



Programa de Doctorado Formación en la Sociedad del
Conocimiento

Factores asociados al rendimiento y a la eficacia escolar: un estudio basado en métodos mixtos a partir de PISA 2015

Tesis Doctoral

Dña. Adriana Gamazo García

Directores

Dra. Dña. Susana Olmos Migueláñez
Dr. D. Fernando Martínez Abad

Salamanca, septiembre de 2019

D^a Susana Olmos-Migueláñez, Profesora Titular de Universidad del departamento de Didáctica, Organización y Métodos de Investigación de la Universidad de Salamanca, y D. Fernando Martínez Abad, Profesor Ayudante Doctor del departamento de Didáctica, Organización y Métodos de Investigación de la Universidad de Salamanca en calidad de directores del trabajo de tesis doctoral titulado “Factores asociados al rendimiento y a la eficacia escolar: un estudio basado en métodos mixtos a partir de PISA 2015” y realizado por Dña. Adriana Gamazo García

HACEN CONSTAR

Que dicho trabajo tiene suficientes méritos teóricos contrastados adecuadamente mediante las validaciones oportunas, publicaciones relacionadas y aportaciones novedosas. Por todo ello consideran que procede su defensa pública.

En Salamanca, a 9 de julio de 2019.

**Dr. D^a. Susana
Olmos Migueláñez**

**Dr. D. Fernando
Martínez Abad**



Programa de Doctorado Formación en la Sociedad del Conocimiento

Factores asociados al rendimiento y a la eficacia escolar: un estudio basado en métodos mixtos a partir de PISA 2015

Tesis Doctoral

Directores

Susana Olmos-Migueláñez

Fernando Martínez Abad

Doctoranda

Adriana Gamazo García

Septiembre, 2019

Agradecimientos

Esta tesis está dedicada a todas las personas que la han inspirado o que la han sufrido junto a mí, y en especial a aquellas que han hecho ambas cosas.

En primer lugar, me gustaría agradecer a mis directores, Susana y Fernando, por su guía y apoyo incondicionales, y por mostrarme a través del ejemplo cómo ser una investigadora inquisitiva, independiente y rigurosa.

A María José por su incansable trabajo, por haber creado un grupo de investigación del que me siento orgullosa de formar parte, y por haberme dado esta oportunidad para desarrollarme y crecer como investigadora.

A los demás compañeros del GE2O, del área de MIDE y de GRIAL, por haberme acompañado y apoyado a lo largo de este proceso.

A mis amigos y familia, que posiblemente no querrán volver a oír a hablar de PISA en lo que les queda de vida, por su apoyo, interés, palabras de ánimo y obligaciones de salir de casa, aunque fuese a rastras. A mis padres por inculcarme el amor por la lectura y el respeto por la ciencia, dos cualidades fundamentales para cualquier investigador.

A Alicia Chaparro, Joaquín Caso y el resto del equipo del IIDE en México, por haberme dado la oportunidad de participar en sus proyectos y conocer otras formas de trabajar.

Al Ministerio de Economía y Competitividad, que a través de la concesión de una ayuda para contratos predoctorales para la formación de doctores (BES-2016-077829) me ha permitido dedicarme en exclusiva a esta tarea.

Suele decirse que una tesis es un proceso solitario, pero en mi caso no ha sido así en absoluto. Mi último y más profundo agradecimiento es para José, que me ha inspirado y sufrido a partes iguales, se ha leído hasta la última coma varias veces, y ha contribuido con su visión y su esfuerzo a que esta tesis llegase a buen puerto. Es un verdadero privilegio caminar esta senda contigo. Gracias por compartir, aportar y comprender siempre.

