursos tienen una influencia sustantiva en el rendimiento de los estudiantes. Por el contrario, factores como el nivel social, económico y cultural de las familias, de los propios centros educativos, del contexto global, son fuentes claras de desigualdades y están fuertemente asociados a las diferencias en rendimiento.

La incorporación del nivel socioeconómico como variable relevante en los estudios e investigaciones sobre el rendimiento académico nos la volvemos a encontrar en el trabajo de Jencks et al. (1972) donde se estudia el efecto del nivel socioeconómico familiar sobre los resultados escolares. Aquí además se reflexiona sobre las implicaciones que se derivan del empleo de diferentes variables como indicadores del SES, distinguiendo entre aquellas que se refieren al entorno meramente familiar de las relativas al barrio o vecindario, regiones o escuelas (pp. 77-78).

El trabajo de White (1982) es igualmente reseñable por varias razones. Con el fin de examinar la relación entre el SES y el rendimiento académico, White se planteó realizar un meta-análisis de los estudios e investigaciones centrados en tal cuestión y en los que el SES fuese un foco primario del estudio. Con esta premisa, empleando el Education Index, el Current Index to Journals in Education (CIJE), la base de datos ERIC y Dissertation Abstracts, se identificaron 248 estudios para su posible inclusión en el meta-análisis, de los que finalmente 101 fueron incluidos en el trabajo. Estos trabajos son investigaciones que se caracterizan por haber sido realizadas en un dilatado periodo de tiempo (el año de publicación oscila entre 1918 y 1976), llevadas a cabo mayoritariamente en Estados Unidos, aunque también hay algunas realizadas en Canadá y en Gran Bretaña; y en las mismas se estudia el problema en diferentes niveles o etapas escolares, con diferentes poblaciones de referencia (en algunos casos centrados en minorías

¹ Este estudio ha sido elaborado en el contexto de los siguientes proyectos de investigación: “Evaluación de la Educación Secundaria en la CAV. Estudio longitudinal y elaboración y validación de un modelo multinivel del rendimiento académico”. Proyecto UPV0218.230-H-15288. “Estudio longitudinal de las habilidades metacognitivas y de las estrategias de aprendizaje de estudiantes de enseñanza secundaria obligatoria y post-obligatoria”. Proyecto UPV 00218.230-H-15886/2004.

² Aunque la lengua de este artículo es el castellano, empleamos el acrónimo en inglés (*SES, Socio-economic Status*) en vez del equivalente español NSE por ser el primero de uso habitual en la literatura.

étnicas), y con diferentes conjuntos de variables empleadas como indicadores del SES. Una de las principales conclusiones de este meta-análisis es que la relación entre el SES familiar y el rendimiento académico es débil (0,22 empleando la mediana de las correlaciones como estimador) cuando se usa al estudiante como unidad de análisis. Pero este valor se incrementa hasta 0,73 cuando se usan unidades de análisis agregadas. Segundo, la fuerza de la relación varía de forma importante en función de cómo se defina el SES y de qué variables se empleen como indicadores.

Jeynes (2002) apunta que los investigadores incluyen la variable SES en sus trabajos por dos razones: primero, porque el SES está correlacionado con muchas variables tanto a nivel individual como grupal; y en segundo lugar, porque el SES explica mucha de la varianza que se encuentra entre las observaciones objeto de estudio. En la literatura podemos encontrar numerosos ejemplos de trabajos que estudian la relación entre el rendimiento académico (en general o en algunas materias específicas) y el SES familiar, tomando el estudiante como unidad de análisis. D'Angiulli et al. (2004a y 2004b) presentan los resultados de un estudio longitudinal sobre las relaciones entre el rendimiento en lecto-escritura y el SES de las familias. Después de realizar una exhaustiva relación de trabajos en esta misma línea concluyen que los resultados de su investigación longitudinal apuntan a que la relación entre ambos constructos es clara aunque tiende a atenuarse con el paso del tiempo. En el mismo campo, Hoff y Tian (2005) centran su trabajo en la relación entre el SES de las familias y las influencias culturales sobre el lenguaje.

Pero estos trabajos, y otros muchos que se podrían citar aquí, tienen en común que su foco de atención es el entorno familiar del estudiante y las dimensiones socio-económica y socio-cultural que habitualmente incorpora el SES. Sin embargo, la investigación educativa en este campo apunta no sólo a esto, sino a que se da un importante *efecto contextual* asociado a las características demográficas del aula o del centro escolar; efecto que tiene o puede tener una influencia incluso superior a los efectos individuales a nivel familiar.

Existen abundantes trabajos sobre el efecto contextual en educación. Y así, nos encontramos con trabajos que prestan atención a los posibles efectos de los centros escolares, efectos centrados en el tamaño de los centros o aulas, el estilo docente de los profesores, el grado de autonomía de los centros, el clima escolar, el nivel y uso de los recursos y la tecnología, etc. (Brookover et al., 1978; Henderson, Mieszkowski, & Sauvageau, 1978; Rumberger & Willms, 1992; Shavit & Williams, 1985; Willms, 1986; Tajalli & Ophelm, 2004; Howley & Howley, 2004; Warschauer et al., 2004). Pero la investigación existente apunta a que, entre todas las variables relativas a los centros, es de nuevo el SES—en este caso el nivel medio del centro—la que más influencia tiene sobre el rendimiento académico de los alumnos. Estudios como el Program for International Student Assessment (PISA) muestran que “el nivel socioeconómico medio de una escuela tiene un efecto sobre el rendimiento del estudiante, efecto que se superpone y supera a los efectos asociados al propio nivel socioeconómico familiar del estudiante” (Willms, 2004). De forma que si un estudiante de SES medio acude a un centro de SES alto, su rendimiento será probablemente mayor que si hubiese asistido a uno de SES bajo.

Con respecto al sistema educativo español, en el trabajo de Álvaro et al. (1990) se propone un modelo causal del rendimiento académico, estableciendo dos tipos de factores contextuales: los del ámbito sociofamiliar y los estrictamente escolares. Por otra parte, tanto las evaluaciones a nivel nacional realizadas por el Ministerio de Educación a través del Instituto de Evaluación, como las evaluaciones a nivel internacional de PISA y TIMSS realizadas en España apuntan en el mismo sentido. En relación a la educación en el País Vasco, España tiene una estructura semifederal basada en las denominadas Comunidades Autónomas, y la educación es una competencia transferida enteramente a las mismas. En función de ello, la gestión y evaluación del sistema educativo vasco es competencia de la Consejería de Educación del Gobierno Vasco y las tareas de evaluación son realizadas por el Instituto Vasco de Evaluación e

Investigación Educativa. Las publicaciones, documentos y bancos de datos relativos a dichas evaluaciones están disponibles en sus respectivos sitios Web.³

Como vemos, hay estudios que han prestado una mayor atención a los efectos individuales, mientras que otros se han centrado en los contextuales, fundamentalmente desde la perspectiva de los centros escolares. Pero la cuestión no radica en comparar estos enfoques, sino en poder estudiar simultáneamente ambos tipos de efectos. Dicho de otra forma, aquí nos encontramos con que los datos se encuentran anidados en estructuras jerárquicas. En un primer nivel jerárquico se encuentran las variables relativas a los estudiantes y sus familias (el nivel socioeconómico familiar). A su vez, estos estudiantes se agrupan por centros educativos, y es a este segundo nivel jerárquico al que pertenece una variable como la media del nivel socioeconómico de las escuelas o centros. Y dado que, según acredita el corpus de investigación al que hemos hecho referencia, ambos niveles resultan relevantes para el problema planteado, parece recomendable emplear un enfoque analítico que nos permita examinar esta situación respetando y asumiendo esta estructura multinivel.

Los modelos jerárquicos lineales (HLM, por sus siglas en inglés) (Raudenbush y Bryk, 2002; Goldstein, 1987; Goldstein, 1995) nos proporcionan una herramienta analítica para examinar y analizar datos anidados, para examinar simultáneamente tanto las variables relativas al individuo como las variables relativas a su contexto, de forma que se puedan tener en cuenta las influencias mutuas entre los diferentes niveles de la estructura. Muchos de los estudios sobre el tema que nos atañe emplean modelos jerárquicos como herramienta de modelización. Entre otros, podemos mencionar los siguientes: Willms (2003, 2004), Kobrin (2003) y Lapsley (2002) analizan modelos a tres niveles (estudiantes, aulas y centros); Välijärvi y Malin (2003) comparan los resultados de PISA de los países nórdicos europeos en relación al nivel socioeconómico de familias y centros escolares; o el también ya citado de Howley y Howley (2004), centrado en el tamaño de los centros escolares y el SES. Gaviria, Arias y Castro (2004) estudian los factores de eficacia escolar en el sistema educativo brasileño mediante la validación de un modelo de tres niveles (estudiantes; profesores y escuelas; estados federados brasileños). Finalmente, centrado en el sistema educativo español, el trabajo de Marchesi y Martínez Arias (2002) presenta un modelo multinivel (estudiantes y centros) para evaluar la educación secundaria obligatoria.

Asumiendo este enfoque multinivel, la finalidad de este trabajo es examinar las relaciones entre el nivel socioeconómico familiar (SESF) y la media del nivel socioeconómico de los centros escolares (SESC) con respecto al rendimiento académico de estudiantes de enseñanza secundaria obligatoria (ESO) en la Comunidad Autónoma Vasca⁴. Más en concreto, nuestro primer objetivo se centra en comprobar la existencia o no del “efecto contextual.” Para tal fin se han formulado y validado diversos modelos jerárquicos lineales (HLM), planteados desde la perspectiva de contrastar las hipótesis de *doble y triple riesgo* relativas a los efectos contextuales del SES (Willms, 2003).

La *hipótesis de doble riesgo* formula el siguiente supuesto: el efecto contextual supone que los estudiantes de nivel socioeconómico familiar bajo están en una situación de *doble riesgo*, porque al hecho de que están en desventaja en razón de su bajo status socioeconómico familiar (primer

³ La página Web oficial del Instituto de Evaluación del Ministerio Español de Educación es <http://www.ince.mec.es>. La documentación relativa a las evaluaciones del Ministerio de Educación de España y a las evaluaciones PISA y TIMSS se encuentra en <http://www.ince.mec.es/pub/index.htm>. La del Instituto Vasco de Evaluación e Investigación Educativa es <http://www.isei-ivei.net>, y la documentación relativa a la Comunidad Autónoma Vasca está disponible en <http://www.isei-ivei.net/cast/pub/indexpub.htm>

⁴ Aunque en el apartado dedicado a las variables, describiremos algunas características de la ESO, adelantemos aquí que en el sistema educativo vasco—al igual que en el conjunto del sistema educativo español—el periodo de educación obligatoria se extiende desde los 6 a los 16 años y está dividido en dos partes: la enseñanza primaria (los primeros seis cursos) y la secundaria, con una duración de cuatro cursos. A su vez, la ESO se estructura en dos ciclos de dos cursos cada uno.

riesgo), se añade el que esta situación desfavorable empeora aún más cuando dichos estudiantes asisten a centros escolares cuyo nivel socioeconómico medio es igualmente bajo (segundo riesgo). Es decir, los sujetos de familias de bajo nivel socioeconómico obtendrán un rendimiento aún peor cuando estudien en centros escolares cuyo status socioeconómico sea también bajo.

La *hipótesis de triple riesgo* se formula desde el supuesto básico de que la posible interacción entre el SESF y el SESC podría tender a favorecer a los estudiantes de mayor nivel socioeconómico familiar. Es decir, de todos los estudiantes que se encuentren en un contexto desfavorable (un centro escolar de bajo nivel socioeconómico medio), aquellos que tengan la posibilidad de poder disponer de recursos adicionales podrían mejorar su situación, es decir, superar o paliar el efecto contextual. Por el contrario, los estudiantes de nivel socioeconómico familiar bajo no pueden contar con dichos recursos y sus expectativas no pueden mejorar (triple riesgo). Por poner un ejemplo, si el efecto contextual se verifica, podríamos aducir que los estudiantes que acudan a un centro escolar de nivel socioeconómico bajo tenderán a obtener un bajo rendimiento académico, pero si las familias de algunos de ellos disponen de medios económicos suficientes, podrían contratar profesores o clases particulares que contribuyesen a mejorar el rendimiento de sus hijos, mientras que los estudiantes de nivel socioeconómico familiar bajo no tendrían esta oportunidad. Dicho más brevemente, la hipótesis de triple riesgo plantea que el efecto contextual es *más fuerte, más intenso* en sujetos de familias de bajo nivel socioeconómico.

En segundo lugar, y para analizar con más detalle los efectos detectados, nos planteamos realizar un “estudio diferencial” que nos permita comparar las diferencias en rendimiento en los diferentes subgrupos resultantes de considerar simultáneamente los dos niveles de SES. Para ello, en primer lugar se categorizan las dos medidas del nivel socioeconómico en sendas nuevas variables de cuatro valores cada una: niveles socioeconómicos altos, medio-altos, medio-bajos y bajos; tanto para las familias como para las escuelas. Al realizar el producto cruzado de ambas variables, se dispone para cada individuo de una nueva variable de 16 valores en función de todas las combinaciones posibles entre ambos niveles del SES. De esos 16 valores, el primer dígito representa al nivel socioeconómico del centro escolar y el segundo al familiar, de forma que la variable oscila entre 11 (centro y familia de niveles socioeconómico bajo) hasta 44 (centro y familia de niveles socioeconómico alto).

Segundo, empleando técnicas estadísticas de segmentación (árboles de clasificación) se realiza el “estudio diferencial” mediante la comparación de los componentes de los diferentes subgrupos que se generan en los diferentes niveles de segmentación en función de las 16 posibles subclases establecidas. Por ejemplo, si partimos la muestra en dos grandes grupos en función de su media aritmética en rendimiento (subgrupos superior e inferior), ¿cuáles de las 16 clases formarán parte de cada uno?, ¿cuáles serán los mayoritarios: los *cuarentas* y los *treintas*?, ¿o los *catorces*, *veinticuatro*, *treinta y cuatro* y *cuarenta y cuatro*? Una vez realizada esta primera segmentación, se puede repetir el procedimiento en el nivel inmediatamente inferior volviendo a partir en dos cada uno de los subgrupos antes obtenidos lo que nos daría una estructura de cuatro subgrupos en función de la media en rendimiento (superior, medio-superior, medio-inferior, inferior). Este proceso continua hasta que se satisface alguno de los criterios de parada del algoritmo, de manera que al final se dispone de una estructura arborescente que resulta de utilidad para caracterizar los diferentes subgrupos obtenidos en función de las variables introducidas en el modelo y así estudiar con más detalle las relaciones entre los niveles socioeconómicos y el rendimiento académico (véase como ejemplo la figura 3).

