



RELIEVE. Revista Electrónica de Investigación y
Evaluación Educativa

ISSN: 1134-4032

relievejournal@gmail.com

Universidad de Granada
España

Elosua, Paula
Diversidad lingüística y evaluación en centros multilingües
RELIEVE. Revista Electrónica de Investigación y
Evaluación Educativa, vol. 24, núm. 2, 2018, Enero-Junio
Universidad de Granada
Valencia, España

DOI: <https://doi.org/10.7203/relieve.24.2.11414>

Disponible en: <https://www.redalyc.org/articulo.oa?id=91664471007>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica Redalyc

Red de Revistas Científicas de América Latina y el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso
abierto

Diversidad lingüística y evaluación en centros multilingües

Linguistic diversity and assessment in multilingual schools

Elosua, Paula

Universidad del País Vasco

Abstract

In a context where linguistic diversity in the classroom is natural, educational assessment raises new challenges that demand the conjunction of different perspectives. The equity in the assessment process requires considering facets related to the psychometric properties of the tests, the development of linguistic competence, or the sociolinguistic context through the incorporation of context variables into the predictive models of performance. By analyzing population data from compulsory secondary education schools which offer more than one linguistic educational model in the Autonomous Community of the Basque Country (N = 45), this paper examines the responses of 2783 students to a mathematical competence test. The methodological approach is twofold; it includes a psychometric study of equivalence through the estimation of the differential item functioning with regard to the family language and the linguistic educational model, and the estimation of mixed linear models for the prediction of the mathematics competence. The results show: a) the equivalence of the assessment test in relation to the family language and educational model, b) the significant impact of the linguistic context variables on the competence estimation, c) the lack of significance of the linguistic models on performance. It is concluded the importance of conducting studies of differential functioning, and the incorporation of linguistic variables in the prediction of mathematical competence.

Keywords: educational assessment, student diversity, language of instruction, differential item functioning.

Reception Date

2018 January 25

Approval Date

2018 December 05

Publication Date:

2018 December 05

Resumen

En un entorno en el que la diversidad lingüística es natural, la metodología de evaluación afronta nuevos retos. La equidad en el proceso evaluativo aconseja la consideración de aspectos relacionados con las características psicométricas de las pruebas, el desarrollo de la competencia lingüística, o el contexto sociolingüístico, a través de modelos explicativos sobre las variables contextuales de naturaleza lingüística que inciden en el rendimiento. Utilizando datos poblacionales de centros educativos de segundo curso de enseñanza secundaria obligatoria que ofertan más de un modelo lingüístico en la Comunidad Autónoma del País Vasco (N=45), este trabajo analiza las respuestas de 2783 estudiantes a una prueba de competencia matemática y estudia el efecto de la lengua familiar (euskera o castellano). La aproximación metodológica incluye un doble acercamiento; un estudio psicométrico de equivalencia a través del análisis del funcionamiento diferencial del ítem en función de lengua familiar y del modelo educativo, y el ajuste de modelos lineales mixtos para la predicción de la competencia. Los resultados muestran: a) la equivalencia de la prueba de evaluación en función del idioma familiar y del modelo educativo, b) el impacto significativo de variables de contexto lingüístico, y c) la falta de significatividad del modelo lingüístico sobre el desempeño. Se concluye la importancia de llevar a cabo estudios de funcionamiento diferencial, y de la incorporación de variables de naturaleza lingüística en la predicción de la competencia matemática.

Palabras clave: evaluación educativa, diversidad, lenguaje de instrucción, funcionamiento diferencial del ítem.

Fecha de recepción

2018 Enero 25

Fecha de aprobación

2018 Diciembre 05

Fecha de publicación

2018 Diciembre 05

La diversidad lingüística en el aula lejos de ser una rareza o una situación inusual es un escenario común en nuestra práctica educativa.

Los proyectos de revitalización de lenguas propias, la inclusión de programas de multilingüismo que incorporan lenguas

Autor de contacto / Corresponding author

Elosua, Paula. Departamento de Psicología Social y Metodología de las Ciencias del Comportamiento. Universidad del País Vasco (España). paula.elosua@ehu.es

extranjerías en las aulas, la migración y la globalización en general, definen espacios educativos de diversidad.

En un país con 17 comunidades autónomas y cinco lenguas co-oficiales que conviven con el español, el aula es multilingüe. El aranés, catalán, gallego, vasco o euskera y valenciano son lenguas propias en seis comunidades autónomas, y el derecho a conocerlas, usarlas y recibir enseñanza en ellas está legalmente reconocido. Con el objetivo común de revitalización de las lenguas minorizadas, cada comunidad autónoma ha diseñado modelos lingüísticos para impulsar su conocimiento a la par que desarrollar las competencias lingüísticas de los estudiantes. Los modelos lingüísticos son dispares; desde la inmersión total en Cataluña, por ejemplo, hasta la coexistencia de tres modelos diferentes en el País Vasco. Un repaso a la legislación que ampara cada modelo lingüístico revela que: a) la ley de educación en Cataluña (Ley 12/2009 de 10 de julio) define al catalán como lengua vehicular; b) en Valencia, la ley de plurilingüismo (Decreto 9/2017 de 27 de Enero; aunque anulada parcialmente por el Tribunal Superior de Justicia de la Comunidad Valenciana) concreta tres niveles lingüísticos en función de la naturaleza de la lengua vehicular; c) en Baleares el decreto de lenguas extranjeras (Decreto 9/2017 de 27 de enero) permite a cada centro a diseñar su proyecto lingüístico y en función de él proponer la proporción de uso vehicular de cada una de las lenguas; d) en Galicia el decreto 79/2010, de 20 de mayo, determina cual es la lengua en la que se imparten algunas asignaturas dejando libertad al centro para elegir el idioma en el resto, siempre que se garantice que las asignaturas en gallego y en castellano se distribuyen en el mismo porcentaje de horas semanales; e) en el País Vasco el decreto 138/1983, de 11 de julio, perfila tres modelos lingüísticos que pueden convivir en un mismo centro y en los cuales la lengua vehicular puede ser el español, el euskera, o ambas; f) la comunidad foral de Navarra adapta el sistema del País Vasco para las zonas vascoparlantes y añade un nuevo modelo para aquellas áreas en las que no hay presencia del euskera (Ley foral 4/2015 de 10 de marzo).