ÍNDICE

Resumen	15
Abstract.....	17
Introducción.....	19
1. El Programa de Evaluación Internacional de los Alumnos.....	27
1.1. Evolución.....	29
1.1.1. Antecedentes: el proyecto INES y <i>Education at a Glance</i>	29
1.1.2. El origen de PISA.....	31
1.2. Estructura.....	32
1.2.1. Definición de las competencias.....	33
1.2.2. Cuestionarios de contexto	41
1.3. Características metodológicas	43
1.3.1. Diseño muestral.....	43
1.3.2. Modelos de Teoría de Respuesta al Ítem y generación de Valores Plausibles de rendimiento	46
1.4. Críticas a PISA	47
2. Usos de PISA en investigación educativa: una revisión sistemática de la literatura..	53
2.1. Metodología.....	56
2.1.1. Preguntas de investigación	56
2.1.2. Criterios selección	57
2.1.3. Cadena de búsqueda	60
2.1.4. Proceso de selección.....	60
2.2. Resultados.....	62
2.2.1. Resultados del estudio <i>mapping</i>	62
2.2.2. Resultados de la SLR	68
2.3. Conclusiones.....	72
3. La investigación en eficacia educativa	75
3.1. Evolución de la disciplina.....	79
3.1.1. Años 70: Los orígenes de la IEE.....	81
3.1.2. Años 80: Consolidación de la disciplina.....	85
3.1.3. Años 90: Los modelos integrados	86
3.1.4. Siglo XXI: Nuevos modelos y métodos	93
3.2. La IEE en España	97
3.2.1. Siglo XX.....	97
3.2.2. Siglo XXI	103
3.3. Lecciones aprendidas.....	109
3. Los modelos multinivel para el análisis de datos educativos	113
4.1. Supuestos	116
4.2. Especificación del modelo.....	117
4.2.1. El modelo nulo	117
4.2.2. El modelo condicional.....	118
4.3. Los efectos escolares y el valor añadido	119
4.4. Características metodológicas	123
4.4.1. Métodos de estimación.....	123
4.4.2. Efectos fijos y aleatorios	124
4.5. Ventajas, limitaciones y comparativa con otras técnicas	125
5. Método	129
5.1. Objetivos.....	132
5.2. Los métodos mixtos.....	132
5.3. Fase 1: estudio cuantitativo	134

5.3.1.	Objetivos	134
5.3.2.	Diseño.....	134
5.3.3.	Población y Muestra.....	134
5.3.4.	Instrumentos	139
5.3.5.	VARIABLES.....	139
5.3.6.	Procedimiento y análisis de datos	143
5.4.	Fase 2: estudio cualitativo	144
5.4.1.	Objetivos	145
5.4.2.	Enfoque	145
5.4.3.	Muestra.....	146
5.4.4.	Informantes.....	146
5.4.5.	Instrumentos	147
5.4.6.	Procedimiento.....	148
5.4.7.	Análisis de datos.....	149
6.	Resultados del análisis de datos cuantitativo.....	153
6.1.	Análisis de los efectos de las variables contextuales en el nivel competencial de los estudiantes españoles.....	155
6.1.1.	Consideraciones previas	156
6.1.2.	Competencia matemática	157
6.1.3.	Competencia lectora.....	165
6.1.4.	Competencia científica.....	174
6.2.	Selección de centros de alta y baja eficacia.....	183
6.3.	Análisis del efecto de las variables de proceso en la eficacia de los centros ...	186
7.	Resultados del análisis de datos cualitativo	191
7.1.	El sistema de categorías.....	194
7.2.	Descripción general de los centros	197
7.3.	Comparación de categorías.....	202
7.3.1.	Planes, proyectos y formación	202
7.3.2.	Metodologías	205
7.3.3.	Atención a la diversidad y seguimiento del alumnado.....	207
7.3.4.	Evaluación.....	210
7.3.5.	Gestión del tiempo	212
7.3.6.	Liderazgo y Equipo Directivo	212
7.3.7.	Gestión y organización del centro	214
7.3.8.	Clima y convivencia.....	216
7.3.9.	Familia y comunidad.....	219
7.4.	Resumen	220
8.	Discusión y conclusiones	225
8.1.	Equidad.....	228
8.1.1.	Nivel socioeconómico	229
8.1.2.	Género	230
8.1.3.	Estatus migratorio	233
8.1.4.	Repetición de curso	233
8.1.5.	Cambios de centro a lo largo del historial académico.....	235
8.1.6.	Limitaciones	235
8.1.7.	Líneas futuras de investigación	237
8.2.	Educación integral	238
8.2.1.	239
8.2.1.	Limitaciones	239
8.2.2.	Líneas futuras de investigación	240