Método

Muestra

Desde 1997, el proyecto REDES (Red de Evaluación de Centros) del Instituto de Evaluación y Asesoramiento Educativo (IDEA) dependiente de la Fundación Santa María, ha

venido realizando evaluaciones externas en aproximadamente 150 centros de educación secundaria de diferentes Comunidades Autónomas españolas. En este contexto, desde el curso 1999-2000 se ha estado llevando a cabo el proyecto denominado *Evaluación de Centros Educativos del País Vasco* en la Comunidad Autónoma Vasca (CAV). Ha sido una evaluación externa realizada a centros de educación secundaria por iniciativa del Gobierno Vasco que fue la entidad financiadora.

El desarrollo del proyecto fue llevado a cabo por profesores e investigadores del departamento de Métodos de Investigación y Diagnóstico en Educación de la Universidad del País Vasco-Euskal Herriko Unibertsitatea contando con la colaboración de IDEA. El plan de evaluación estaba diseñado para 4 cursos, finalizando en Junio de 2003. Las características de estos proyectos de evaluación auspiciados por REDES e IDEA, sus niveles y dimensiones, así como la presentación de diversos resultados obtenidos se encuentran en las siguientes publicaciones: Equipo REDES (1999 y 2000) y Marchesi y Martín (2002) para el proyecto estatal; y en Lizasoain et al. (2003); Lukas y Santiago (2004) y Santiago y Lukas (2005) para el proyecto realizado en la CAV.

Para iniciar el proceso de evaluación, la Consejería de Educación del Gobierno Vasco⁵ tomó la decisión de realizar una convocatoria pública dirigida a todos los centros educativos—tanto públicos como privados—de Enseñanza Secundaria Obligatoria de la Comunidad Autónoma Vasca que quisiesen tomar parte en el proyecto. Del total de centros educativos de este nivel de la CAV (346 en esa fecha), 55 expresaron su intención de tomar parte en dicha evaluación. Al final, las autoridades educativas vascas optaron por no excluir a ningún centro de los que se habían presentado, por lo que finalmente los 55 centros escolares han tomado parte en la evaluación.

En este estudio, las pruebas a los estudiantes se aplican en tres momentos. La primera aplicación tiene lugar al comienzo del 1º curso de la ESO (normalmente en el mes de octubre, casi recién iniciado el curso), por lo que equivale a hacer una evaluación final de la educación primaria o, si se quiere, una evaluación inicial de la educación secundaria obligatoria. La segunda aplicación tiene lugar al terminar el 2º curso de la ESO (meses de abril o mayo), de forma que este segundo momento coincide con el final del primer ciclo de la ESO. Por último, la tercera aplicación se lleva a cabo al terminar 4º, lo que se corresponde con el final del segundo ciclo de la ESO y de la propia ESO en su conjunto. Por tanto, los estudiantes pertenecen a tres cursos: 1º, 2º y 4º, aunque insistimos que los datos de primer curso han de considerarse como la evaluación final de la primaria.

La Consejería de Educación, de la misma manera que decidió que fuesen evaluados todos los centros que se habían presentado a la convocatoria, optó por seleccionar a *todos* los estudiantes de 1º, 2º y 4º de los 55 centros escolares participantes. A este respecto, está claro que esto constituye una primera limitación de nuestro trabajo, pues, a la vista del procedimiento seguido, en ningún caso podemos hablar de una muestra aleatoria de la población escolar de secundaria de la CAV. De cualquier manera, establecido esto, también hay que hacer notar que tanto por el tamaño final de la muestra como por la distribución de variables como el curso o la titularidad de los centros escolares, podemos afirmar que los datos obtenidos pueden considerarse representativos de los estudiantes y centros de enseñanza secundaria obligatoria en la CAV, aunque con una ligera sobre-representación de centros privados en detrimento de los públicos (Lukas, 2000).

Esta afirmación se justifica si examinamos las tablas que aparecen a continuación. En la Tabla 1 aparece el número total de estudiantes distribuido por cursos, tanto los participantes en la evaluación como el total de la población estudiantil de la secundaria obligatoria. Las Tablas 2 y 3 presentan, también a nivel muestral y censal, la distribución de estudiantes y centros escolares

⁵ La Consejería de Educación del Gobierno Vasco es un órgano administrativo equivalente a un ministerio, pero a nivel de Comunidad Autónoma.

respectivamente en función de la titularidad pública o privada de estos últimos. A este respecto hay que hacer notar que en el conjunto de la CAV en la secundaria obligatoria, las proporciones son de casi el 58% para centros privados y el 42% para públicos. Son, junto con Madrid y Barcelona, las proporciones más elevadas de enseñanza privada en España. La Tabla 4 proporciona la distribución conjunta del curso y la titularidad⁶. Por último, en la Tabla 5 cruzamos la titularidad (pública/privada) con el número de estudiantes (de 15 años de edad) y el número de centros educativos de la muestra seleccionada para realizar el estudio PISA 2003 en la CAV.⁷ Pensamos que estos datos contribuyen a reafirmar lo dicho con respecto a una cierta representatividad de la muestra a pesar del procedimiento seguido.

Tabla 1

Distribución de los Estudiantes por Cursos en la Muestra y en la Población

Curso	Muestra (%)	Población (%)
1º ESO	2812 (29,2)	17997 (28,6)
2º ESO	3451 (35,8)	22288 (35,4)
4º ESO	3378 (35,0)	22713 (36,1)
Total	9641	62998

Tabla 2

Distribución de los Estudiantes según la Titularidad en la Muestra y en la Población

Tipo	Muestra (%)	Población (%)
Público	3223 (33,4)	26611 (42,2)
Privado	6418 (66,6)	36387 (57,8)
Total	9641	62998

Tabla 3

Distribución de los Centros según Titularidad en la Muestra y en la Población

Tipo	Muestra (%)	Población (%)
Público	17 (30,9)	133 (38,4)
Privado	38 (69,1)	213 (61,6)
Total	55	346

Tabla 4

Distribución de los Estudiantes según Cursos y Titularidad en la Muestra y en la Población

Curso	Muestra (%)		Población (%)	
	Público	Privado	Público	Privado
1º ESO	933 (33,2)	1879 (66,8)	7140 (39,7)	10857 (60,3)
2º ESO	1148 (33,3)	2303 (66,7)	9486 (42,6)	12802 (57,4)
4º ESO	1142 (33,8)	2236 (66,2)	9985 (44,0)	12728 (56,0)
Total	3223 (33,4)	6418 (66,6)	26611 (42,2)	36387 (57,8)

⁶ Los datos de la población de ESO de la CAV han sido extraídos de las estadísticas oficiales sobre el sistema educativo vasco disponibles en el Instituto Vasco de Estadística (EUSTAT) y en la propia Consejería de Educación del Gobierno Vasco. Sus direcciones respectivas son:

<http://www.eustat.es/estad/temalista.asp?tema=19&idioma=c&opt=0>

http://www.hezkuntza.ejgv.euskadi.net/r43-573/es/contenidos/informacion/dia4/es_2025/a4_c.html

⁷ Estos datos han sido extraídos del informe PISA 2003 en la CAV, disponible en <http://www.isci-ivei.net/cast/pub/PISA2003euskadic1.pdf>.

Tabla 5
Muestra Elaborada por el Consorcio PISA 2003 para la CAV

Tipo	Centros	Estudiantes
Público	58 (40,8)	1502 (38,4)
Privado	84 (59,2)	2407 (61,6)
Total	142	3909

Variables

Ya hemos apuntado que nuestro objetivo general es examinar la relación entre el nivel socioeconómico de estudiantes y centros escolares y el rendimiento académico de estudiantes de la ESO en la CAV. Más concretamente, las variables dependientes son el rendimiento académico en las materias de Matemáticas y Lengua. Pero antes de pasar a exponer las cuestiones relativas a éstas y a las demás variables, conviene que contextualicemos describiendo brevemente la organización y currículo de la ESO en el sistema educativo vasco y español.

El sistema educativo español, al igual que el sistema educativo vasco, cuenta con la siguiente estructura. La Educación Infantil abarca el periodo comprendido entre los cero y los seis años y es una etapa formativa de carácter voluntario. De los 6 a los 16 años se extiende el periodo obligatorio y gratuito estructurado en dos partes: la Educación Primaria (6 cursos, de los 6 a los 12 años) y de los 13-16 años la Educación Secundaria Obligatoria (ESO) que completa la educación básica y que abarca cuatro cursos académicos estructurados en dos ciclos de dos cursos cada uno. A los 16 años finaliza la escolaridad obligatoria y la finalidad de la ESO es que a esa edad los alumnos hayan adquirido y dominen los elementos fundamentales de la cultura, de forma que al finalizar la etapa estén preparados, bien para la incorporación a la vida activa, o bien para continuar sus estudios accediendo a la formación profesional específica de grado medio o al bachillerato (educación secundaria post-obligatoria). El nivel más alto lo ocupa la enseñanza superior (estudios universitarios o formación profesional de grado superior).⁸

Las materias que configuran el currículo de la ESO y que son comunes a los cuatro cursos son las Ciencias de la Naturaleza; las Ciencias Sociales, Geografía e Historia; la Educación Física; la Educación Plástica y Visual; la Lengua Castellana y Literatura; la Lengua Extranjera; las Matemáticas; la Música, la Tecnología y la Religión para aquellas familias que así lo decidan o en su caso actividades de estudio. En aquellas Comunidades Autónomas con idioma propio además del español, también es materia básica la Lengua y Literatura específica de la misma. Y éste es el caso de la Comunidad Autónoma Vasca donde la lengua vasca es lengua co-oficial y forma parte del currículo tanto como materia de estudio como lengua vehicular del mismo.

Una de las especificidades más notables del sistema educativo vasco es su organización en modelos lingüísticos establecidos en función de cuál sea la lengua vehicular del proceso de enseñanza-aprendizaje. Así existen tres modelos: el "A", el "B" y el "D" (en lengua vasca no existe la letra "C", de ahí que el tercero sea el "D"). Como ha quedado dicho, en los tres modelos se estudian las dos asignaturas relativas a ambas lenguas y literaturas. La diferencia estriba en que en el modelo "A" la enseñanza se imparte en castellano y la lengua vasca se enseña como una asignatura, mientras que en el modelo "D" la situación es a la inversa: todas las materias se imparten en vasco y el castellano es una asignatura. En el modelo "B", unas materias se enseñan en vasco y otras en español.

En otro trabajo (Lizasoain et al., 2003) ya abordamos lo relativo al papel del modelo lingüístico en la evaluación del rendimiento de los estudiantes de esta comunidad. En éste, y con estos antecedentes, vamos a centrar nuestra atención en analizar fundamentalmente la influencia

⁸ Un esquema del sistema educativo español se encuentra en <http://www.mec.es/educa/sistema-educativo/logse/siseduc.html>.

del nivel socioeconómico familiar y del nivel socioeconómico medio de los centros con respecto al rendimiento académico en matemáticas y lengua española, dos de las materias obligatorias y básicas del currículo de la ESO. Nos hemos limitado a estas dos materias porque son consideradas como “materias instrumentales básicas” por las regulaciones normativas de los sistemas vasco y español. Y hemos optado por la lengua española en vez de la lengua vasca porque, tal y como estudiamos en el trabajo antes citado, con respecto al aprendizaje de la lengua vasca el *modelo* es crucial. Al igual que sucede en el aprendizaje de otras lenguas, existe una clara diferencia entre los sujetos que estudian la lengua vasca solo como asignatura, frente aquellos modelos en que—total o parcialmente—se estudia *en* dicha lengua. En cambio, en lo relativo al español, el modelo lingüístico no resulta ser una variable relevante porque dada la situación socio-lingüística de la Comunidad Autónoma Vasca, todos los sujetos aprenden el español al margen de que sea considerado como asignatura o como lengua vehicular. Ésta es la razón básica por la que entre las dos materias de lengua posibles, hemos optado por la de lengua española.

Los datos del rendimiento de los estudiantes se obtienen como resultado de la aplicación de pruebas curriculares en las que se ha tratado de mantener un cierto equilibrio entre los contenidos conceptuales y los procedimentales. Ya hemos apuntado anteriormente cómo esta evaluación se inserta en el contexto más amplio del proyecto REDES. En una primera etapa, las pruebas fueron preparadas por IDEA⁹. Posteriormente, y antes de su aplicación en la CAV, dichas pruebas fueron revisadas por un conjunto de expertos integrado por profesorado de primaria y secundaria (que analiza la adecuación de los ítems a las líneas curriculares de cada área), por profesorado universitario del área de Medición y Evaluación de nuestro departamento y por personal especializado de la Consejería de Educación del Gobierno Vasco. Como resultado final del trabajo de estos tres grupos de expertos, se envía a IDEA un informe que sirve de base para la elaboración definitiva de cada una de las pruebas. La última fase consiste en la traducción de las mismas a la lengua vasca. En todo este proceso se han seguido las directrices señaladas por la *International Test Commission* (Hambleton, 1996) para la traducción y adaptación de pruebas. En definitiva, las pruebas utilizadas pueden ser consideradas como *Pruebas Estandarizadas de Indicadores de Resultados*.

En la Tabla 6 se muestran las distribuciones de ambas variables para cada uno de los tres cursos en que se aplicaron las pruebas de rendimiento. No hay que olvidar que nuestro trabajo no es un estudio longitudinal, sino que las pruebas han sido administradas a los estudiantes en tiempos pre-definidos (octubre al 1er curso, abril-mayo del siguiente año al 2º y 4º cursos). En este caso, y en función de la distribución de la población por cursos (ver Tabla 1), el número de estudiantes se incrementa con cada curso, mientras que en un estudio longitudinal lo habitual suele ser que el número de efectivos vaya decreciendo con el paso del tiempo.

Con respecto a las variables predictoras, en el nivel de los estudiantes disponemos del género, el curso y el nivel socioeconómico familiar (SESF). En el nivel de los centros escolares se han considerado la titularidad (pública, privada) (TIPO) y el nivel socioeconómico medio de los centros (SESC). En trabajos previos del equipo REDES (Marchesi y Martínez Arias, 2002) se estudiaron otras variables del centro escolar como por ejemplo el tamaño del centro y las valoraciones de padres y alumnos sobre los diferentes procesos de centro, no encontrándose influencias significativas sobre el rendimiento. En el caso de las variables que acabamos de citar, el tamaño del centro sólo influye en la valoración que del centro hacen tanto los padres como los estudiantes, de forma que, en general, la valoración es mejor en los centros escolares pequeños. En cualquier caso, pensamos que una de las limitaciones de este trabajo es el limitado número de variables del nivel 2 que hemos considerado. Como al final apuntamos, en los trabajos actualmente en curso estamos abordando esta cuestión.

⁹ En la parte quinta de la obra citada de Marchesi y Martín (2002) se pueden encontrar descripciones más detalladas de los procedimientos de obtención de datos en las diferentes áreas curriculares. Más en concreto, en los capítulos noveno y décimo se aborda lo relativo a las Matemáticas y la Lengua.