La implementación de lenguas propias en el sistema educativo ha tenido que franquear entre otras dificultades, las relacionadas con la formación del profesorado, la elaboración de materiales docentes y/o la normalización lingüística. Junto a ellas, la evaluación educativa afronta asimismo desafíos específicos a los que ha de responder con metodologías adecuadas. La atención a la diversidad requiere considerar que en el grupo de estudiantes evaluados: a) la lengua materna y la lengua de instrucción pueden diferir; b) el nivel de competencia lingüística en el idioma en el que se lleva a cabo la evaluación puede ser insuficiente, c) las pruebas evaluativas pueden presentarse en más de un idioma, y/o, d) las lenguas en contacto pueden variar en características tan importantes como el estatus y el prestigio. Estos aspectos en su conjunto, son un claro reflejo del carácter multifacético que caracteriza la evaluación en situaciones de diversidad, y que abarca disciplinas como la psicolingüística, la psicometría, y la sociolingüística (Elosua, 2016). Desde cada una de ellas es posible considerar fuentes potenciales de sesgo que interactúan y amenazan la ecuanimidad del proceso, y que podrían estar ligadas a la competencia lingüística del estudiante, las características de las pruebas de evaluación, y el entorno sociolingüístico.

La competencia lingüística como fuente de sesgo, hace referencia al nivel mínimo de competencia, umbral, (Cummins, 1981, 2000) que un estudiante debe alcanzar para desenvolverse con fluidez en el entorno académico. La denominada competencia lingüística cognitiva y académica (CALP) se alcanzaría en un período de entre cinco y siete años, e implicaría la habilidad para expresar ideas y conceptos relevantes en el medio académico que son precisos para un desempeño óptimo. La segunda fuente de sesgo, está ligada a las propiedades psicométricas de las pruebas utilizadas en la evaluación. Las características de los instrumentos han de mantenerse invariantes tanto con relación a los grupos lingüísticos definidos dentro de la población evaluada como con relación al idioma en el cual se lleve a cabo la evaluación. Desde la psicometría se

ha desarrollado toda una metodología, tanto cuantitativa como cualitativa para el estudio de la equivalencia psicométrica que se ha centrado sobre todo en la adaptación de tests (Muñiz, Elosua, & Hambleton, 2013) y en la detección del funcionamiento diferencial del ítem (FDI; Dorans & Holland, 1992). El tercer elemento de la triada, el sesgo sociolingüístico estaría relacionado con el impacto que variables como el proceso de normalización lingüística y/o actitudes hacia las lenguas ejercerían sobre el proceso evaluativo (Elosua, 2016). Desde un punto de vista amplio del concepto de validez de las puntuaciones, los tres aspectos, el psicolingüístico, psicométrico y sociolingüístico impactan en el proceso evaluativo, y el significado de las puntuaciones puede verse comprometido si las competencias relacionadas con el dominio del lenguaje académico no alcanzan un nivel mínimo, el idioma no está normalizado y/o existe funcionamiento diferencial del ítem.

El análisis del sesgo psicolingüístico y psicométrico requiere considerar el efecto de recibir instrucción en un idioma diferente al idioma familiar. Es una cuestión abordada sobre todo para el caso de estudiantes migrantes pero que afecta también a los alumnos matriculados en un modelo lingüístico cuyo idioma vehicular no coincide con el idioma familiar. En Estados Unidos, por ejemplo, las investigaciones que comparan el desempeño entre estudiantes inmigrantes para los cuales el inglés es su segunda lengua (ELS; *English Learning Students*), y no-inmigrantes aprecian de forma unánime un rendimiento menor para los primeros (Abedi, 2009; Abedi & Lord, 2001; Solano-Flores & Trumbull, 2003) que en el caso de la competencia matemática se asocia al empleo de un nivel de complejidad lingüística irrelevante (sesgo psicométrico). Con relación a la equivalencia entre pruebas administradas en diferentes idiomas es común reportar funcionamiento diferencial del ítem que se acentúa cuando las lenguas son no-indoeuropeas (Grisay, de Jong, Gebhardt, Berezner, & Halleux-Monseur, 2007) y cuando el país objeto de estudio es multilingüe (Monseur & Halleux, 2009).

En entornos de contacto de lenguas más próximos, Pifarré, Sanuy, Huguet y Vendrell (2003) concluyen que los alumnos instruidos en catalán cuyo idioma familiar es el castellano obtienen en matemáticas promedios algo menores que aquellos compañeros que hablan catalán en sus hogares. Los mismos autores informan de una relación significativa entre el dominio del catalán y el rendimiento en matemáticas entre alumnos de segundo curso de educación secundaria obligatoria (Sanuy, Huguet, Pifarré & Vendrell, 2002). Se han conseguido resultados similares con relación al euskera; se aprecian diferencias en el desempeño en matemáticas entre estudiantes de cuarto de educación primaria cuyo idioma familiar es el castellano y reciben instrucción en euskera frente a alumnos para los cuales la lengua familiar y lengua de instrucción es el euskera (Elosua & Egaña, 2017). Entre las hipótesis que se formulan para explicar estos resultados se apela a la perspectiva psicolingüística relacionada con el proceso de adquisición de competencias lingüísticas (CALP).