8.3. Valor añadido	241
8.3.1. Limitaciones	245
8.3.2. Líneas futuras de investigación	246
8.4. Conclusiones.....	247
8.5. Publicaciones derivadas del trabajo de investigación.....	249
Referencias	253
Anexos.....	275
Anexo I. Publicaciones incluidas en la revisión sistemática de la literatura.....	277
Anexo II. Ítems que componen los índices utilizados en la regresión logística.....	283
Anexo III. Documentos presentados en los centros durante la fase cualitativa.....	291

ÍNDICE DE TABLAS

Tabla 1.1. Competencias principales y secundarias en PISA, por ciclo.....	32
Tabla 1.2. Contextos y temas de la prueba de competencia científica.....	34
Tabla 1.3. Relaciones entre los procesos y las habilidades matemáticas fundamentales.....	40
Tabla 2.1. Criterios de calidad.....	58
Tabla 2.2. Rúbrica para la evaluación de las preguntas de calidad.....	59
Tabla 2.3. Número de publicaciones por autor/a.....	64
Tabla 2.4. Fuente de las publicaciones.....	64
Tabla 2.5. Competencias analizadas como variable criterio en función del ciclo de PISA.....	66
Tabla 2.6. Número de publicaciones por ciclo de PISA.....	67
Tabla 2.7. Número de publicaciones por metodología empleada.....	68
Tabla 2.8. Resumen de las respuestas a las preguntas del <i>mapping</i>	68
Tabla 2.9. Variables predictoras a nivel de alumno más estudiadas.....	69
Tabla 2.10. Variables predictoras a nivel de centro más estudiadas.....	69
Tabla 2.11. Resumen de las respuestas a las preguntas de la SLR.....	72
Tabla 3.1. Características de los tipos de investigación en eficacia educativa.....	78
Tabla 3.2. Principales conclusiones de los estudios de los 70 agrupados por tema.....	84
Tabla 3.3. Factores escolares relacionados con variables de producto en los estudios neerlandeses.....	87
Tabla 3.4. Características de los estudios de IEE más relevantes del siglo XX en España.....	101
Tabla 3.5. Porcentaje de varianza del logro académico explicado por el aula, la escuela y el país para el rendimiento de matemáticas y lengua.....	104
Tabla 3.6. Resumen de resultados de los estudios de caso en el estudio iberoamericano.....	104
Tabla 3.7. Resumen de las buenas prácticas detectadas en el estudio de las escuelas de alta eficacia, primera fase del estudio UPV/ISEI.....	107
Tabla 3.8. Resumen de las prácticas educativas detectadas en los centros de alta y baja eficacia, segunda fase del estudio UPV/ISEI.....	108
Tabla 5.1. Número de estudiantes y centros participantes y porcentaje que representan de la muestra total, por Comunidad Autónoma.....	136
Tabla 5.2. Variables de contexto.....	140
Tabla 5.3. Características de los centros de la fase cualitativa.....	146
Tabla 5.4. Informantes de la fase cualitativa.....	147
Tabla 6.1. Correlaciones entre las covariables de nivel 1.....	156
Tabla 6.2. Correlaciones entre las covariables de nivel 2.....	156
Tabla 6.3. Componentes de varianza del modelo nulo para la competencia matemática.....	157
Tabla 6.4. Modelo de intercepto aleatorio y pendientes fijas para matemáticas.....	159
Tabla 6.5. Modelo de intercepto y pendientes aleatorias para matemáticas.....	160
Tabla 6.6. Diferencias entre los residuos de los centros en los modelos de pendientes fijas y aleatorias para matemáticas.....	161
Tabla 6.7. Modelo final para matemáticas, estimación robusta de los errores.....	161
Tabla 6.8. Componentes de varianza del modelo condicional para matemáticas.....	161
Tabla 6.9. Estadísticos descriptivos de los residuos de nivel 1 para matemáticas.....	162