Tabla 6
Distribución de las Puntuaciones de las Pruebas de Rendimiento en Matemáticas y Lengua

Curso	Estadístico	Variables dependientes: Rendimiento en las asignaturas	
		Matemáticas	Lengua
1º ESO (N = 2812)	Media	15,91	16,28
	Desviación típica	5,40	5,49
	Escala (min. – max.)	0–29	3–30
2º ESO (N = 3451)	Media	15,25	15,39
	Desviación típica	5,54	4,46
	Escala (min. – max.)	0–32	0–28
4º ESO (N = 3378)	Media	13,97	17,41
	Desviación típica	5,38	4,66
	Escala (min. – max.)	0–32	0–31

En referencia a la variable relativa a la titularidad o tipo del centro, en el sistema educativo español y en el vasco se garantiza la gratuidad de la enseñanza obligatoria. Con el fin de salvaguardar también el derecho de los padres a la libre elección de centro escolar, a los centros privados se les ofrece la posibilidad de establecer un concierto con el Estado, de forma que la enseñanza obligatoria en dichos centros es también gratuita. En tal caso se habla de centro escolar privado *concertado*, y, en razón de ello, es lógico pensar que la composición social de su alumnado sea más variada que la de un centro privado de pago no acogido a la concertación (estos también existen pero son una minoría). En nuestra muestra, todos los centros escolares con título de privados son centros concertados y en su gran mayoría pertenecen a órdenes religiosas católicas tradicionalmente dedicadas a la enseñanza. En España en el curso 2003-2004, el 69,6% de los estudiantes de enseñanza obligatoria (primaria y secundaria) cursaban sus estudios en centros públicos, el 25,8% en centros privados concertados y sólo el 4,6% en centros privados no concertados.¹⁰ En el País Vasco ya hemos visto cómo la distribución de estudiantes entre centros públicos y privados es distinta (ver Tabla 2). Pero la gran mayoría de estos centros privados son concertados y la proporción de no concertados es muy similar al global de España.

Aclarado esto, en lo que se refiere al tratamiento de las variables predictoras, y tal y como muchos autores preconizan (ver entre otros, Raudenbush, 2002; Willms, 2003), la interpretación de los coeficientes del modelo debe tener un significado claro en relación al marco teórico. Dado que, en nuestro caso, las variables SESF y SESC toman valores entre 1 y 4, careciendo de sentido el valor 0, es preciso centrarlas. Para ello, la covariable SESC no plantea especiales problemas, pues la opción más razonable es hacerlo en su media global. En cambio, la covariable SESF presenta la disyuntiva de usar el centrado en la media global o en la de la escuela. Hemos optado por centrar en la gran media en la medida en que creemos que es más acorde con la percepción general del contexto familiar; es decir, las familias sitúan sus propios contextos más acertadamente en un contexto global que en el del centro educativo. Respecto de las variables de género y titularidad, éstas son variables dicotómicas y están codificadas en $\{0, 1\}$, por lo que la interpretación de sus coeficientes es inmediata, no siendo preciso centrarlas. Este esquema de codificación permite que los coeficientes de los modelos indiquen la diferencia entre hombres y mujeres y entre centros públicos y privados, de forma que coeficientes positivos denotan una mayor puntuación para mujeres y para centros privados respectivamente.

¹⁰Fuente: Ministerio español de Educación. Recuperado de http://www.mec.es/mecd/estadisticas/educativas/dcce/DATOS_Y_CIFRAS_WEB.pdf

Dado que este trabajo presta una atención especial a las variables relativas a los niveles socioeconómicos, procedemos a explicar más en detalle lo relativo a las mismas. Son muchos los autores que coinciden al señalar que el SES es una variable con, al menos, dos dimensiones claramente definidas: el nivel socioeconómico y el nivel sociocultural. Los trabajos de Yang (2003) y Yang y Gustafsson (2004) apuntan a que a nivel del estudiante estas dos dimensiones son claras, mientras que a nivel de los centros educativos la dimensión preponderante es el aspecto económico.

Al igual que en el caso de las pruebas de rendimiento, el cuestionario para recabar datos sobre el nivel socioeconómico fue elaborado por el equipo REDES. A continuación apuntamos las principales características del mismo (una descripción pormenorizada del proceso se puede encontrar en Tiana (2002)). El SES de las familias puede ser operativizado como una medida compuesta de indicadores relativos al nivel de ingresos económicos, el nivel de educación, el nivel cultural y el prestigio y nivel ocupacional. (Willms, 2003; Dutton & Levine, 1989; Mueller & Parcel, 1981). En nuestro caso, la primera decisión adoptada fue la de recoger la información relativa al nivel socioeconómico familiar mediante la aplicación de un cuestionario que habría de ser respondido por los propios estudiantes en el momento de la aplicación en el aula. Dada la edad de los sujetos en los tres cursos, se pensó que era posible recabar información significativa de los propios estudiantes sobre sus familias sin necesidad de dirigirse a los padres. Aunque esto conlleva algunas limitaciones (no se pudo incluir el nivel de ingresos de los padres, por ejemplo), también presenta ventajas, siendo la principal que la tasa de respuesta es obviamente mayor. Para comprobar la viabilidad de la opción escogida, el primer año el cuestionario elaborado fue aplicado tanto a padres como a estudiantes, comprobándose un alto grado de coincidencia entre ambos (Tiana, 2002, p. 74).

A la vista de estudios y evaluaciones anteriores (Álvaro, 1990; y todas las evaluaciones institucionales realizadas por el Ministerio de Educación), al final, el cuestionario contenía ítems sobre las siguientes variables. En primer lugar, *el nivel de estudios del padre y de la madre*; donde el estudiante debe indicar el nivel máximo de estudios cursado por sus padres desde el nivel mínimo (estudios primarios), hasta el máximo (nivel universitario) pasando por estudios secundarios, formación profesional. Segundo, *la situación profesional del padre y de la madre*. Aquí el estudiante debe describir el tipo de actividad profesional que sus padres desempeñan, escogiendo entre cinco opciones: la primera agrupa a empresarios, directivos, profesores, profesiones liberales, médicos, etc.; la segunda se refiere a pequeños empresarios y técnicos medios; la tercera a técnicos auxiliares, empleados de oficina; la cuarta incluye a los obreros cualificados, conductores, etc., y por último, la quinta, obreros no cualificados, empleados domésticos, etc. Estos indicadores están ligados tanto a la dimensión sociocultural como a la socioeconómica. Sin embargo, más directamente vinculados a la dimensión socioeconómica, nos encontramos con *el número de baños del hogar familiar* y *el número de automóviles de la familia* que son indicadores del ámbito socioeconómico que en anteriores estudios demostraron su potencialidad.

Por último nos encontramos con tres variables relativas a la posesión de tecnologías de la información y de bienes culturales. En primer lugar, *la posesión de computadora electrónica en el domicilio familiar*; segundo, *la frecuencia en la compra del periódico*; y, por último, *el número de libros en el domicilio familiar*. Aunque sea *grosso modo* (indicando el número aproximado mediante la selección entre intervalos), este último descriptor proporcionado por los estudiantes es un buen indicador del nivel sociocultural habida cuenta de que en la evaluación de la educación primaria realizada por el Instituto de Evaluación del Ministerio de Educación, los resultados fueron que el 37 por ciento de los hogares poseen entre 0-50 libros; el 29 por ciento entre 50-100; y el 32 por ciento más de un centenar libros (INCE.MEC, 2000). El empleo de los materiales de lectura como indicador sociocultural es un recurso habitual. Por ejemplo, Lubienski (2006) opta por una aproximación más cualitativa e incluye el tipo de material de lectura (periódicos, revistas, enciclopedias, etc.) disponible en los domicilios.

La puntuación base de cada sujeto en el nivel socioeconómico familiar se obtiene de aplicar a la respuesta a cada ítem del cuestionario la ponderación establecida para la variable en cuestión. A su vez, estas ponderaciones son el resultado de los pesos factoriales obtenidos por cada variable en el análisis factorial que el equipo REDES realizó en el momento inicial del diseño del instrumento (Tiana, 2002). En el extremo inferior derecho de la Tabla 7 (totales fila y columna), podemos ver que este índice oscila para toda la muestra entre un mínimo de 57,8 y un máximo de 146,5.

Como en su momento anticipamos, además de disponer de la puntuación original de cada individuo, y con el fin de facilitar la interpretación de los resultados, el nivel socioeconómico fue categorizado en cuatro valores empleando el criterio de los cuartiles, y es a esta nueva variable recodificada a la que denominamos SESF. Este procedimiento ha sido usado por autores como D'Angiulli et al. (2004) y Howley & Howley (2004). En la Tabla 7 se muestran, para los tres cursos, los límites superiores e inferiores de los cuatro niveles del nivel socioeconómico familiar resultantes de la aplicación de dicho criterio de categorización.

Tabla 7

Escalas (Mínimo y Máximo) de los Valores Originales de SES Familiar por Curso y Categorías

SES Familiar	1º ESO		2º ESO		4º ESO		Todos	
	Mínimo	Máximo	Mínimo	Máximo	Mínimo	Máximo	Mínimo	Máximo
Bajo	65,0	91,2	62,5	88,2	57,8	89,6	57,8	91,2
Medio-bajo	91,2	101,0	88,3	98,6	89,7	98,7	88,3	101,0
Medio-alto	101,0	111,8	98,6	110,0	98,7	109,0	98,6	111,8
Alto	111,8	146,5	110,1	134,8	109,1	133,5	109,1	146,5
Total	65,0	146,5	62,5	134,8	57,8	133,5	57,8	146,5

Con respecto al nivel socioeconómico de la escuela, aquí nos encontramos con dos opciones: o se opera exclusivamente con el índice promedio de los centros a partir de los índices obtenidos por los individuos en el nivel socioeconómico familiar; o, al mismo se incorporan además otras características de la escuela. Por ejemplo, en el ya citado trabajo de Lubienski (2006) se incorpora al SES promedio del centro el porcentaje de estudiantes solicitantes de becas y de ayudas económicas para el almuerzo. Como algunos de estos indicadores no son válidos para nuestro sistema educativo hemos optado por la primera opción y hemos construido un índice promedio de los centros. Este criterio de calcular el nivel socioeconómico medio de los centros educativos usando únicamente el promedio de los valores del nivel familiar de los estudiantes pertenecientes a dichos centros ha sido empleado, por ejemplo, en los trabajos de Willms y Raudenbush (1989); De Fraine, Van Damme & Onghena (2002); Rumberger y Palardy (2005) y Willms (2006). En nuestro entorno, ha sido empleado también en el informe PISA 2003 en la CAV, donde se afirma que así obtenido es un “predictor poderoso del rendimiento” (ISEI, 2005, p. 94).

En consecuencia, una vez promediados para cada centro escolar los índices de los estudiantes, nos encontramos con una puntuación que se distribuye con un mínimo de 85, un máximo de 124 y una media aritmética de 100. Por último, al igual que con el SESF, el nivel socioeconómico de los centros (que hemos denominado SESC) se obtiene de la categorización de esta puntuación en 4 intervalos. En la Tabla 8 se muestra el criterio de categorización empleado así como la distribución por cursos. Y la Tabla 9 presenta la distribución de los estudiantes en estas cuatro categorías por cursos y según el tipo de los centros escolares. Como puede observarse, no hay ningún centro educativo de titularidad pública que alcance la categoría promedio de *alto*.

Tabla 8
Categorización del SES del Centro Educativo por Cursos

SES Centro		1º ESO (%)	2º ESO (%)	4º ESO (%)
1 Bajo	$X \leq 93$	540 (19,2)	673 (19,5)	705 (20,9)
2 Medio-bajo	$93 < X \leq 100$	989 (35,2)	1262 (36,6)	1154 (34,2)
3 Medio-alto	$100 < X \leq 108$	750 (26,7)	825 (23,9)	806 (23,9)
4 Alto	$X > 108$	533 (19,0)	691 (20,0)	713 (21,1)

Tabla 9
Distribución según SES del Centro Educativo por Cursos y Titularidad

SES Centro	1º ESO			2º ESO			4º ESO		
	Público	Privado	Total	Público	Privado	Total	Público	Privado	Total
Bajo	335	205	540	446	227	673	489	216	705
Medio-bajo	325	664	989	416	846	1262	384	770	1154
Medio-alto	273	477	750	286	539	825	269	537	806
Alto	0	533	533	0	691	691	0	713	713
Total	933	1879	2812	1148	2303	3451	1142	2236	3378

Modelización de los Efectos Contextuales

El Modelo de Doble Riesgo

Tal y como señalamos al formular los objetivos de este trabajo, en la literatura sobre la influencia del SES –tanto familiar como de centro escolar- en el rendimiento académico (por ejemplo, Willms, 2003) se suele hablar de la hipótesis de doble riesgo. Ésta se basa en el supuesto de que los estudiantes de nivel socioeconómico familiar bajo están en una situación de doble riesgo porque, a la desventaja que supone su bajo status familiar, se añade el hecho de que esta situación desfavorable empeora cuando aquellos estudian en centros educativos cuyo nivel socioeconómico promedio es igualmente bajo. Es decir, esta hipótesis plantea que los sujetos de familias de bajo nivel socioeconómico obtendrán un rendimiento académico aún peor cuando asistan a centros escolares cuyo nivel socioeconómico promedio sea también bajo. El modelo subyacente a esta hipótesis es el siguiente:

Nivel 1:

$$Y_{ij} = \beta_{0j} + \beta_{1j} \text{SES Familiar}_{ij} + r_{ij}$$

Nivel 2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} \text{SES Centro}_j + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + u_{1j}$$

En el que Y_{ij} representa las puntuaciones en una asignatura (Matemáticas o Lengua) del estudiante “i” perteneciente al centro “j”. En cuanto a los parámetros de la parte fija, γ_{00} es la puntuación media en la asignatura de los estudiantes de contexto familiar y contexto del centro educativo medios (recordemos que ambos contextos han sido centrados en la media global). Por su parte, γ_{01} representa el efecto contextual del centro escolar y γ_{10} el efecto contextual del contexto familiar, de manera análoga a los coeficientes de regresión. En lo que se refiere a la parte aleatoria del modelo, hay que señalar que u_{0j} evalúa la desviación respecto de la puntuación media en la asignatura (γ_{00}) de cada centro, y u_{1j} la desviación de cada centro respecto de la influencia media del SES Familiar (γ_{10}).

El Modelo de Triple Riesgo

Como al inicio apuntamos, se habla de hipótesis de *triple riesgo* porque el supuesto básico es que, al doble riesgo que acabamos de explicar en la sección anterior, se añade el hecho de que el efecto contextual es más fuerte para estudiantes de bajo nivel socioeconómico familiar. El supuesto subyacente es que la posible interacción entre el SES familiar y el SES centro educativo podría tender a favorecer a los estudiantes de mayor SES familiar ya que, situados en un contexto desfavorable (centro educativo de bajo nivel socioeconómico), su mejor situación familiar les permitiría poder disponer de otros recursos adicionales que paliasen o mejorasen su situación. Por el contrario, los estudiantes de bajo nivel socioeconómico familiar no pueden disponer de dichos recursos adicionales. No es sólo que el efecto contextual varíe dependiendo de la escala del nivel socioeconómico familiar, sino que esta variación opera de forma que el efecto es más intenso en sujetos de nivel socioeconómico familiar bajo. Ahora tenemos un modelo en el que se incorpora también el SES centro educativo como predictor.