En este contexto de investigación el objetivo de este trabajo es profundizar en las fuentes potenciales de sesgo que pueden afectar a los resultados de la evaluación educativa en contextos de diversidad lingüística. El estudio se centra en la Comunidad Autónoma del País Vasco (CAPV) y en la competencia matemática. Desde el año 1979 el euskera es, junto con el castellano, un idioma oficial. En 1983 el llamado decreto de bilingüismo diseñó tres programas de educación bilingüe conocidos como modelos A, B y D; cada uno incorpora una proporción cada vez mayor de euskera como lengua de instrucción. En los tres modelos, la lengua y literatura española, lengua y literatura vasca y lenguas modernas se enseñan principalmente en el idioma correspondiente. Los modelos se definieron de la siguiente manera:

- Modelo A: Todas las asignaturas, excepto el euskera, se imparten básicamente en español. El euskera se trata como cualquier otra asignatura con el número de horas de clase semanales estipuladas por el Departamento de Educación y Cultura.

- Modelo B: Tanto el español como el euskera se utilizan como lengua vehicular en otras asignaturas.
- Modelo D: Todas las materias, excepto la lengua española, se impartirán básicamente en euskera, siendo el euskera tratado de la misma manera que cualquier otra asignatura.

La elección de la competencia matemática no es arbitraria; Si bien es importante reconocer que el lenguaje es un vehículo de comunicación entre el evaluador y el evaluado, elegimos las matemáticas porque teóricamente tiene una carga lingüística menor y por tanto menor ruido en el estudio de las variables lingüísticas.

Con base a la literatura revisada, se plantean dos hipótesis de investigación:

- En los modelos en los que el euskera es vehicular, bien sea de forma parcial (modelo B) o total (modelo D), se espera que no exista relación significativa entre el idioma familiar y el idioma de instrucción; es decir, el desempeño en matemáticas será similar entre alumnos independientemente de su lengua materna (perspectiva psicolingüística).
- Dado que las competencias lingüísticas mínimas exigidas en la prueba de evaluación han sido adquiridas por todos los estudiantes de educación secundaria (homogeneidad lingüística básica), no se encontrará funcionamiento diferencial del ítem (perspectiva psicométrica).

Método

Muestra

La muestra está definida por los 2783 estudiantes de segundo curso de enseñanza secundaria matriculados en centros educativos de la comunidad autónoma del País Vasco que recibiendo fondos públicos ofrecen de forma simultánea los modelos lingüísticos B y D. Se trata de 45 centros, de los cuales 18 son concertados y 27 son públicos. Del total de estudiantes, 1341 son mujeres y 1442 son varones. Todos ellos son nativos de la Comunidad Autónoma del País Vasco, y su idioma familiar es el euskera (N = 624) o el

castellano (N = 2159). La distribución de estudiantes por modelo educativo indica que 854 están matriculados en el modelo B y el resto, 1929, estudian en el modelo D (ver tabla 1).

Instrumento

En el marco de las evaluaciones diagnósticas, el instituto vasco de evaluación educativa construyó un cuestionario de 24 ítems para la evaluación de la competencia matemática (ISEI-IVEI, 2012a). La competencia matemática se define como la habilidad para usar y relacionar los números, sus operaciones básicas, los símbolos y las formas de expresión y razonamiento matemático, tanto para producir e interpretar información, como para ampliar el conocimiento sobre aspectos cuantitativos y espaciales, y para resolver problemas relacionados con la vida cotidiana y con el mundo laboral. El curriculum para el desarrollo de la competencia matemática es el mismo para los estudiantes de los modelos B y D. La prueba fue administrada en euskera a todos los participantes.

Variables

Se especificaron dos tipos de variables; variables de nivel individual o variables de primer nivel y variables de centro o variables de segundo nivel.

Variables de nivel individual. Se definieron cinco variables predictoras en este nivel: índice socioeconómico y cultural (ISEC), sexo, idoneidad e idioma familiar.

- ISEC es un indicador compuesto construido sobre la información aportada por los padres de los alumnos participantes en la evaluación diagnóstica que se obtiene a partir de medidas relacionadas con el nivel educativo de los padres, el estatus social de sus profesiones, y los recursos familiares. El índice está medido en una escala estandarizada (M = 0, DT = 1).
- Sexo: es una variable con amplia tradición de uso en la evaluación educativa. Los estudios sobre competencia matemática y/o rendimiento matemático y sexo concluyen una relación significativa entre ambas

variables con promedios mayores para los varones que para las mujeres (OECD, 2016).

- **Idoneidad:** Refleja si el alumno está matriculado en el curso académico teóricamente adecuado por su año de nacimiento.
- **Idioma familiar.** El idioma familiar es el español o el euskera. La variable que es autoinformada por las familias, identifica el idioma utilizado regularmente en casa.

Variables de centro. Se delimitaron tres variables relacionadas con las características del centro educativo; titularidad, índice socioeconómico del centro (CentroISEC), y modelo lingüístico.

- **CentroISEC.** Sobre los valores individuales ISEC se construyó la variable agregada CentroISEC para cada uno de los centros participantes en esta investigación. Cabe esperar que a medida que se incrementen los recursos aportados por los estudiantes mejor será el rendimiento medio de los alumnos (Coleman et al., 1966; Willms, 2006).
- **Modelo lingüístico.** Los estudiantes dentro de un centro pueden recibir instrucción en el modelo lingüístico B o en el modelo lingüístico D.
- **Titularidad:** indica si el centro es público o concertado. 1725 alumnos estudian en centros públicos y 1058 en centros concertados.

Procedimiento y análisis

La administración de las pruebas se llevó a cabo por personal externo a los centros durante el horario lectivo. Al tratarse de una prueba diagnóstica los datos son censales.