Tabla 6.10. Estadísticos descriptivos de los residuos de nivel 2 para matemáticas.....	163
Tabla 6.11. Componentes de varianza del modelo nulo para lectura.....	166
Tabla 6.12. Modelo de intercepto aleatorio y pendientes fijas para lectura.....	167
Tabla 6.13. Modelo de intercepto y pendientes aleatorias para lectura.....	168
Tabla 6.14. Diferencias entre los residuos de los centros en los modelos de pendientes fijas y aleatorias para lectura.....	169
Tabla 6.15. Modelo final para lectura, con estimación robusta de los errores.....	169
Tabla 6.16. Componentes de varianza del modelo condicional para lectura.....	170
Tabla 6.17. Estadísticos descriptivos de los residuos de nivel 1 para lectura.....	170
Tabla 6.18. Estadísticos descriptivos de los residuos de nivel 2 para lectura.....	172
Tabla 6.19. Componentes de varianza del modelo nulo para ciencia.....	175
Tabla 6.20. Modelo de intercepto aleatorio y pendientes fijas para ciencia.....	176
Tabla 6.21. Modelo de intercepto y pendientes aleatorias para ciencia.....	177
Tabla 6.22. Diferencias entre los residuos de los centros en los modelos de pendientes fijas y aleatorias para lectura.....	178
Tabla 6.23. Modelo final para ciencia, con estimación robusta de los errores.....	178
Tabla 6.24. Componentes de varianza del modelo condicional para ciencia.....	179
Tabla 6.25. Estadísticos descriptivos de los residuos de nivel 1 para ciencia.....	179
Tabla 6.26. Estadísticos descriptivos de los residuos de nivel 2 para ciencia.....	181
Tabla 6.27. Estadísticos descriptivos de las variables escolares en función de su eficacia.....	184
Tabla 6.28. Distribución de los centros en función de su titularidad.....	185
Tabla 6.29. Regresión logística para variables de proceso de nivel 1 agregadas.....	187
Tabla 6.30. Precisión del modelo predictivo para variables de proceso de nivel 1.....	188
Tabla 7.1. Sistemas de categorías a priori y final para el análisis cualitativo.....	195
Tabla 7.1. Frecuencia de palabras más mencionadas por centro.....	198
Tabla 7.2. Porcentaje de codificación de categorías según los informantes.....	198
Tabla 7.3. Porcentaje de codificación de subcategorías según los informantes.....	200
Tabla 7.4. Resumen de los resultados de la fase cualitativa.....	221

ÍNDICE DE FIGURAS

Figura 1.1. Relaciones entre los cuatro elementos del dominio científico en PISA 2015.....	34
Figura 1.2. Relación entre el marco de la competencia lectora y la subescala de aspecto.....	38
Figura 1.3. Estructura modular de los cuestionarios de contexto en PISA 2015.....	42
Figura 1.4. Síntesis de las principales críticas a PISA.....	51
Figura 2.1. Proceso de selección.....	62
Figura 2.2. Número de publicaciones por año.....	63
Figura 2.3. Número de publicaciones por país.....	65
Figura 3.1. Hitos principales de la investigación en eficacia educativa.....	80
Figura 3.2. Modelo de eficacia escolar de Scheerens.....	89
Figura 3.3. Modelo QAIT/MACRO.....	90
Figura 3.4. Modelo de Creemers.....	92
Figura 3.5. Modelo dinámico de eficacia educativa. Adaptado de Creemers & Kyriakides.....	95
Figura 3.6. Modelo de eficacia de los centros de primaria, Muñoz-Repiso et al.....	99
Figura 3.7. Modelo Iberoamericano de Eficacia Escolar.....	105
Figura 4.1. Representación gráfica del valor añadido.....	120
Figura 4.2. Modelos para el cálculo de los efectos escolares.....	121
Figura 5.1. Resumen de las fases de la investigación.....	131
Figura 5.2. Enfoques de investigación.....	133
Figura 5.3. Características de la muestra de estudiantes.....	137
Figura 5.4. Distribución de la variable ESCS para la muestra española.....	138
Figura 5.5. Características de la muestra de centros.....	139
Figura 5.6. Procedimiento de análisis cualitativo de datos.....	150
Figura 6.1. Distribución de los residuos de nivel 1 de matemáticas.....	162