Nivel 1:

$$Y_{ij} = \beta_{0j} + \beta_{1j} \text{SES Familiar}_{ij} + r_{ij}$$

Nivel 2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} \text{SES Centro}_j + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} \text{SES Centro}_j + u_{1j}$$

La representación compacta a través del modelo mixto, quizá más intuitiva, es la siguiente:

$$Y_{ij} = \gamma_{00} + \gamma_{01} \text{SES Centro}_j + \gamma_{10} \text{SES Familiar}_{ij} + \gamma_{11} \text{SES Centro}_j \times \text{SES Familiar}_{ij} + (u_{0j} + u_{1j} \text{SES Familiar}_{ij} + r_{ij}).$$

Aquí se puede observar, como nuevo efecto, la mencionada interacción entre niveles, es decir, entre el SES familiar y el SES del centro educativo, representada por el parámetro γ_{11} .

Modelo con Inclusión de Variables Predictoras en el Nivel 1 y en el Nivel 2

El objeto de estos análisis es el de examinar los efectos contextuales controlando las variables predictoras de que se dispone entre los datos. Así, se incluirá en el nivel 1 la variable dicotómica *sexo* y en el nivel 2, la también variable dicotómica *titularidad* del centro educativo (tipo: público, privado); ambas variables bajo código binario. El modelo completo se representa con la siguiente ecuación:

Nivel 1:

$$Y_{ij} = \beta_{0j} + \beta_{1j} \text{SES Familiar}_{ij} + \beta_{2j} \text{Sexo}_{ij} + r_{ij}$$

Nivel 2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} \text{SES Centro}_j + \gamma_{02} \text{Tipo}_j + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} \text{SES Centro}_j + \gamma_{12} \text{Tipo}_j + u_{1j}$$

$$\beta_{2j} = \gamma_{20} + \gamma_{21} \text{SES Centro}_j + \gamma_{22} \text{Tipo}_j + u_{2j}$$

Sabemos previamente que en el presente modelo bastantes de los efectos especificados no van a ser significativos, pero también es verdadero que conviene partir del modelo completo hasta llegar a uno definitivo en el que sólo tengan presencia los efectos significativos¹¹. Los principales parámetros se interpretan como sigue:

γ_{00} : puntuación media de los chicos de centros públicos de SES familiar y SES centro educativo medios; para las chicas se prevé un incremento de γ_{20} y para los centros privados de γ_{02}

γ_{01} : influencia del SES centro educativo; para las chicas se incrementa en γ_{21}

¹¹ Hay que subrayar que con el fin de incluir de forma exhaustiva los posibles efectos derivados de las variables disponibles, en un principio se introdujeron las interacciones SES familiar \times sexo en el nivel 1 y SES centro educativo \times tipo en el nivel 2. Sin embargo, en general resultaron no ser significativos.

γ_{10} : influencia del SES familiar en centros públicos; en los centros privados se prevé un incremento de γ_{12}

γ_{11} : interacción entre SES centro educativo y SES familiar.

γ_{22} : incremento de la puntuación media para las chicas de centros privados (no confundir con $\gamma_{20} + \gamma_{02}$, pues se trata de una interacción).

Resultados

Los Modelos de Doble y Triple Riesgo

Una vez estimados los parámetros del modelo nulo y tras constatar que hay variación en el rendimiento en ambas asignaturas, tanto entre los estudiantes en su propio centro educativo como entre los centros mismos, pasamos a crear los modelos de los posibles efectos contextuales en función del SES, tanto a nivel familiar como de los centros. En todos estos análisis se ha empleado el programa HLM (Raudenbush, Bryk, Cheong y Congdon, 2005). La Tabla 10 muestra los resultados manteniendo sólo los efectos significativos para cada una de las asignaturas y cada nivel educativo.

A la vista de estos resultados, constatamos que se dan los dos efectos contextuales, el familiar y el escolar. Además, utilizando el criterio χ^2 , basado en la diferencia entre las razones de verosimilitud entre los modelos a contrastar, todos los modelos de doble riesgo ajustan mejor que los correspondientes modelos nulos. Dado que, en mayor o menor grado, se observan diferencias en la influencia de los efectos contextuales de la familia y del centro escolar, se han realizado las correspondientes pruebas de hipótesis entre las estimaciones de los parámetros de tales efectos. Para ello se ha utilizado, dentro de las pruebas de hipótesis multivariantes para efectos fijos, la prueba de hipótesis lineal general para pruebas de múltiples parámetros, especificando la siguiente hipótesis nula:

$$H_0 : \gamma_{\text{SES Centro}} - \gamma_{\text{SES Familiar}} = 0$$

Los resultados de contrastar esta hipótesis no permiten concluir de forma global que uno de los contextos tenga un efecto significativamente mayor que el otro. En el 1º curso, por ejemplo, el efecto de SES centro educativo sí es significativamente mayor que el de SES familiar, pero sólo en la asignatura de Lengua. Por el contrario, en el 2º curso la diferencia es significativa sólo en Matemáticas (siendo también mayor el efecto del SES de centro) y en el 4º curso las diferencias no resultan significativas. En cualquier caso, sólo la intersección tiene un componente aleatorio significativo; es decir, la influencia de SES familiar sobre el rendimiento de cada asignatura es análoga en todos los centros educativos. Se observa también que el incremento de un punto en el SES familiar origina un incremento de aproximadamente medio punto en los distintos rendimientos. Por su parte, el incremento de un punto en SES centro educativo, se traduce en un incremento de aproximadamente un punto en el rendimiento de las dos asignaturas. En lo que se refiere a los resultados de los estudiantes de 2º curso, y sólo en el rendimiento en Matemáticas, el efecto de SES familiar varía entre los centros. En resumen, la influencia de SES Familiar se eleva a aproximadamente a más de 0,70 puntos y la de SES Centro, más irregular, oscila entre más de un punto en Matemáticas y algo más de medio punto en Lengua. Además sólo en Matemáticas la influencia de SES Familiar varía significativamente entre los centros. Como tendencia, las diferencias entre los estudiantes de un mismo centro disminuyen con respecto a los estudiantes de 1º.

En los estudiantes de 4º curso se dan ambos efectos. El incremento de un punto en SES familiar origina un aumento de más de 0,60 en el rendimiento en Matemáticas; pero en Lengua no llega a medio punto. Respecto al SES centro educativo, su influencia oscila entre casi 0,75.

Tabla 10

Hipótesis de Doble Riesgo: 1º, 2º y 4º del Curso 1999-2000 Analizados por Separado

Curso	Matemáticas		Lengua	
	Parámetro	Estimación	Parámetro	Estimación
1° ESO				
Parte Fija				
	Constante	15,74 (0,23)**	Constante	16,07 (0,23)**
	SES Familiar	0,47 (0,10)**	SES Familiar	0,48 (0,10)**
	SES Centro	0,97 (0,27)**	SES Centro	1,05 (0,29)**
Parte Aleatoria				
	Nivel 2		Nivel 2	
	Constante	2,43 (0,56)**	Constante	2,27 (0,54)**
	Nivel 1	23,92	Nivel 1	25,27
	Razón de verosimilitud	17004,8 ₅	Razón de verosimilitud	17154,2 ₅
2° ESO				
Parte Fija				
	Constante	15,08 (0,25)**	Constante	15,28 (0,16)**
	SES Familiar	0,80 (0,10)**	SES Familiar	0,66 (0,06)**
	SES Centro	1,12 (0,24)**	SES Centro	0,56 (0,18)**
Parte Aleatoria				
	Nivel 2		Nivel 2	
	Constante	3,05 (0,67)**	Constante	1,05 (0,25)**
	SES Familiar	0,19 (0,10)**	-----	
	Nivel 1	21,04	Nivel 1	15,25
	Razón de verosimilitud	20113,1 ₇	Razón de verosimilitud	18966,0 ₅
4° ESO				
Parte Fija				
	Constante	13,77 (0,31)**	Constante	17,28 (0,16)**
	SES Familiar	0,64 (0,07)**	SES Familiar	0,46 (0,08)**
	SES Centro	1,09 (0,31)**	SES Centro	0,75 (0,16)**
Parte Aleatoria				
	Nivel 2		Nivel 2	
	Constante	4,73 (0,98)**	Constante	1,13 (0,28)**
	Nivel 1	18,94	Nivel 1	17,41
	Razón de verosimilitud	19344,5 ₅	Razón de verosimilitud	19003,1 ₅

Nota. El valor entre paréntesis representa el error típico de estimación del parámetro.

Las variables predictoras SES familiar y SES centro educativo están centradas en la gran media./

El subíndice de la Razón de verosimilitud representa los grados de libertad.

* $p < 0,05$ / ** $p < 0,01$

Los Modelos Definitivos

Tabla 11

1°, 2° y 4° del Curso 1999-2000 Analizados por Separado

Curso	Matemáticas		Lengua	
	Parámetro	Estimación	Parámetro	Estimación

		1º ESO			
Parte Fija		Constante	14,68 (0,39)**	Constante	15,35 (0,26)**
		SES Familiar	0,47 (0,10)**	SES Familiar	0,48 (0,10)**
		SES Centro	0,75 (0,24)**	SES Centro	1,05 (0,29)**
		Tipo	1,59 (0,46)**	Sexo	1,60 (0,21)**
Parte Aleatoria		Nivel 2		Constante	2,24 (0,53)**
		Constante	2,79 (0,71)**	Sexo	-
		Sexo	0,80 (0,51)*	Nivel 1	24,65
		Nivel 1	23,72	Razón de verosimilitud	17084,8 ₆
		Razón de verosimilitud	16985,9 ₈		
		2º ESO			
Parte Fija		Constante	14,95 (0,27)**	Constante	14,84 (0,18)**
		SES Familiar	1,12 (0,18)**	SES Familiar	0,91 (0,10)**
		SES Centro	1,06 (0,24)**	SES Centro	0,58 (0,18)**
		Sexo	0,36 (0,17)*	Sexo	1,05 (0,16)**
		SES Familiar x Tipo	-0,49 (0,21)*	SES Familiar x Tipo	-0,37 (0,13)**
Parte Aleatoria		Nivel 2		Constante	1,23 (0,33)**
		Constante	3,09 (0,67)**	Sexo	0,44 (0,27)*
		SES Familiar	0,17 (0,20)**	Nivel 1	14,85
		Nivel 1	21,01	Razón de verosimilitud	18897,7 ₉
		Razón de verosimilitud	20102,5 ₉		
		4º ESO			
Parte Fija		Constante	14,09 (0,31)**	Constante	16,23 (0,29)**
		SES Familiar	0,61 (0,07)**	SES Familiar	0,48 (0,08)**
		SES Centro	1,07 (0,31)**	SES Centro	0,67 (0,16)**
		Sexo	-0,69 (0,19)**	Sexo	1,10 (0,15)**
				Tipo	0,77 (0,33)*
Parte Aleatoria		Nivel 2		Constante	1,01 (0,25)**
		Constante	4,67 (1,02)**	Sexo	-
		Sexo	0,71 (0,39)**	Nivel 1	17,12
		Nivel 1	18,65	Razón de verosimilitud	18943,4 ₇
		Razón de verosimilitud	19317,2 ₈		

Nota. El valor entre paréntesis representa el error típico de estimación del parámetro.

Las variables predictoras SES Familiar y SES Centro están centradas en la gran media.

El subíndice de la Razón de verosimilitud representa los grados de libertad.

* $p < 0,05$ / ** $p < 0,01$

Debido a que los modelos resultan más complejos, se comentarán los resultados en apartados separados para las dos asignaturas, aunque en la Tabla 11 se presenten juntos.

Basándose en el criterio de χ^2 , todos estos modelos ajustan mejor que los correspondientes modelos de doble riesgo. Al igual que en el modelo relativo a la hipótesis de doble riesgo, también aquí se han realizado las correspondientes pruebas de hipótesis entre las estimaciones de los parámetros de los efectos contextuales de familia y centro educativo. Y en esta ocasión los resultados son más claros, pues prácticamente en ningún caso se dan diferencias

significativas entre ambos efectos. Debido a que se trata de 6 modelos (2 asignaturas y 3 cursos), para simplificar, la presentación de los resultados se estructura en dos subapartados, uno para cada asignatura.

Los Modelos para Matemáticas

Como ilustración de la interpretación de los parámetros, se incidirá en los resultados de primer curso. Los otros se representan esquemáticamente. En 1° de la ESO la constante 14,68 representa la línea base y corresponde a la puntuación media en Matemáticas de los estudiantes de nivel socioeconómico familiar y de centro escolar medios y que asisten a centros públicos. Bajo las mismas condiciones de SES, los estudiantes de centros privados obtienen 1,586 puntos más, es decir, 16,266 puntos. El incremento de un punto en SES familiar origina un incremento de 0,474 en la puntuación de Matemáticas. Así, la diferencia entre los estudiantes de menor y mayor SES familiar es de 1,422 puntos. En cuanto al SES centro escolar, un incremento de un punto da lugar a un incremento de 0,754 puntos en Matemáticas, por lo que la diferencia en matemáticas entre los dos extremos del SES centro escolar es de 2,262 puntos. Observando la parte aleatoria, en las chicas se da mayor dispersión que en los chicos.

En 2° curso la constante, cuyo valor es 14,947, corresponde a los chicos de SES familiar y de centro escolar medios; para las chicas este valor aumenta en 0,358; es decir, 15,305. Debido a la interacción entre *tipo* de centro y SES familiar, éste influye de distinta manera en los centros públicos y en los privados. En los públicos, incrementar el SES familiar en un punto supone que la puntuación en Matemáticas aumenta en 1,123 puntos; mientras que en los privados, el aumento es sólo de 0,633 (1,123 – 0,490). En cuanto al SES centro escolar, incrementarlo en un punto produce un incremento medio de 1,057 en Matemáticas. En consecuencia, la diferencia entre los estudiantes de SES familiar más bajo y más alto es de 3,369 en los centros públicos y de 1,9 en los privados. En cuanto a los valores extremos del SES centro escolar, las correspondientes puntuaciones en la asignatura se diferencian en 3,171 puntos. Además, a la vista de la parte aleatoria del modelo, sólo la influencia de SES familiar es variable entre los centros educativos.

En el 4° curso la puntuación media es de 14,088 para los chicos con valores medios tanto en el SES familiar como en el del centro escolar. Las chicas ven disminuida su puntuación hasta 13,40. La influencia del SES familiar se traduce en un aumento de 0,606 puntos y la del SES centro educativo en 1,073. La comparación entre los valores extremos de ambos SES representa una diferencia en Matemáticas de 1,818 y de 3,219 puntos respectivamente. La parte aleatoria del modelo indica que las chicas tienen puntuaciones más dispersas. Observando particularmente la parte aleatoria, en 1° curso se da mayor dispersión en las chicas que en los chicos; en 2° la influencia de SES familiar es variable entre los centros escolares y en 4° otra vez las chicas tienen puntuaciones más dispersas. Estas observaciones de la parte aleatoria las representamos esquemáticamente por medio de la Figura 1.