Los análisis integran una perspectiva psicométrica con un acercamiento predictivo basado en el ajuste de modelos lineales mixtos (Goldstein, 2003). Dentro la primera aproximación se analizó la equivalencia de la prueba de competencia en función del idioma familiar y del modelo educativo. Para ello una vez estimados los coeficientes de consistencia interna, se evaluaron los porcentajes de varianza asociados con el primer componente

extraído de un análisis de componentes principales, para finalmente examinar la presencia de funcionamiento diferencial por medio de la aplicación de dos procedimientos diferentes: el estadístico Mantel-Haenszel y la estandarización. Para la clasificación de la severidad del funcionamiento diferencial se utilizó el criterio del *Educational Testing Service* que define tres niveles: A (severo), B (moderado) y C (intangible).

Para la modelización de la competencia, además de las variables utilizadas en toda evaluación como predictoras de la competencia (ISEC; sexo, titularidad del centro, idoneidad) se incluyeron variables lingüísticas de nivel individual (idioma familiar) y de segundo nivel (modelo educativo). Para todos los efectos fijos se utilizó una codificación *dummy*. Los niveles de referencia fueron respectivamente el castellano para el idioma familiar, el modelo B para el modelo lingüístico, la red pública frente a la concertada, mujer para la variable sexo y alumno no idóneo. Tras el ajuste del modelo mixto se estimaron las medias mínimo cuadráticas (*lsm*). Se utilizó la función *lmer* del paquete *lme4* para la estimación del modelo multinivel (Bates, Maechler, Bolker & Walker 2015).

Resultados

Análisis preliminares

En el proceso de mecanización de las respuestas no se diferenció entre valor faltante e ítem no alcanzado. Del total de la muestra 1597 estudiantes (57.02%) respondieron a todos los ítems de la prueba. El porcentaje de valores ausentes para cada uno de los primeros 22 ítems osciló entre el 1% del tercer ítem al 6% del ítem 19. Los dos últimos ítems de la prueba se asociaron con porcentajes de no-respuesta del 8% y del 20% respectivamente.

De la distribución de estudiantes por modelos educativos se desprende que las familias optan en mayor medida por el modelo D. El porcentaje de estudiantes con el euskera como lengua familiar es el 22.4% (N = 624), sin embargo, el número de alumnos matriculados en el modelo D, es superior a

esta cifra y alcanza el 30.71% (N = 855). La relación entre ambas variables no es independiente ($\chi^2(1) = 41.01$; $p < .01$). Del conjunto de alumnos cuya lengua familiar es el euskera un 79.80% (N = 498) elige el modelo D como modelo de instrucción; y este porcentaje es del 66.2% (N = 1431) para aquellos estudiantes que hablan castellano en sus hogares.

Excepto en la variable modelo educativo, en la cual no se han observado diferencias en la competencia matemática entre los modelos lingüísticos B y D (Tabla 1), en el resto de variables las diferencias son significativas (p

$< .05$). Los estudiantes provenientes de centros concertados obtienen una media aritmética mayor (MA = 13.34) que aquellos que estudian en centros públicos (MA = 12.72); el rendimiento de los varones (MA = 13.15) es más alto que el de las mujeres (MA = 12.74); y aquellos estudiantes cuyo idioma familiar es el euskera tiene un promedio mayor (MA = 14.13) que sus compañeros cuya lengua familiar es el castellano (MA = 12.62). La significación práctica de estas diferencias se evaluó con la d de Cohen, que mostró un tamaño del efecto mayor asociado con la variable idioma familiar.

Tabla 1 - Muestra, estadísticos descriptivos y diferencias en competencia matemática

	N	Media aritmética	Contraste	p	d de Cohen
Modelo educativo					
Modelo B	854	12.83 (5.0)			
Modelo D	1929	13.01 (4.64)	$t(1530) = -0.88$.37	-0.03
Lengua familiar					
Euskera	624	14.13 (4.69)			
Castellano	2159	12.62 (4.72)	$t(1016) = 7.09$	$< .01$	-0.32
Sexo					
Mujer	1341	12.74 (4.71)			
Varón	1442	13.15 (4.79)	$t(2772) = -2.28$.02	0.08
Red					
Pública	1725	12.72 (4.75)			
Concertada	1058	13.34 (4.73)	$t(2244) = -3.33$	$< .01$	-0.12

Nota: Entre paréntesis las desviaciones típicas y grados de libertad

Equivalencia psicométrica en función del idioma familiar

Los coeficientes de fiabilidad estimados por el procedimiento alfa de Cronbach sobre la matriz de correlaciones tetracóricas arrojó los valores de 0.87 y 0.88 para las muestras de estudiantes cuyo idioma familiar es el castellano y para aquellos que hablan euskera de forma regular y dominante. Para el primer grupo, el porcentaje de varianza asociado al primer componente es del 27%, y en el segundo grupo el valor estimado es 26.5%. En ambos casos las pruebas superan los criterios habituales para la determinación de unidimensionalidad.

Ninguno de los procedimientos de detección de funcionamiento diferencial concluyó la presencia de ítems que favorecieran de forma severa a alguno de los grupos (Tabla 2). Únicamente el ítem número

3 mostró un funcionamiento diferencial moderado, que favorecía al grupo de estudiantes cuyo idioma familiar es el euskera. Los índices de dificultad para este ítem fueron de .90 y .81 con valores más altos para los estudiantes que hablan euskera en su entorno familiar. Los índices de discriminación respectivamente fueron .23 y .39. Para una mejor apreciación del comportamiento de los ítems se calcularon las medias aritméticas condicionadas sobre los niveles de competencia (Figura 1). En el gráfico se aprecia que el ítem 3 apenas discrimina en uno de los grupos, y que las mayores diferencias se sitúan en los niveles bajos de competencia. Para el resto de los ítems las medias aritméticas de respuesta al ítem o índices de dificultad para cada nivel de competencia fueron similares con independencia del idioma familiar; se muestra un ítem como ejemplo de los patrones de respuestas observados.