Parámetro	Línea de base	SES Familiar	SES Centro
Descripción	Es la puntuación de los estudiantes cuyo SES Familiar y del Centro son medios.	Refleja el incremento en Matemáticas producido por el aumento de un punto en SES Familiar.	Refleja el incremento en Matemáticas producido por el aumento de un punto en SES Centro.

1º ESO			
Estimación	14,680 público 16,266 privado	0,474	0,754
Máxima diferencia en Matemáticas para los valores mínimo y máximo de SES.		1,422	2,262
2º ESO			
Estimación	14,947 chicos 15,305 chicas	1,123 público 0,633 privado	1,057
Máxima diferencia en Matemáticas para los valores mínimo y máximo de SES.		3,369 público 1,900 privado	3,171
4º ESO			
Estimación	14,088 chicos 13,400 chicas	0,606	1,073
Máxima diferencia en Matemáticas para los valores mínimo y máximo de SES.		1,818	3,219

Figura 1. Representación esquemática del modelo para matemáticas

Los Modelos para Lengua

Siguiendo el esquema presentado en Matemáticas, se desarrollarán los resultados curso por curso. En 1º de la ESO la línea base corresponde a los chicos, de manera que los que tienen niveles medios en SES familiar y SES centro educativo obtienen 15,347 puntos en Lengua; las chicas, por su parte, aumentan hasta 16,950 puntos. La influencia del nivel socioeconómico consiste en que el incremento de un punto en el SES familiar supone un incremento en el rendimiento académico de 0,477 puntos en Lengua, mientras que un incremento análogo en el SES centro educativo genera un incremento en rendimiento de 1,046 en Lengua. Esto implica que la máxima diferencia de puntuación en Lengua entre los valores mínimo y máximo de SES familiar y SES centro educativo es de 1,431 y 3,138 puntos respectivamente.

En segundo curso de la ESO las diferencias vuelven a darse entre chicos y chicas. Así, para niveles medios de ambos SESs, los chicos obtienen una media de 14,845 puntos en Lengua, mientras que las chicas llegan hasta 15,895 puntos. Por otra parte, la influencia de SES familiar no es homogénea, pues a los centros públicos corresponde un incremento en el rendimiento académico de 0,913 puntos en Lengua por cada punto que aumenta el SES familiar, mientras que en los privados este valor se reduce hasta 0,539 puntos. De esta manera la máxima diferencia en Lengua para los SES familiares extremos es de 2,739 en los centros educativos públicos y de 1,617 en los privados. En cuanto al SES centro educativo, por cada punto de aumento en el SESE, hay 0,583 puntos de aumento en el rendimiento en Lengua. Esto se traduce en una diferencia máxima de 2,739 para los centros educativos públicos y de 1,617 para los privados al comparar los centros de SES mínimo y máximo.

En 4º curso de ESO las interacciones son algo más complejas. Se puede observar que hay cuatro grupos. La línea base, SES familiar y de centro educativo medios, corresponde a los

chicos de centros públicos, obteniendo éstos 16,231 puntos en Lengua. Los chicos de centros privados de niveles medios en ambos SESs obtienen 17 puntos. Bajo las mismas condiciones de nivel socioeconómico, las chicas de los centros públicos obtienen 17,336 puntos y las de los centros privados 18,105. Un punto de incremento en el SES familiar genera un aumento de 0,483 puntos en el rendimiento académico en Lengua, mientras que un punto de incremento en el SES centro educativo, genera 0,669 puntos de aumento. Los efectos máximos de SES familiar y SES centro educativo—es decir, la diferencia para el SES mínimo y máximo—son de 1,449 y 2 puntos respectivamente. El único efecto aleatorio de las variables predictoras se da en 2º curso y corresponde a la variable *sexo*, volviendo a ser más dispersas las puntuaciones de las chicas en la asignatura Lengua. Esto lo representamos esquemáticamente en la Figura 2.

Parámetro	Línea de base	SES Familiar	SES Centro
Descripción	Es la puntuación de los estudiantes cuyo SES Familiar y del Centro son medios.	Refleja el incremento en Lengua producido por el aumento de un punto en SES Familiar.	Refleja el incremento en Lengua producido por el aumento de un punto en SES Centro.
1º ESO			
Estimación	15,347 chicos 16,950 chicas	0,477	1,046
Máxima diferencia en Lengua para los valores mínimo y máximo de SES		1,431	3,138
2º ESO			
Estimación	14,845 chicos 15,895 chicas	0,913 público 0,539 privado	0,583
Máxima diferencia en Lengua para los valores mínimo y máximo de SES		2,739 público 1,617 privado	1,749
4º ESO			
Estimación	16,231 chicos-público 17,000 chicos-privado 17,336 chicas-público 18,105 chicas-privado	0,483	0,669
Máxima diferencia en Lengua para los valores mínimo y máximo de SES		1,449	2,007

Figura 2. Representación esquemática del modelo para lengua

Como resumen final de los resultados obtenidos, apuntamos que los efectos, tanto del nivel socioeconómico familiar como del centro educativo son significativos. En consecuencia: *se comprueba la existencia de efecto contextual y el cumplimiento de la hipótesis de doble riesgo*. En cambio, el efecto de interacción entre SES familiar y SES centro educativo no resulta significativo, es decir: *no se ha podido probar que se cumpla la hipótesis de triple riesgo*. Si comparamos ambos efectos, el efecto del nivel socioeconómico promedio del centro educativo es significativamente mayor que el SES familiar en primer curso en la asignatura de Lengua; sin embargo, no hay suficiente evidencia de

que en 2º y 4º cursos se den estas diferencias entre ambos efectos en ninguna de las dos asignaturas. Con respecto a la variable *sexo*, las chicas obtienen mejores puntuaciones salvo en Matemáticas de 1º curso (las puntuaciones son sólo más dispersas) y en Matemáticas de 4º curso (las puntuaciones son menores). Y con respecto a la *titularidad*, nos encontramos con que en 2º curso en los centros privados el efecto del nivel socioeconómico familiar es menor que en los públicos. Las únicas diferencias encontradas entre centros educativos públicos y privados fueron a favor de éstos últimos en Matemáticas de 1º curso y Lengua de 4º. Por tanto, una vez que los modelos jerárquicos lineales nos han permitido comprobar la hipótesis de doble riesgo, así como el efecto contextual asociado, y los efectos del resto de variables predictoras, en el siguiente apartado vamos a tratar de analizar los mismos de una forma más matizada.

Estudio Conjunto del SESF y SESC

Modelización Mediante Segmentación

Acabamos de ver cómo uno de los principales resultados de los modelos jerárquicos que hemos formulado es establecer la importancia que el nivel socioeconómico del centro educativo tiene como variable predictora del rendimiento escolar; es decir, la constatación del efecto contextual en el rendimiento académico de los estudiantes. Esto en sí mismo no constituye ninguna novedad pues, como vimos al inicio de este artículo en la revisión de la literatura, este hecho ha sido comprobado en muchas investigaciones. Pero, aunque el efecto de interacción entre SESF y SESC no ha resultado significativo (la hipótesis de triple riesgo no ha sido comprobada), ello no es óbice para que la acción conjunta de estas dos variables pueda ser examinada con mayor detenimiento. Y para ello, en primer lugar hemos procedido a generar una nueva variable denominada “Combina” (ver Tabla 12), cuyos 16 valores son el resultado del producto cruzado de ambas variables categorizadas en cuatro valores, de forma que el primer dígito representa el valor del SES del centro educativo y el segundo dígito el del SES familiar.

Tabla 12

Creación de la Variable “Combina” (SES Centro Educativo \times SES Familiar)

		SES Familiar			
		1 (Bajo)	2 (Medio-bajo)	3 (Medio-alto)	4 (Alto)
SESC	1 (Bajo)	11	12	13	14
	2 (Medio-bajo)	21	22	23	24
	3 (Medio-alto)	31	32	33	34
	4 (Alto)	41	42	43	44

Para analizar el comportamiento de esta nueva variable en el contexto general de los modelos formulados procedemos a emplear técnicas estadísticas de segmentación, también conocidas como árboles de decisión o clasificación. Se trata de un conjunto de técnicas que permiten definir y validar modelos de forma que se pueda determinar qué variables (predictoras) inciden o explican los cambios de una variable dependiente. Son técnicas estadísticas explicativas de la familia de la regresión o el análisis discriminante, pero tienen la ventaja de que tanto la variable criterio como las predictoras pueden ser de cualquier tipo (cuantitativas o cualitativas). Este tipo de técnicas permiten abordar problemas y cuestiones como la propia segmentación de

poblaciones, la validación de modelos predictivos, la reducción de la dimensionalidad o la identificación de la interacción (SPSS, 2001).

La Figura 3 muestra el primer nivel del árbol resultante de la aplicación del método CART (Kim, 1991). La variable dependiente es el rendimiento académico de los estudiantes de primer curso en la asignatura Lengua y las variables predictoras son el tipo de centro, el género y la nueva variable combinada de ambos SES. Para generar el árbol, el algoritmo ha buscado qué dicotomización de todas las variables predictoras origina una máxima diferencia con respecto a la variable dependiente.

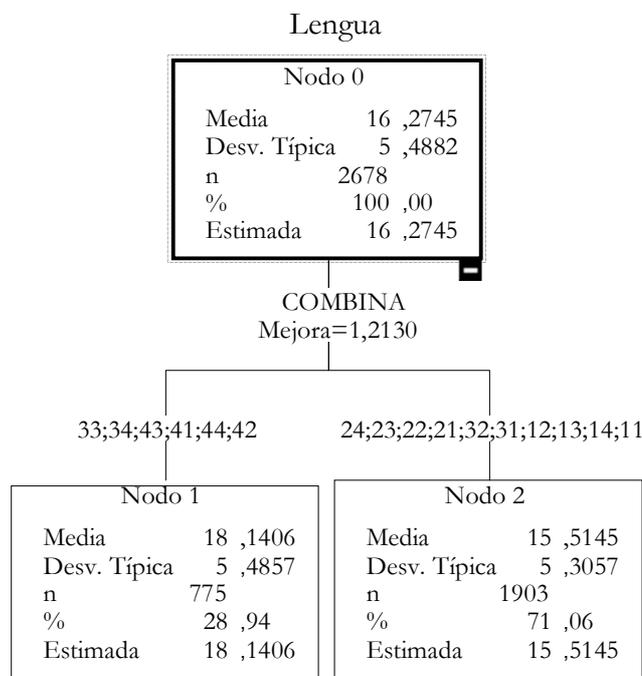


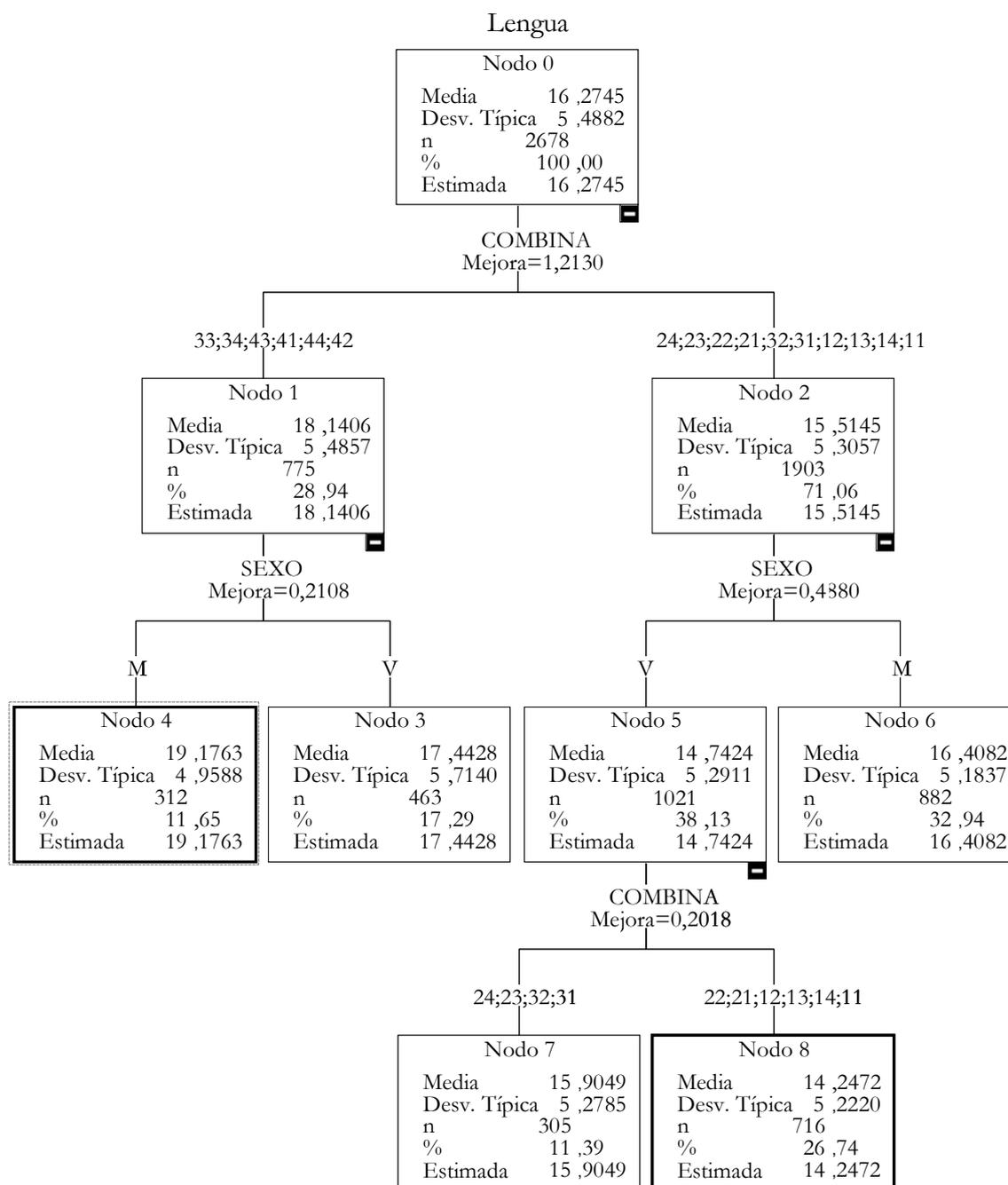
Figura 3. Segmentación inicial (Lengua 1º curso). Árbol de primer nivel.

En este caso, la solución alcanzada segmenta en dos subgrupos (nodos) a los 2678 estudiantes de primer curso con respecto a su rendimiento en Lengua, de forma que se distingue entre el subgrupo superior (media = 18,14) y el inferior (media = 15,51). Si observamos la composición de ambos nodos con respecto a los 16 valores de la variable combinada, vemos cómo operan los efectos de ambas medidas del SES. Si el peso más importante correspondiese al nivel socioeconómico familiar, la dicotomización tendería a establecerse distinguiendo entre sujetos de status socioeconómico familiar alto frente a los de nivel socioeconómico bajo. Si fuese el efecto contextual del centro educativo el más relevante, la segmentación usaría como criterio el nivel medio del centro escolar, agrupando en el nodo superior a las escuelas de status socioeconómico alto y medio-alto y en el nodo inferior a las escuelas de nivel socioeconómico bajo y medio-bajo.

Como vemos, lo que ocurre es que el subgrupo superior está constituido por 6 clases: las 4 correspondientes a los centros de nivel socioeconómico alto, independientemente del nivel socioeconómico familiar (41, 42, 43 y 44), y las 2 restantes que corresponden a los centros educativos de nivel medio-alto con SES familiar medio-alto y alto (33 y 34). El examen de las 10 clases del subgrupo inferior desvela de manera análoga la caracterización del mismo.

Una vez obtenido este primer nivel de segmentación, estamos en condiciones de generar el modelo final. Para ello, el algoritmo procede a ir segmentando cada nodo buscando la mejor partición dicotómica; es decir, la partición que arroje la mayor diferencia con respecto a la variable dependiente, en este caso el rendimiento en Lengua. El proceso de segmentación finaliza cuando se satisface algún criterio de parada (tamaño del nodo, nivel alfa de la diferencia, etc.) y

cuando se eliminan aquellos nodos y ramas que no aportan una mejora significativa de la capacidad explicativa del modelo. En este caso, la solución es la que muestra la Figura 4.



Resultados de la Segmentación

Aquí vemos cómo opera cada una de las variables predictoras segmentando los diferentes subgrupos a los diferentes niveles de profundidad del árbol. Así, el segundo nivel es generado en ambos casos (en ambos nodos) segmentando en función del género de los estudiantes. Y para obtener el tercer, y último nivel, el nodo 5 es segmentado de nuevo en función de la variable que combina ambos niveles del SES. El criterio de segmentación es el siguiente: en el nodo 7 nos

encontramos con los sujetos de nivel socioeconómico familiar bajo y medio-bajo que acuden a centros de nivel socioeconómico medio alto (subgrupos 31 y 32) y a los sujetos de nivel socioeconómico familiar alto y medio-alto que acuden a centros educativos de nivel socioeconómico medio-bajo (subgrupos 23 y 24). Por último, en el nodo 8 se encuentran los 716 sujetos que obtienen la media más baja en el rendimiento académico en Lengua (14,25). Su caracterización es clara, se trata de los chicos que pertenecen a los seis subgrupos más desfavorecidos en términos socioeconómicos: los cuatro que acuden a centros educativos de nivel socioeconómico bajo, independientemente de su nivel familiar (11, 12, 13 y 14), y los dos que asisten a centros educativos de nivel socioeconómico medio-bajo y cuyo nivel socioeconómico familiar es bajo o medio-bajo (21 y 22). En contraposición, el nodo 4 obtiene la media aritmética más alta (19,33), siendo un nodo terminal compuesto por las 312 chicas cuyos valores de SES familiar y de SES de centro educativo fueron descritos en el primer nivel de segmentación.

El modelo obtenido mediante técnicas de segmentación aporta resultados similares a los modelos jerárquicos en lo que a la influencia de las variables predictoras se refiere. Para esta asignatura y para primer curso, operan el sexo y la variable combinada, lo que es básicamente coincidente con lo que acabamos de apuntar en las conclusiones de los modelos jerárquicos. Y en el resto de los grupos y asignaturas ocurre lo mismo. Por ejemplo, en el curso 4º para esta misma asignatura, las variables de segmentación son en primer lugar la combinada, y en segundo el sexo y la titularidad de los centros.

Es obviamente esperable que ambos enfoques (HLM y segmentación) ofrezcan soluciones convergentes y similares por lo que no vamos a insistir en ello. En cambio, son otras las cuestiones que queremos reseñar aquí. En primer lugar, mediante la solución gráfica aportada podemos fácilmente caracterizar los diferentes subgrupos e ir evaluando paso a paso las soluciones a los diferentes niveles. Segundo, el empleo de la variable combinada nos permite analizar más matizadamente el efecto conjunto de ambos niveles del SES. Tercero, en todos los cursos y asignaturas, la nueva variable combinada es la que segmenta al primer nivel y distinguiendo siempre y con muy pocas variaciones entre los mismos subgrupos. Además esta variable vuelve a aparecer con frecuencia en niveles posteriores de segmentación. Y en técnicas de segmentación, estos dos hechos (aparecer como primera variable de segmentación y aparecer frecuentemente en diversos niveles) son claros indicadores de que la variable en cuestión es importante a la hora de elaborar e interpretar un modelo en la medida en que ello debe considerarse a la luz, no sólo de la diferencia de medias, sino también del incremento de la capacidad explicativa del modelo y de los criterios de parada del algoritmo. Esto contribuye a informarnos de la importancia relativa de la variable en el contexto global del modelo planteado y nos reafirma en lo apropiado del enfoque adoptado al generar esta nueva variable.

Como prueba de lo dicho, examinemos la Figura 5. En ella se muestra el árbol final resultante para las puntuaciones de matemáticas de 4º curso. Como vemos, es un árbol de 2 niveles con 4 nodos terminales, resultante del efecto de una única variable de segmentación, la nueva variable combinada, lo que confirma lo anteriormente expresado. Aquí los nodos más extremos son el 3 y el 5. El primero contiene los 566 sujetos con menor media aritmética en rendimiento en matemáticas (11,48); el segundo está compuesto por los 333 sujetos que obtienen las mejores puntuaciones (media = 17,85). Y estos últimos sujetos pertenecen exclusivamente a los subgrupos 44 y 41. Si segmentásemos este nodo, veríamos que estos 333 estudiantes se dividen entre los 307 del subgrupo 44 que son los que, como cabía esperar, obtienen el rendimiento más alto (media = 17,86) y los 26 del 41 (media = 17,73). Lo que nos resulta llamativo es que sean éstos los que obtengan la segunda media más alta de todo el grupo por delante de los otros subgrupos de centros educativos de nivel socioeconómico alto como el 42 y el 43, o de aquellos estudiantes que acuden a centros de nivel medio-alto pero cuyo status socioeconómico familiar es igualmente alto o medio-alto (33 y 34).

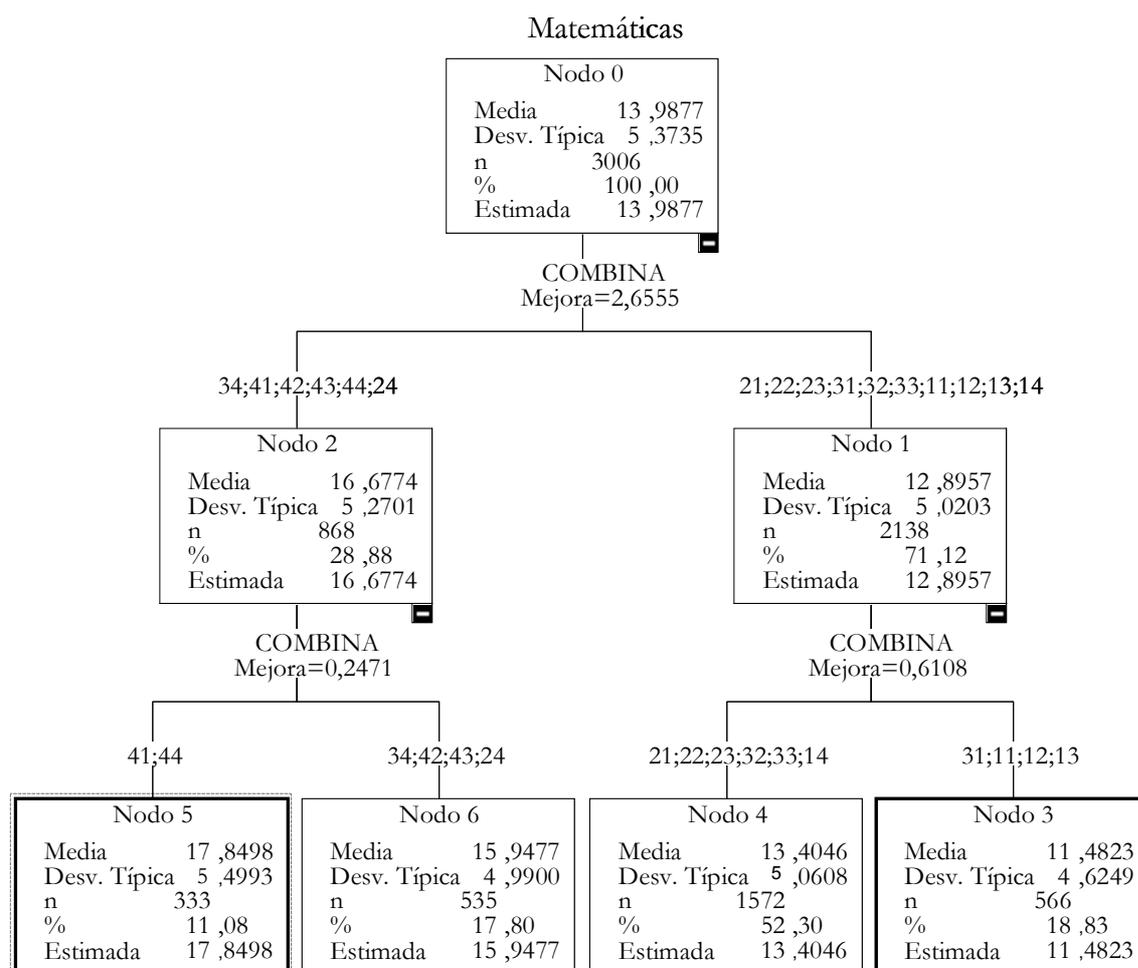


Figura 5. Segmentación final (Matemáticas 4º curso). Árbol de dos niveles

Está claro que se trata de un grupo formado por pocos casos, pero lo notable es que esta tendencia, con ligeras diferencias, se da en todos los cursos y en las dos asignaturas, lo que puede comprobarse examinando la Tabla 13. En la misma hemos resaltado los datos de los grupos 41 en los tres cursos y en ambas materias añadiendo el ordinal del valor de la media aritmética con respecto a los 16 subgrupos. Como ya apuntamos, la tendencia es clara, aunque los grupos son pequeños y hay diferencias según las asignaturas y los cursos, la media de este subgrupo es siempre de las más altas. Siempre por detrás del grupo 44 (excepto en Lengua de 4º donde resulta ser la primera) y habitualmente por encima de grupos como el resto de los de nivel socioeconómico de centro educativo alto (42 y 43) o de los 33 y 34.

Tabla 13

Descripción de la Puntuación en las Asignaturas por Cursos

SES Centro	SES Familiar	\bar{X} N s	1º ESO		2º ESO		4º ESO	
			Matemáticas	Lengua	Matemáticas	Lengua	Matemáticas	Lengua
Bajo	Bajo		13,75	14,42	12,67	13,54	10,50	15,47
1	1		195	194	244	240	243	244
			5,60	5,23	5,04	4,57	4,39	4,66

Bajo	Medio-bajo	\bar{X}	14,87	15,39	13,01	14,73	12,33	16,24
1	2	N	127	127	146	147	144	150
		s	5,10	5,27	5,08	4,07	4,84	4,99
Bajo	Medio-alto	\bar{X}	14,69	15,37	14,21	15,79	12,06	16,73
1	3	N	71	72	72	72	71	79
		s	5,27	5,60	5,99	4,72	5,04	5,62
Bajo	Alto	\bar{X}	15,00	15,61	15,75	17,09	12,81	17,90
1	4	N	22	23	36	35	42	41
		s	6,29	6,07	5,69	4,07	4,67	5,13
Medio-bajo	Bajo	\bar{X}	14,62	14,94	13,01	13,73	13,03	16,31
2	1	N	242	241	261	263	261	256
		s	5,05	5,06	5,00	4,35	5,30	4,37
Medio-bajo	Medio-bajo	\bar{X}	15,62	15,60	14,86	14,59	13,87	17,22
2	2	N	229	231	264	263	263	262
		s	5,37	5,40	4,92	4,48	5,19	4,29
Medio-bajo	Medio-alto	\bar{X}	16,60	16,66	15,50	15,16	14,28	17,30
2	3	N	194	192	230	234	235	233
		s	5,39	5,28	5,13	4,09	5,08	4,45
Medio-bajo	Alto	\bar{X}	16,15	16,38	16,84	16,37	15,63	18,14
2	4	N	118	119	116	114	130	130
		s	5,32	5,70	4,92	3,98	4,77	4,17
Medio-alto	Bajo	\bar{X}	15,76	16,18	13,88	14,80	12,18	16,70
3	1	N	80	79	100	99	108	110
		s	5,27	5,16	5,24	4,44	4,19	4,17
Medio-alto	Medio-bajo	\bar{X}	15,57	16,18	15,54	15,90	13,40	17,20
3	2	N	137	137	181	181	162	163
		s	5,07	5,43	4,99	4,15	4,59	4,75
Medio-alto	Medio-alto	\bar{X}	16,71	17,10	15,91	15,52	14,00	17,55
3	3	N	164	164	186	184	191	194
		s	4,82	4,78	5,64	4,14	4,42	4,20
Medio-alto	Alto	\bar{X}	17,34	17,79	17,43	17,48	15,66	18,74
3	4	N	178	179	210	210	187	188
		s	4,90	5,66	5,08	3,72	4,91	4,38
Alto	Bajo	\bar{X}	17,83	18,93	18,41	16,11	17,73	19,88
		<i>Orden</i>	(4°)	(2°)	(2°)	(4°)	(2°)	(1°)
4	1	N	29	29	17	18	26	26
		s	4,50	5,90	5,80	3,53	5,85	3,30
Alto	Medio-bajo	\bar{X}	18,60	17,71	17,10	15,96	16,65	18,99
4	2	N	55	56	50	50	79	79
		s	5,06	5,70	4,40	4,44	4,63	4,32
Alto	Medio-alto	\bar{X}	17,99	18,44	17,88	17,03	16,24	18,68
4	3	N	116	117	151	151	139	139
		s	5,17	5,83	4,63	3,88	5,46	4,79
Alto	Alto	\bar{X}	19,00	19,00	19,57	17,94	17,86	19,77
4	4	N	229	230	281	281	307	307
		s	4,82	5,43	5,07	4,07	5,48	4,16
Total		\bar{X}	16,19	16,49	15,56	15,57	14,20	17,53
		N	2186	2190	2545	2542	2588	2601
		s	5,37	5,54	5,52	4,43	5,37	4,64

Conclusiones, Discusión e Investigación Adicional

A la vista de los resultados obtenidos tanto en la formulación y estimación de los modelos jerárquicos, como de las técnicas de segmentación; y en función de los objetivos planteados, de forma resumida podemos presentar las siguientes conclusiones finales.