Tabla 2 - Funcionamiento diferencial del ítem

Ítem	Idioma familiar				Modelo educativo			
	MH	Tipo	Std	Tipo	MH	Tipo	Std	Tipo
1	0.97	A	0.00	A	1.15	A	-0.03	A
2	1.09	A	-0.01	A	1.13	A	-0.02	A
3	1.62	B	-0.06	B	0.86	A	0.01	A
4	0.95	A	0.02	A	0.94	A	0.02	A
5	0.90	A	0.00	A	0.93	A	0.01	A
6	1.01	A	0.00	A	0.98	A	0.01	A
7	0.77	A	0.06	A	1.05	A	-0.02	A
8	0.91	A	0.04	A	0.83	A	0.03	A
9	1.00	A	0.01	A	0.96	A	0.01	A
10	0.98	A	0.00	A	0.96	A	0.01	A
11	0.89	A	0.03	A	1.15	A	-0.04	A
12	1.04	A	-0.00	A	1.09	A	-0.03	A
13	1.00	A	-0.01	A	0.97	A	0.00	A
14	0.88	A	0.03	A	1.20	A	-0.03	A
15	0.97	A	0.01	A	0.77	A	0.04	A
16	0.96	A	0.01	A	1.01	A	-0.00	A
17	0.97	A	0.01	A	1.15	A	-0.03	A
18	1.14	A	-0.02	A	1.00	A	0.00	A
19	0.94	A	0.01	A	1.12	A	-0.02	A
20	0.95	A	0.01	A	1.02	A	-0.00	A
21	1.24	A	-0.03	A	0.87	A	0.02	A
22	1.09	A	-0.01	A	1.09	A	-0.02	A
23	0.99	A	0.02	A	1.11	A	-0.03	A
24	0.98	A	0.01	A	1.15	A	-0.05	A

Nota: MH. Estadístico alfa de Mantel-Haenszel. Std=Estandarización.

Equivalencia psicométrica en función del modelo educativo

Los coeficientes de fiabilidad estimados por el procedimiento alfa de Cronbach sobre la matriz de correlaciones tetracóricas arrojó los valores de 0.89 y 0.87 para los estudiantes de los modelos B y D repectivamente. En el modelo B la varianza asociada al primer componente fue 28.5%, y en el modelo D el valor indicador de unidimensionalidad fue 26.6%.

El análisis del funcionamiento diferencial del ítem con relación al modelo lingüístico rechazó la presencia de FDI; ninguno de los ítems alcanzó los criterios especificados.

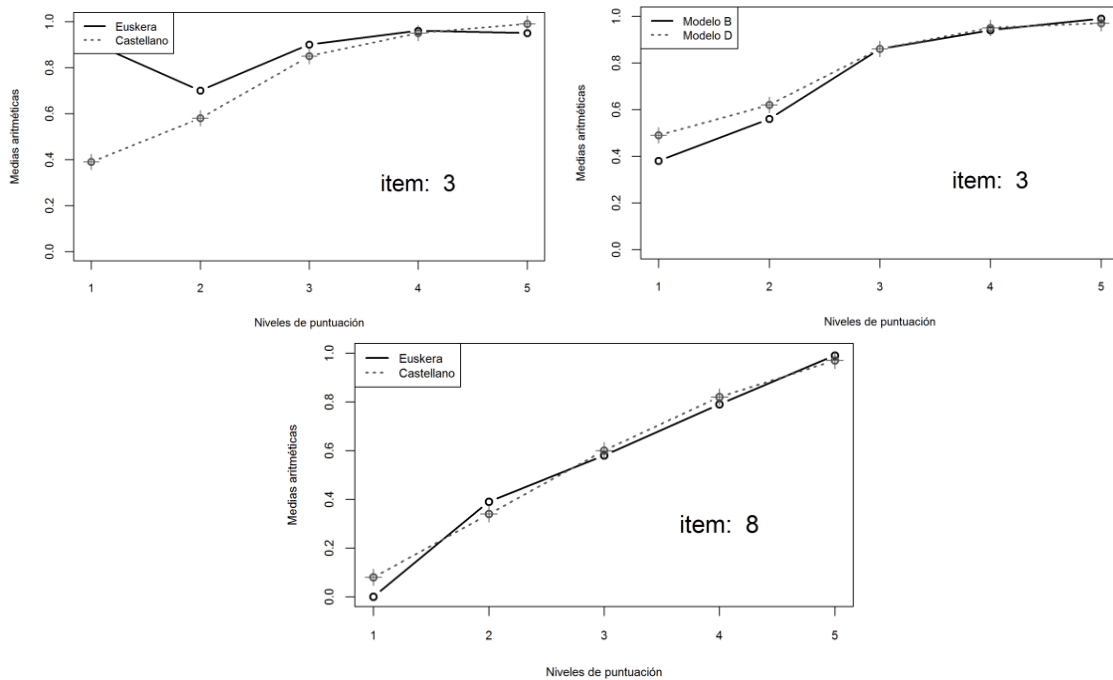


Figura 1 - Funcionamiento diferencial del ítem en función de la lengua familiar y del modelo educativo. Índices de dificultad condicionados.

Modelo predictivo multinivel

En primer lugar se estimó el modelo nulo, que arrojó una varianza residual de 20.51. El modelo que incluye como efectos fijos las variables sexo, ISEC, CentroISEC, tipo de centro, modelo lingüístico, idioma familiar e idoneidad

se estimó por el método de máxima verosimilitud restringida (REML) (Tabla 3). La correlación intraclase para este modelo fue .065, lo cual indica que el 6.5% de la variabilidad entre puntuaciones está asociada a las características de los centros.