En primer lugar, hemos obtenido evidencia empírica suficiente del efecto contextual del nivel socioeconómico medio de los centros escolares en el rendimiento académico, y la hipótesis de doble riesgo ha quedado comprobada. En consecuencia, en nuestra muestra sí podemos afirmar que los estudiantes de nivel socioeconómico familiar bajo obtienen un rendimiento inferior cuando son escolarizados en centros educativos cuyo nivel socioeconómico medio es también bajo. Por el contrario, la hipótesis de triple riesgo no ha sido verificada al no haber detectado efectos significativos de interacción entre ambas medidas del SES, de forma que no podemos hablar de que este efecto contextual sea más fuerte para estos estudiantes.

Segundo, el modelo definitivo estimado incluye también como variables relevantes una de cada nivel considerado, el género (Sexo) de los estudiantes y la titularidad (pública/privada) de los centros (Tipo). Aquí los resultados son básicamente coincidentes con las investigaciones citadas, siendo preciso señalar como limitación de este trabajo el limitado número de variables explicativas que hemos podido incorporar a los modelos estudiados.

Tercero, el generar la nueva variable combinada—resultado del producto cruzado de las dos medidas categorizadas del SES de estudiantes y centros—se ha revelado como una estrategia apropiada que ha facilitado una interpretación más fina del modelo jerárquico estimado. Se trata de una variable de gran capacidad explicativa en la medida en que para las dos asignaturas y todos los cursos es la que con mayor frecuencia aparece como variable de segmentación. El empleo de la misma nos ha permitido detectar además que el rendimiento académico medio del subgrupo 41 (SESC alto y SESF bajo) ha resultado ser superior al de subgrupos con iniciales mejores expectativas. Aunque el tamaño de dichos subgrupos es pequeño, el hecho de que esto ocurra en todos los cursos y en las dos asignaturas nos induce a pensar que es posible que nos encontremos ante una tendencia que puede merecer un estudio más amplio y detallado sin las limitaciones muestrales y de diseño que hemos señalado. Y, para acabar, vamos brevemente ahora a detenernos en estas cuestiones.

Aunque con las limitaciones y precauciones que ya hemos señalado (fundamentalmente en lo relativo a la no aleatoriedad de la muestra y al bajo número de casos en la clase 41), pensamos que el comportamiento de este subgrupo merece un estudio detenido en la medida en que no deja de ser llamativo que los sujetos de menor nivel de SES familiar y que acuden a centros escolares de alto nivel socioeconómico obtengan tan excelentes resultados.

Y esto es así en todos los cursos y en ambas materias, como antes apuntábamos. El subgrupo 41 se sitúa por encima de subgrupos de SES familiar medio-alto o alto, como por ejemplo los grupos 33 y 34. En este caso, bien pudiera argüirse que esto se debe al efecto del SES del centro educativo. Sin embargo, cabe preguntarse entonces, ¿por qué casi en la totalidad de los casos obtienen promedios superiores a los de los grupos 42 y 43, cuando éstos subgrupos también asisten a centros escolares de alto nivel socioeconómico y además, poseen un mayor status socioeconómico familiar?

Estos interrogantes nos llevan, en primer lugar, a examinar con detalle a estos subgrupos para lograr un conocimiento más preciso de los mismos. A continuación, finalizaremos este trabajo apuntando algunas posibles hipótesis explicativas que pueden guiar posteriores investigaciones.

Procedamos en primer lugar a caracterizar a estos subgrupos. Se trata, en los tres cursos, de grupos formados por un pequeño número de estudiantes que cursan sus estudios en centros privados religiosos concertados. Esto en sí mismo puede ser un primer indicador, pero, por otra parte, hay que tener presente que en nuestra muestra no hay ningún centro público cuyo nivel socioeconómico medio sea alto. Esta carencia puede atribuirse a las características de la muestra donde, como dijimos al inicio, los centros privados están sobre-representados. Sin embargo, nuestro conocimiento del sistema educativo de la CAV nos induce a pensar que han de ser muy pocos los centros públicos de nivel socioeconómico alto, pues una gran mayoría de las familias con estos niveles de status económico opta por los centros privados a la hora de escolarizar a sus hijos.

En cualquier caso, un análisis de las características diferenciales de los centros educativos tendría sentido si los estudiantes de este grupo estuviesen adscritos, total o mayoritariamente, a un subconjunto de estos colegios. Pero, examinada la distribución de estos estudiantes por centro escolar en los tres cursos, hemos constatado que se distribuyen entre *todos* los centros de estas características incluidos en la muestra. Por ejemplo, los 29 estudiantes de 1er curso se distribuyen entre los 6 centros escolares de la muestra de nivel 4.

Desde un punto de vista meramente estadístico, estos subgrupos obtienen un promedio tan elevado en rendimiento porque se trata de pequeños grupos muy homogéneos con un máximo y, sobre todo, con un mínimo superiores al de la gran mayoría del resto de los subgrupos. La Tabla 14 ilustra lo dicho para el caso del rendimiento en Lengua en primer curso.

Tabla 14

Descripción de las Modalidades de "Combina" en Lengua de 1er Curso

SES Centro	SES Familiar	Media	N	Desv. típ.	Mínimo	Máximo
Bajo	Bajo	14,42	194	5,23	4,0	25,0
Bajo	Medio-bajo	15,39	127	5,27	3,0	30,0
Bajo	Medio-alto	15,37	72	5,60	3,0	26,0
Bajo	Alto	15,61	23	6,07	6,0	25,0
Medio-bajo	Bajo	14,94	241	5,06	3,0	27,0
Medio-bajo	Medio-bajo	15,60	231	5,40	3,0	30,0
Medio-bajo	Medio-alto	16,66	192	5,28	6,0	28,0
Medio-bajo	Alto	16,38	119	5,70	6,0	30,0
Medio-alto	Bajo	16,18	79	5,16	7,0	27,0
Medio-alto	Medio-bajo	16,18	137	5,43	6,0	28,0
Medio-alto	Medio-alto	17,10	164	4,78	5,0	27,0
Medio-alto	Alto	17,79	179	5,66	3,0	30,0
Alto	Bajo	18,93	29	5,90	9,0	30,0
Alto	Medio-bajo	17,71	56	5,70	5,0	30,0
Alto	Medio-alto	18,44	117	5,83	4,0	29,0
Alto	Alto	19,00	230	5,43	5,0	29,0
Total		16,49	2190	5,54	3,0	30,0

Si comparamos el grupo 41 con el extremo más bajo de la escala (subgrupo 11), vemos como los 4,5 puntos de diferencia entre las medias se explican en razón de lo dicho. En el diagrama de la Figura 6 se muestra la distribución de las puntuaciones de ambos subgrupos. Lógicamente, no se trata de que en el subgrupo 11 no haya individuos con puntuaciones altas (los hay en todos), sino que son minoría en comparación con el resto. Por el contrario, los sujetos del subgrupo 41 distribuyen sus puntuaciones en un rango menor, rango cuyos mínimo y máximo son además bastante superiores.

Figura 6. Histograma de las distribuciones de Lengua en 1er curso para los subgrupos 11 y 41.

Ante todo esto, una primera hipótesis que es natural e inmediatamente planteable es si, dado que se trata de centros privados, no habrá un mecanismo de selección inicial de estudiantes en el momento en que estos se incorporan a la enseñanza secundaria. Pero ante esto hay que decir claramente que en el sistema educativo español, como norma general e independientemente de si el centro es público o privado concertado, ni se selecciona a la entrada, ni se filtra a lo largo del proceso, al menos en el periodo obligatorio. Mizala et al. (2004) presentan una situación muy distinta en el caso del sistema educativo chileno, incluso en el caso de las escuelas privadas concertadas.

Por el contrario, y con las cautelas que imponen nuestras limitaciones muestrales y la propia distribución poblacional entre centros escolares públicos y privados en la CAV, quizá el hecho de que sean centros religiosos sí pueda ser significativo. Esto sería coincidente con los trabajos de autores como Coleman, Hoffer & Kilgore (1982 a y b); Raudenbush y Bryk (1986) y Lee y Bryk (1989) en los que se plantea que los institutos católicos norteamericanos parecen favorecer una más equitativa distribución del rendimiento escolar. Como posibles causas apuntan a una estructura académica más rígida o constreñida, a un entorno normativo que genera un clima escolar que para la mayoría de los estudiantes resulta servir de apoyo, independientemente de su origen social y de sus capacidades académicas. De cualquier forma, insistimos que en nuestra muestra no hay centros públicos cuyo nivel socioeconómico medio sea alto y que todos los privados son centros concertados regidos por órdenes religiosas, por lo que carecemos de elementos suficientes de comparación.

Resultados como los obtenidos aquí, remiten también al enfoque del estudio de la resiliencia académica. Está claro que no habría que limitarse al estudio de este subgrupo (11), sino que una línea de trabajo podría consistir en establecer la distinción entre estudiantes resilientes y no resilientes (en todos los subgrupos) y posteriormente, mediante el estudio diferencial de ambos, tratar de caracterizar a los primeros con el fin de contrastar hipótesis que expliquen sus buenos resultados. En nuestro caso hemos aplicado el criterio que Borman y

Overman (2004) proponen en su estudio sobre resiliencia académica en matemáticas. Hemos calculado la puntuación esperada para cada sujeto en función del modelo definitivo estimado y luego hemos obtenido los residuos. Aquellos estudiantes con residuos tipificados iguales o superiores a 0,33 son asignados al grupo de resilientes, mientras que los de valores iguales o inferiores a -0,33 al de no resilientes.

Con las variables disponibles en nuestro estudio no hemos encontrado ninguna característica diferencial entre estos dos grupos en relación con las 16 clases generadas por la variable combinada, tomando en consideración el resto de las variables predictoras. Por ejemplo, ni los estudiantes resilientes están asignados a un subconjunto específico de centros, ni su proporción varía significativamente entre las 16 clases.

Tanto los estudios sobre resiliencia académica, como los centrados en el efecto contextual, apuntan a factores explicativos que a nivel de los individuos se centran en la actitud hacia el estudio, la motivación académica, la implicación en las actividades del centro escolar y un adecuado nivel de autoestima. En lo relativo al entorno familiar, destacan el nivel de implicación y apoyo paterno, las actitudes favorables de los padres hacia el hecho educativo y las propias expectativas de éstos hacia sus hijos. Por último, a nivel de las escuelas destacan cuestiones como el clima escolar, el efecto de los compañeros (*peer effect*), interacción positiva profesor-alumno, ambiente positivo de disciplina y un adecuado entorno normativo. Los centros escolares de nivel socioeconómico alto tienden a tener un cuadro de docentes más capacitado y motivado. Willms (1989, citado en Coleman, 1990) subsume muchos de estos factores en el concepto de capital social lo que le lleva a afirmar que en comunidades donde se da un alto nivel de capital social los resultados mejoran y las desigualdades se reducen.

Pero todas estas hipótesis plantean también algunas dificultades. Ma y Kingler (2000) apuntan a que algunos de estos indicadores son problemáticos y ponen como ejemplo el caso de la implicación paterna ya que, según afirman los autores, una alta frecuencia de contactos entre profesores y familias puede deberse a casos de problemas conductuales o de rendimiento, que usualmente tienden a afectar en mayor grado a familias de bajo nivel socioeconómico.

Sin embargo, por encima de dificultades concretas, en muchas de las hipótesis explicativas sugeridas planea un problema de más envergadura. Pensemos en factores como la autoestima, el clima escolar favorable, la valoración paterna, la actitud y la motivación, etc. que acabamos de citar; creemos que es arriesgado plantear aquí un esquema de causalidad unidireccional. Jeynes (2002) hace un análisis de estos estudios y de las variables incorporadas a los mismos en calidad de explicativas, llegando a concluir que es difícil establecer esta atribución de manera unívoca y formulando para el caso del constructo SES un modelo de *causalidad recíproca*.

En definitiva, si algo parece evidente es que en el campo de las relaciones entre el SES y el rendimiento académico sigue siendo necesaria investigación adicional. Con respecto a las cuestiones que en nuestro estudio hemos abordado, dicha investigación debería ir encaminada a obtener evidencia empírica de la consistencia de los buenos resultados de grupos similares a los que nosotros hemos etiquetado como subgrupo 41. Por nuestra parte, en la actualidad estamos trabajando con mayores y más representativas muestras e incorporando a los modelos un conjunto más amplio de variables tanto en el nivel de los estudiantes como en el de los centros educativos.

Referencias

- Álvaro Page, M., Bueno, M. J., Calleja, J. A., Cerdán, J., Echeverría, M. J., García, C. et al. (1990). *Hacia un modelo causal del rendimiento académico*. Madrid: Centro de Publicaciones del Ministerio de Educación y Ciencia (C.I.D.E).
- Borman, G. D. & Overman, L. (2004). Academic Resilience in Mathematics among Poor and Minority Students. *The Elementary School Journal* 104(3), pp. 177-195.
- Brookover, W. B., Schweitzer, J. H., Schneider, J. M., Beady, C. H., Flood, P. K. & Wisenbaker, J. M. (1978). Elementary school social climate and school achievement. *American Educational Research Journal* 15(2), pp. 301-318.
- Coleman, J. S., Hoffer, T. & Kilgore, S. B. (1966). *Equality of educational opportunity*. Washington: US Government Printing Office.
- Coleman, J. S., Hoffer, T. & Kilgore, S. B. (1982a). Cognitive outcomes in public and private schools. *Sociology of Education*, 55, pp. 65-76.
- Coleman, J. S., Hoffer, T. & Kilgore, S. B. (1982b). *High School Achievement: Public, Catholic and Other Schools Compared*. New York: Basic Books.
- Coleman, J. S. (1990). *Foundations of Social Theory*. Cambridge, MA: Harvard University Press.
- D'Angiulli, A., Siegel, L. S. & Maggi, S. (2004a). Literacy Instruction, SES, and Word-Reading Achievement in English-Language Learners and Children with English as a First Language: A Longitudinal Study. *Learning Disabilities Research & Practice*, 19(4), pp. 202-213.
- D'Angiulli, A., Siegel, L. S. & Hertzman, C. (2004b). Schooling, Socioeconomic Context and Literacy Development. *Educational Psychology* 24(6), pp. 867-883.
- De Fraine, B.; Van Damme, J.; Onghena, P. (2002). Accountability of Schools and Teachers: what should be taken into account? *European Educational Research Journal*, Vol. 1, No. 3.
- Dutton, D. B., & Levine, S. (1989). Overview, Methodological Critique, and Reformulation. In J. P. Bunker, D. S. Gomby & B. H. Kehrer (Eds.), *Pathways to Health* (pp. 29-69). Menlo Park, CA: The Henry J. Kaiser Family Foundation.
- Gaviria, J., Martínez-Arias, R. & Castro, M. (2004). Un estudio multinivel sobre los factores de eficacia escolar en países en desarrollo: el caso de los recursos en Brasil. *Education Policy Analysis Archives*, 12 (20). Recuperado de <http://epaa.asu.edu/epaa/v12n20/>
- Goldstein, H. (1987). *Multilevel Models in Educational and Social Research*. London: Griffin and New York: Oxford Press.
- Goldstein, H. (1995). *Multilevel Statistical Models*. London: Edward Arnold.
- Hambleton, R. K. (1996). Adaptación de tests para su uso en diferentes idiomas y culturas. Fuentes de error, posibles soluciones y directrices prácticas. En J. Muñiz (Coord.), *Psicometría* (pp. 207-238). Madrid: Universitas.
- Henderson, V., Mieszkowski, P. & Sauvageau, Y. (1978). Peer Group Effects and Educational Production Functions. *Journal of Public Economics* 10(1), pp.97-106.
- Hoff, E. & Tian, C. (2005). Socioeconomic status and cultural influences on language. *Journal of Communication Disorders*. 38, pp. 271-278.
- Howley, C. B. & Howley, A. A. (2004). School size and the influence of socioeconomic status on student achievement: Confronting the threat of size bias in national data sets. *Education Policy Analysis Archives*, 12(52). Recuperado de <http://epaa.asu.edu/epaa/v12n52/>
- INCE. MEC. (2000): *Evaluación de la educación primaria 1999*. Recuperado el 21 de Septiembre de 2006 de <http://www.ince.mec.es/pub/evalprim99.pdf>
- ISEI-IVEI. (2005). *Segundo informe de la evaluación PISA 2003. Resultados en Euskadi*. Recuperado el 21 de Septiembre de 2006 de <http://www.isei-ivei.net>
- Jencks, C, Smith, M., Ackland, H. M., Cohen, J. D., Heyns, H. G. B. & Michelson, S. (1972). *Inequality: A reassessment of the effect of family and schooling in America*. New York: Basic Books.
- Jeynes, W. H. (2002). The Challenge of Controlling for SES in Social Science and Education Research. *Educational Psychology Review*, 14 (2), pp. 205-221.

- Kim, S. H. (1991). *An extension of CART's Pruning Algorithm*. Program Statistics Research Technical Report No 91-11. Educational Testing Service, Princeton, New Jersey.
- Kobrin, M. (2003). *An Investigation of School-Level Factors for Students with Discrepant High School GPA and SAT Scores*. Paper presented at the Annual Meeting of the National Council on Measurement in Education. Chicago. 2003. (ERIC Document Reproduction Service No. ED 476921).
- Lapsley, D. K. (2002). *Teacher aides, class size and academic achievement: a preliminary evaluation of Indiana's Prime Time*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans. (ERIC Document Reproduction Service No. ED 466679).
- Lee, V. & Bryk, A. (1989). A multilevel model of the social distribution of high school achievement. *Sociology of Education* 62(3), pp. 172-192.
- Lizasoain, L., Joaristi, L., Santiago, C., Lukas, J. F., Moyano, N., Sedano, M., & Munárriz, B. (2003). El uso de las técnicas de segmentación en la evaluación del rendimiento en lenguas. Un estudio en la Comunidad Autónoma Vasca. *Revista de Investigación Educativa (RIE)*, 21(1), pp. 93-111.
- Lubienski, S. T. (2006). Examining instruction, achievement, and equity with NAEP mathematics data. *Education Policy Analysis Archives*, 14(14). Retrieved September 21, 2006 from <http://epaa.asu.edu/epaa/v14n14/>.
- Lukas, J. F. et al. (2000). Euskal Autonomia Erkidegoko ikastetxeen kanpo-ebaluazioa. *Tantak*, 24, pp. 39-63.
- Lukas, J. F. & Santiago, K. (2004). Evaluación de centros escolares de educación secundaria del País Vasco. *Revista Electrónica de Investigación Educativa*. 6(1). Recuperado de (<http://redie.uabc.mx/vol6no2/contenido-lukas.html>)
- Ma, X. & Klinger, D. A. (2000). Hierarchical Linear Modelling of Student and School Effects on Academic Achievement. *Canadian Journal of Education*, 25 (1), pp. 41-55.
- Marchesi, A. & Martín, E. (Comp.) (2002). *Evaluación de la educación secundaria. Fotografía de una etapa polémica*. Madrid: SM.
- Marchesi, A. & Martínez Arias, R. (2002). Un modelo multinivel para evaluar la educación secundaria. En A. Marchesi & E. Martín (Comp.), *Evaluación de la educación secundaria. Fotografía de una etapa polémica*. Madrid: SM.
- Mizala, A., Romaguera, P. & Ostoic, C. (2004). *Equity and Achievement in the Chilean School Choice System*. Econometric Society 2004 Latin American Meetings 232, Econometric Society. Retrieved April, 2006 from <http://repec.org/esLATM04/up.13209.1082053943.pdf>
- Mueller, C. W., & Parcel, T. L. (1981). Measures of socioeconomic status: Alternatives and recommendations. *Child Development*, 52, pp. 13-30.
- Raudenbush, S. W. & Bryk, A. S. (1986). A hierarchical model for studying school effects. *Sociology of Education* 59 (1), pp. 1-17.
- Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods (2nd ed)*. Thousand Oaks, CA: Sage.
- Raudenbush, S. W., Bryk, A. S., Cheong, Y. F. & Congdon, R. T. (2005). *HLM 6: Hierarchical Linear and Nonlinear Modelling*. Chicago, IL: Scientific Software International.
- REDES equipo (1999). Una red de evaluación de centros de Educación Secundaria. *Infancia y Aprendizaje*, 85, pp. 59-73.
- REDES equipo (2000). Variables predictoras de la calidad de la educación secundaria. *Revista de Educación*, 323, pp. 395-420.
- Rumberger, R., & Willms, J. D. (1992). The impact of racial and ethnic segregation on the achievement gap in California high schools. *Educational Evaluation and Policy Analysis*, 14 (4), pp. 337-396.

- Rumberger, R. W. & Palardy, G. J. (2005). Does Segregation Still Matter? The Impact of Student Composition on Academic Achievement in High School. *Teachers College Record*, 107 (9), pp. 1999–2045.
- Santiago, C. & Lukas, J. F. (2005). Evaluación externa de centros en la Comunidad Autónoma Vasca. En J. Jornet & G. Ramos (Coord.), *Problemas de la medición y evaluación educativa. Estándares e indicadores para analizar la realidad educativa* (pp. 326-349). Valencia, España: Universidad de Valencia.
- Shavit, Y. & Williams, R. A. (1985). Ability Grouping and Contextual Determinants of Educational Expectations in Israel. *American Sociological Review* 50 (1), pp. 62-73.
- SPSS (2001). *Answer Tree 3.0 User's guide*. Chicago, IL: SPSS Inc.
- Tajalli, H. & Opheim, C. (2004). Strategies for Closing the Gap: Predicting Student Performance in Economically Disadvantaged Schools. *Educational Research Quarterly* 28 (4), pp.44-54.
- Tiana, A. (2002). El contexto sociocultural en la evaluación de los centros educativos. En A. Marchesi & E. Martín (Comp.), *Evaluación de la educación secundaria. Fotografía de una etapa polémica*. Madrid: SM.
- Väljjarvi, J. & Malin, A. (2003). The two-level effect of socio-economic background. In S. Lie, P. Linnakylä & A. Roe (Eds.), *Northern Lights on PISA: Unity and diversity in Nordic countries in PISA 2000* (pp. 123–132). Oslo, Norway: University of Oslo, Department of Teacher Education and School Development.
- Warschauer, M., Knobel, M. & Stone, L. (2004). Technology and Equity in Schooling: Deconstructing the Digital Divide. *Educational Policy*, 18 (4), 562-588.
- Willms, J. D. (1986). Social Class Segregation and Its Relationship to Pupil's Examination Results in Scotland. *American Educational Review* 51 (2), pp.224-241.
- Willms, J. D., Raudenbush, S. W. (1989). A Longitudinal Hierarchical Linear Model for Estimating School Effects and Their Stability. *Journal of Educational Measurement* 26 (3), pp. 209- 232.
- Willms, J. D. (2001). Three Hypotheses about Community Effects on Social Outcomes. *ISUMA*, 2(1), pp. 53-62.
- Willms, J. D. (2003). *Ten Hypotheses about Socioeconomic Gradients and Community Differences in Children's Developmental Outcomes*. Final Report. Applied Research Branch. Strategic Policy. Human Resources Development. Canada. Retrieved April 26, 2006 from <http://www.hrdc-drhc.gc.ca/sp-ps/arb-dgra>
- Willms, J. D. (2004). *Reading Achievement in Canada and the United States: Findings from the OECD Programme of International Student Assessment*. Final Report. Learning Policy Directorate Strategic Policy and Planning Human Resources and Skills Development. Canada. Retrieved April 25, 2006 from <http://www.hrsdc-rhdcc.gc.ca/sp-ps/arb-dgra>
- Willms, J. D. (2006). *Learning Divides: Ten Policy Questions About The Performance And Equity Of Schools And Schooling Systems*. UNESCO Institute for Statistics, Montreal.
- White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin*, 91 (3), pp. 461-481.
- Yang, Y. (2003). Dimensions of Socioeconomic Status and their Relationship to Mathematics and Science Achievement at Individual and Collective Levels. *Scandinavian Journal of Educational Research*, 47 (1), pp. 21-41.
- Yang, Y. & Gustafsson J. (2004). Measuring Socioeconomic Status at Individual and Collective Levels. *Educational Research and Evaluation* 10(3), pp. 259-288.

Datos Biográficos de los Autores:

Los cuatro autores son Profesores Titulares del Departamento de Métodos de Investigación y Diagnóstico en Educación de la Facultad de Filosofía y Ciencias de la Educación de la Universidad del País Vasco-Euskal Herriko Unibertsitatea.

Luis Lizasoain es Doctor en Filosofía y Ciencias de la Educación por la Universidad Complutense de Madrid. Su labor investigadora se centra en la evaluación de programas y sistemas educativos, en la aplicación de los modelos multinivel a la investigación educativa y en la recuperación de la información en bases de datos, informática y lingüística documental.

Luis Joaristi es Ingeniero Industrial y Doctor en Ciencias de la Educación por la Universidad del País Vasco. Su labor investigadora se centra en la evaluación de sistemas educativos, en los métodos estadísticos multivariados, en la aplicación de los modelos multinivel a la investigación educativa y en las técnicas estadísticas para los estudios longitudinales y el estudio del cambio.

José Francisco Lukas es Doctor en Ciencias de la Educación por la Universidad del País Vasco. Su labor investigadora se centra en la evaluación de programas y sistemas educativos, en los métodos y técnicas de medición en educación y en la evaluación de la acción educativa de las entidades locales (ayuntamientos).

Karlos Santiago es Doctor en Ciencias de la Educación por la Universidad del País Vasco. Su labor investigadora se centra en la evaluación de programas y sistemas educativos, en los métodos y diseños cuasi-experimentales en la investigación educativa y en la evaluación de los programas y de la acción educativa de las entidades locales (ayuntamientos).

La correspondencia relacionada con este artículo debe dirigirse a Luis Lizasoain, Departamento de Métodos de Investigación y Diagnóstico en Educación, Facultad de Filosofía y Ciencias de la Educación, Universidad del País Vasco-Euskal Herriko Unibertsitatea, Avda. de Tolosa 70, 20019 San Sebastián, Spain. Tfno: +34 943 015 514. E-mail: luis.lizasoain@ehu.es

Archivos Analíticos de Políticas Educativas <http://epaa.asu.edu>

Editores

Gustavo E. Fischman Arizona State University

Pablo Gentili Universidade do Estado do Rio de Janeiro

Asistentes editoriales: Rafael O. Serrano (ASU) & Lucia Terra (UBC)

Hugo Aboites

UAM-Xochimilco, México

Claudio Almonacid Avila

UMCE, Chile

Alejandra Birgin

FLACSO-UBA, Argentina

Mariano Fernández Enguita

Universidad de Salamanca. España

Roberto Leher

UFRJ, Brasil

Pia Lindquist Wong

CSUS, USA

Alma Maldonado

University of Arizona, USA

Imanol Ordorika

IIE-UNAM, México

Miguel A. Pereyra

Universidad de Granada, España

Romualdo Portella de Oliveira

Universidade de São Paulo, Brasil

José Ignacio Rivas Flores

Universidad de Málaga, España

José Gimeno Sacristán

Universidad de Valencia, España

Susan Street

CIESAS Occidente, México

Daniel Suárez

LPP-UBA, Argentina

Jurjo Torres Santomé

Universidad de la Coruña, España

Armando Alcántara Santuario

CESU, México

Dalila Andrade de Oliveira

UFMG, Brasil

Sigfredo Chiroque

IPP, Perú

Gaudêncio Frigotto

UERJ, Brasil

Nilma Lino Gomes

UFMG, Brasil

María Loreto Egaña

PIIE, Chile

José Felipe Martínez Fernández

UCLA, USA

Vanilda Paiva

UERJ, Brasil

Mónica Pini

UNSAM, Argentina

Paula Razquin

UNESCO, Francia

Diana Rhoten

SSRC, USA

Daniel Schugurensky

UT-OISE Canadá

Nelly P. Stromquist

USC, USA

Antonio Teodoro

Universidade Lusófona, Lisboa

Lílian do Valle

UERJ, Brasil

EDUCATION POLICY ANALYSIS ARCHIVES <http://epaa.asu.edu>**Editor: Sherman Dorn, University of South Florida**

Production Assistant: Chris Murrell, Arizona State University

General questions about appropriateness of topics or particular articles may be addressed to the Editor, Sherman Dorn, epaa-editor@shermamdorn.com.

Editorial Board

Noga Admon	Jessica Allen
Cheryl Aman	Michael W. Apple
David C. Berliner	Damian Betebenner
Robert Bickel	Robert Bifulco
Anne Black	Henry Braun
Nick Burbules	Marisa Cannata
Casey Cobb	Arnold Danzig
Linda Darling-Hammond	Chad d'Entremont
John Diamond	Amy Garrett Dickers
Tara Donohue	Gunapala Edirisooriya
Camille Farrington	Gustavo Fischman
Chris Frey	Richard Garlikov
Misty Ginicola	Gene V Glass
Harvey Goldstein	Jake Gross
Hee Kyung Hong	Aimee Howley
Craig B. Howley	William Hunter
Jaekyung Lee	Benjamin Levin
Jennifer Lloyd	Sarah Lubienski
Susan Maller	Les McLean
Roslyn Arlin Mickelson	Heinrich Mintrop
Shereeza Mohammed	Michele Moses
Sharon L. Nichols	Sean Reardon
A.G. Rud	Lorrie Shepard
Ben Superfine	Cally Waite
John Weathers	Kevin Welner
Ed Wiley	Terrence G. Wiley
Kyo Yamashiro	Stuart Yeh