Tabla 3 - Resumen del modelo predictivo: Coeficientes (β), Errores estándar ($Se(\beta)$), valores t y significación (p).

Efectos fijos	β	$Se(\beta)$	t	p
Intercepto	10.90	.43	25.29	<.01
ISEC	0.98	.09	10.46	<.01
Sexo	0.54	.07	3.34	<.01
Red educativa	-0.42	.42	-0.97	.33
Centro _{ESCS}	1.45	.57	2.55	.01
Modelo educativo	-0.39	.23	-1.67	.10
Idioma familiar	-0.50	.22	-2.23	.02
Idoneidad	3.04	.24	12.28	<.01
Efectos aleatorios		Varianza	DS	
Centro		1.21	1.10	
Residual		17.38	4.16	

Como era esperable, el índice socioeconómico y cultural (ISEC) mostró un efecto significativo tanto en

el nivel individual ($\beta = 0.98$; $Se = 0.09$; $t(2550) = 10.46$; $p < .01$), como en el nivel del centro ($\beta = 1.45$; $Se = 0.57$;

$t(42) = 2.55$; $p = .01$). El parámetro relacionado con el sexo también confirmó los resultados de investigaciones previas, que muestran el mejor desempeño de los varones con respecto a las mujeres ($\beta = 0.54$; $Se = 0.16$; $t(2550) = 3.34$; $p < .01$). La red educativa, concertada o pública, no mostró diferencias ($\beta = -0.42$; $Se = 0.42$; $t(42) = -0.97$; $p = .33$). En cuanto al idioma familiar los resultados evidenciaron diferencias en favor del grupo cuyo idioma familiar es el euskera ($\beta = -0.50$; $Se = 0.22$; $t(2550) = -2.23$; $p = .02$). Los efectos relacionados con el modelo lingüístico no fueron significativos ($\beta = -0.39$; $Se = 0.23$; $t(44) = -1.67$; $p = .10$), mostrando que las medias estimadas en competencia matemática son similares para los estudiantes matriculados en los modelos B y modelos D en aquellos centros que ofertan ambas opciones. Finalmente, los estudiantes que cursan en el nivel que les corresponde teóricamente por edad obtienen mejores promedios que sus compañeros ($\beta = 3.04$; $Se = 0.24$; $t(2550) = 12.28$; $p < .01$).

Tras el ajuste del modelo se estimaron las medias mínimo cuadráticas o medias ajustadas en

competencia matemática; es decir, los valores medios detrayendo los efectos fijos (Tabla 4). Las medias en función del modelo educativo ($M_{\text{modeloB}} = 12.32$; $M_{\text{modeloD}} = 11.93$; $t(36) = 1.67$; $p = .10$) o de la red educativa ($M_{\text{pública}} = 12.34$; $M_{\text{concertada}} = 11.92$; $t(39) = 0.97$; $p = .33$), no resultaron ser estadísticamente significativas. Con relación al idioma familiar las medias son superiores en el grupo de estudiantes cuyo idioma familiar coincide con el idioma de instrucción, el euskera ($M = 12.38$) sobre aquellos que combinan dos idiomas ($M = 11.88$) siendo la diferencia significativa con un nivel de confianza del 5% ($t(2358) = 2.23$; $p = .02$). La diferencia en competencia matemática entre sexos, estimada en -0.55 puntos favorece a los chicos ($t(2589) = -3.34$; $p < .01$).

De acuerdo a los resultados obtenidos, los niveles más altos de competencia estarían asociados con aquellos estudiantes varones provenientes de familias con altos niveles socioeconómicos cuya lengua familiar es el euskera y cursan el nivel que les corresponde por edad, independientemente de la red educativa o del modelo lingüístico.

Tabla 4 - Medias ajustadas en competencia matemática

	Media ajustada	Contraste (t)	p
Modelo educativo			
Modelo B	12.32 (.26)		
Modelo D	11.93 (.24)	1.67	.10
Lengua familiar			
Euskera	12.38 (.28)		
Castellano	11.88 (.22)	2.23	.02
Sexo			
Mujer	11.85 (.24)		
Varón	12.40 (.23)	-3.34	< .01
Red			
Pública	12.34 (.29)		
Concertada	11.92 (.33)	0.97	0.33

Nota: Entre paréntesis los errores estándar de las medias ajustadas.

Conclusiones

La evaluación educativa es una herramienta clave en la mejora de la calidad de la educación (Tiana-Ferrer, 2018). En cuanto instrumento de mejora, la evaluación educativa en contextos de diversidad lingüística ha de responder a cuestiones específicas derivadas de la pluralidad del universo de estudio. Reconociendo la diversidad lingüística como común y enriquecedora, la equidad en el proceso evaluativo exige la consideración de aspectos relacionados con las características psicométricas de las pruebas, el desarrollo de la competencia lingüística, o el contexto lingüístico familiar del alumnado. La comparabilidad entre puntuaciones y por ende la calidad del proceso evaluativo demanda el examen conjunto de distintas fuentes de variación.

Este trabajo intenta profundizar en el problema de la comparabilidad en estas circunstancias. Con referencia al idioma familiar y su impacto en los resultados de la evaluación, se ha concluido la equivalencia psicométrica de la prueba en función del idioma familiar y del modelo educativo. Únicamente un ítem ha mostrado un funcionamiento diferencial de carácter moderado que afecta a los niveles más bajos de competencia. Si bien la equivalencia no se ve afectada, convendría analizar el contenido del ítem y estudiar porque no discrimina entre alumnos cuya lengua familiar es el euskera. La conclusión de ausencia de sesgo psicométrico es importante puesto que permite avanzar en un estudio predictivo para el análisis de las fuentes de variabilidad en la competencia matemática. En caso de que se hubieran detectado ítems con funcionamiento diferencial habría que analizar el modo de equiparar las métricas (Elosua & Hambleton, 2018).

Con referencia al modelo predictivo, en tanto en cuanto se considera a los alumnos integrados en unidades mayores, los centros educativos, se ha utilizado un acercamiento metodológico que posibilita la estimación de efectos fijos y de efectos aleatorios. Los

resultados muestran que la varianza o variabilidad en la competencia matemática explicada por las características de los centros educativos, es únicamente del 6,5%. Un valor alto en este índice, el coeficiente de correlación intraclase, se interpretaría en términos de relación estrecha entre los centros y la competencia matemática, lo cual no sería deseable desde un punto de vista de equidad educativa. Sin embargo, es necesario matizar este dato. La variabilidad inter-centros estimada en la evaluación diagnóstica de la competencia matemática con la población completa de estudiantes matriculados en el País Vasco fue del 16,9% (ISEI-IVEI, 2012a). En este estudio el valor estimado se refiere a un subconjunto de centros del total de los 330 que forman parte de la red educativa vasca; únicamente aquellos que ofrecen dos modelos lingüísticos de forma concurrente, los modelos B y D.

Inciendo en el efecto de las características de los centros, es destacable el hecho de que si se comparan las medias aritméticas y las medias ajustadas referidas a la red educativa, la diferencia por titularidad se diluye. Es decir, atendiendo a los datos analizados, una vez detraído el efecto composicional no hay diferencias en el valor añadido de los centros. Esta circunstancia puede interpretarse admitiendo que la diferencia observada entre las medias aritméticas no proviene de la calidad de la oferta educativa de la red pública y de la red concertada pues los promedios de ambas se igualan al descontar los efectos de la matrícula; el origen de las diferencias pudiera situarse en el hecho de que a la hora de elegir centro/red, las familias con índices ISEC más altos prefieren una oferta distinta a la pública.

Las variables de primer nivel relacionadas con las características individuales o familiares corroboran: a) la relación entre el ISEC y la competencia matemática, b) el impacto de la no-idoneidad del alumno sobre el desempeño, c) las diferencias en competencia matemática en función del sexo, y d) el efecto negativo de la no concordancia entre el idioma familiar y la lengua vehicular. El ISEC, el sexo y la idoneidad han sido ampliamente debatidas en el seno de la evaluación educativa y los

resultados sobre sus efectos son concordantes (OECD, 2016). En cuanto al impacto del idioma familiar sobre el rendimiento académico en contextos de diversidad lingüística, los resultados contravienen la hipótesis de partida y refuerzan los estudios que muestran la relación entre ambas variables.

Los resultados apuntalan la necesidad de apoyar la adquisición de competencias lingüísticas de los alumnos cuyo idioma familiar es distinto del idioma vehicular; conclusión que se desprende también de las investigaciones llevadas a cabo en Cataluña (Sanuy, Huguet, Pifarré & Vendrell, 2002). La hipótesis básica que garantiza la equidad en el proceso evaluativo -a excepción de la evaluación de las competencias lingüísticas orales y afines- es que no existen diferencias entre los evaluados en el nivel lingüístico exigido para el óptimo desarrollo de la prueba. Este principio básico de homogeneidad lingüística mínima puede ser violado en contextos multilingües y en consecuencia convertirse en una fuente de varianza irrelevante.

El hecho de encontrar diferencias en el desempeño en matemáticas entre participantes de enseñanza secundaria obligatoria que se caracterizan por ser estudiantes nativos y cursar estudios en los modelos B y D, parece cuestionar el proceso de adquisición de CALP; pero tal y como señala Cummins (2000) el dominio académico del idioma no es equivalente al dominio de uso idioma, sino que hace referencia a la competencia para comprender y utilizar el tipo concreto de lenguaje en contextos educativos. Esta reflexión es el origen de un debate que deriva hacia el efecto de la lengua de administración de las pruebas (Abedi & Lord, 2001; Solano-Flores, 2016), y que parece concluir que en contextos multilingües el idioma que maximiza la ejecución es el idioma familiar (Elosua & Mujika, 2014; ISEI-IVEI, 2012b).

Extendiendo la cuestión de la equidad a aquellos centros que comparten más de un modelo lingüístico, los resultados han mostrado que no existen diferencias con

respecto a esta variable. Si bien el objetivo del estudio no ha sido la evaluación de los modelos lingüísticos, en el desarrollo de la investigación ha sido importante considerar su posible efecto sobre los resultados. Con relación a las características de la población estudiada – alumnos de segundo de enseñanza obligatoria matriculados en centros que ofrecen de forma conjunta los modelos B y D – el factor modelo educativo ha resultado no ser significativo. Esta conclusión es altamente satisfactoria desde la perspectiva de la equidad educativa. Sin embargo, es importante apuntar que sería interesante extender el estudio e incluir los centros que ofertan el modelo A, caracterizado por el uso del castellano en la enseñanza como lengua vehicular.

En definitiva, la comparabilidad entre puntuaciones en situaciones de diversidad lingüística exige la conjunción de áreas de conocimiento diferentes. La comparabilidad se torna aún más espinosa cuando un mismo sistema educativo dispone de distintos modelos o programas, que pueden además ser ofertados por un mismo centro. Es una cuestión con innumerables connotaciones políticas, educativas y sociales, abordado de forma esquivada y confusa por la bibliografía especializada. La puesta a disposición pública de las bases de datos educativas podría ayudar a profundizar en un tema tan controvertido como cardinal.

Referencias

- Abedi, J. (2009). Validity of Assessment for English Language Learning Students in a National/International Context. *Estudios sobre Educación*, 16, 167-183. doi: https://doi.org/10.1207/S15324818AME1403_2
- Abedi, J., & Lord, C. (2001). The language factor in mathematics. *Applied Measurement in Education*, 14, 219-234.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 671, 1-48. doi: <https://doi.org/10.18637/jss.v067.i01>

- Coleman, J., Campbell, E. Q., Hobson, C. J., McPartland, J., Mod A. M., Weinfeld, F. D., & York, R. L. (1996). *Equality of Educational Opportunity*. Washington, DC: US Department of Health, Education and Welfare.
- Cummins, J. (1981). Four misconceptions about language proficiency in bilingual education. *NABE Journal*, 5, 31-45. doi: <https://doi.org/10.1080/08855072.1981.10668409>
- Cummins, J. (2000). *Language, power and pedagogy: Bilingual children in the crossfire*. Clevedon, England: Multilingual Matters, Ltd. doi: <https://doi.org/10.1080/15235882.2001.10162800>
- Dorans, N.J., & Holland, P.W. (1992). *DIF detection and description: Mantel-Haenszel and standardization*. (RR-92-10). Princeton, NJ: Educational Testing Service.
- Elosua, P. (2016). Minority Language Revitalization and Educational Assessment. Do Language-related factors impact performance?. *Journal of Sociolinguistics*, 20, 212-228. doi: <https://doi.org/10.1111/josl.12176>
- Elosua, P., & Egaña, M. (2017). Socioeconomic patterns and educational performance based on language-of-instruction models in the Basque Autonomous Community. *Journal of Multilingual and Multicultural Development*, 38, 686-698. doi: <https://doi.org/10.1080/01434632.2016.1224239>
- Elosua, P., & Hambleton, R. K. (2018). Psychological and Educational Test Score Comparability Across Language and Cultural Groups in the Presence of Item Bias. *Revista de Psicología y Educación* 13, 23-32. doi: <https://doi.org/10.23923/rpye2018.01.155>
- Elosua, P., & Mujika, J. (2014). Impact of family language and testing language on reading performance in a bilingual educational context. *Psicothema*, 26, 328-335. doi: <https://doi.org/10.7334/psicothema2013.344>
- Grisay, A., de Jong, J.H.A.L., Gebhardt, E., Berezner, A., & Halleux-Monseur, B. (2007). Translation equivalence across PISA countries. *Journal of Applied Measurement*, 8, 249-266.
- Goldstein, H. 2003. *Multilevel Statistical Models*. Third Edition. London: Edward Arnold.
- ISEI-IVEI. (2012a). *Evaluación Diagnóstica. Informe de Resultados*. Bilbao: ISEI-IVEI.
- ISEI-IVEI. (2012b). *PISA-L*. Bilbao: ISEI-IVEI.
- Monseur, C., & Halleux, B. (2009). *Translation and verification outcomes: National versions quality*. In OECD technical report (pp. 96-104). Paris: Organisation for Economic Cooperation and Development.
- Muñiz, J., Elosua, P., & Hambleton, R. K. (2013). Directrices para la traducción y adaptación de tests: Segunda edición. *Psicothema*, 25, 149-155. doi: <https://doi.org/10.7334/psicothema2013.24>
- OECD. (2016). *PISA 2015 Results: Excellence and Equity in Education*. Paris: OECD Publishing, Paris. doi: <https://doi.org/10.1787/9789264267510-en>
- Pifarré, M., Sanuy, J., Huguet, A., & Vendrell, C. (2003). Rendimiento matemático en contextos bilingües. Análisis de la incidencia de algunas variables del contexto socio-educativo. *Revista de Investigación Educativa*, 21, 183-199.
- Sanuy, J., Huguet, A., Pifarré, M., & Vendrell, C. (2002). Relaciones entre el conocimiento del catalán, el castellano y el rendimiento matemático. Un estudio empírico. *Revista de Educación*, 328, 383-394.
- Solano-Flores, G. (2016). *Assessing English Language Learners*. New York: Routledge. doi: <https://doi.org/10.4324/9780203521953>
- Solano-Flores, G., & Trumbul, E. (2003). Examining language in context: The need for new research and practice paradigms in the testing of English Language Learners.

Educational Researcher, 32, 3-13. doi: <https://doi.org/10.3102/0013189X032002003>

Tiana-Ferrer, A. (2018). Treinta años de evaluación de centros educativos en España. *Educación XXI*, 21(2), 17-36.

UNESCO (2003). *Education in a multilingual world. UNESCO Education Position Paper*. Paris: UNESCO. Recuperado de <http://unesdoc.unesco.org/images/0012/001297/129728e.pdf>

Willms, J. (2006). *Learning Divides: Ten policy Questions About the Performance and Equity of Schools and Schooling Systems*. Montreal: UNESCO Institute for Statistics.

Zalbide, M., & Cenoz, J. (2008). Bilingual Education in the Basque Autonomous Community: Achievements and Challenges. *Language, Culture and Curriculum*, 21, 5-20. doi: <https://doi.org/10.2167/lcc339.0>

Agradecimientos

Trabajo financiado parcialmente por la UPV/EHU GIU 15/24, GIU18/88

Autores / Authors

To know more /
Saber más

Elosua, Paula (paula.elosua@ehu.es)

Doctora en Psicología en el Departamento de Psicología Social y Metodología de las Ciencias del Comportamiento de la Facultad de Psicología de la Universidad del País Vasco. Dirección postal: Facultad de Psicología. Avda. Tolosa, nº 70. 20018-San Sebastián (Guipúzcoa, España).


[0000-0002-9651-9100](https://orcid.org/0000-0002-9651-9100)



Revista ELección de Investigación y **E**valuación **E**ducativa
E-Journal of Educational Research, Assessment and Evaluation

[ISSN: 1134-4032]



Esta obra tiene [licencia de Creative Commons Reconocimiento-NoComercial 4.0 Internacional](https://creativecommons.org/licenses/by-nc/4.0/).
This work is under a [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by-nc/4.0/).