Figura 6.2. Comprobación del supuesto de homocedasticidad para los residuos de nivel 1 en matemáticas.....	163
Figura 6.3. Distribución de los residuos de nivel 2 de matemáticas.....	164
Figura 6.4. Gráfico Q-Q para la distribución de los residuos de nivel 2 de matemáticas.....	164
Figura 6.5. Comprobación del supuesto de homocedasticidad para los residuos de nivel 2 de matemáticas.....	165
Figura 6.6. Distribución de los residuos de nivel 1 de lectura.....	171
Figura 6.7. Gráfico Q-Q para los residuos de nivel 1 de lectura.....	171
Figura 6.8. Verificación del supuesto de homocedasticidad para los residuos de nivel 1 de lectura.....	172
Figura 6.9. Distribución de los residuos de nivel 2 de lectura.....	173
Figura 6.10. Gráfico Q-Q de los residuos de nivel 2 de lectura.....	173
Figura 6.11. Verificación del supuesto de homocedasticidad para los residuos de nivel 2 de matemáticas.....	174
Figura 6.12. Distribución de los residuos de nivel 1 para ciencia.....	180
Figura 6.13. Verificación del supuesto de homocedasticidad para los residuos de nivel 1 de ciencia.....	181
Figura 6.14. Distribución de los residuos de nivel 2 para ciencia.....	182
Figura 6.15. Gráfico Q-Q de los residuos de nivel 2 de ciencia.....	182
Figura 6.16. Verificación del supuesto de homocedasticidad para los residuos de nivel 2 de ciencia.....	183
Figura 6.17. Distribución de centros de alta y baja eficacia por Comunidad Autónoma.....	185
Figura 7.1. Nubes de palabras.....	196
Figura 7.2. Porcentaje dedicado a cada categoría en función del centro y los informantes.....	199
Figura 7.3. Porcentaje dedicado a cada subcategoría en función del centro.....	201
Figura 7.4. Similitud de codificación de los recursos en función del centro y el informante.....	202
Figura 8.1. Condicionantes del desempeño e intereses de las alumnas por materias STEM.....	232

Resumen

El estudio de la eficacia escolar tiene un largo recorrido dentro del campo de la investigación educativa, aunque en las últimas dos décadas se ha visto aumentado debido al surgimiento de las evaluaciones de rendimiento a gran escala como PISA. Estas evaluaciones proporcionan a los investigadores una gran cantidad de variables de contexto (p. ej. género, nivel socioeconómico, estatus migratorio), de proceso (p. ej. liderazgo, prácticas educativas, gestión del currículo) y de producto (rendimiento en diversas competencias básicas como lectura, matemáticas o ciencia), las cuales permiten una gran profundización en el estudio del rendimiento académico y sus determinantes.

Esta investigación hace uso de los métodos mixtos, concretamente de un método mixto secuencial explicativo con predominancia cuantitativa, para tratar de determinar dos cuestiones fundamentales. En primer lugar se lleva a cabo un análisis multinivel de las tres competencias principales evaluadas por PISA (lectura, matemáticas y ciencias) para estudiar el impacto de variables contextuales de estudiantes y de centros en el nivel competencial presentado por los alumnos participantes en la prueba. A partir de dichos análisis, se estima el nivel de eficacia de cada centro, definido como la diferencia entre la puntuación esperada en función de sus características contextuales y su puntuación real. Los centros se ordenan en función de esta variable para seleccionar aquellos con niveles más altos y bajos de eficacia, a partir de los cuales se realiza la siguiente fase de estudio. Dicha fase consiste en la estimación de la relación entre los factores de proceso de los estudiantes y los centros recogidos por los cuestionarios de contexto de PISA y el nivel de eficacia presentado por los centros. Por último, se desarrolla también un estudio cualitativo de casos en un centro de alta eficacia y uno de baja eficacia localizados en la Comunidad Autónoma de Castilla y León. A través de grupos de discusión con los docentes y equipos directivos de ambos centros se pretende profundizar en el estudio de los determinantes escolares de la eficacia.

Los resultados de la fase cuantitativa muestran la relevancia del impacto sobre el rendimiento que tienen factores como el género, el nivel socioeconómico (a nivel de estudiante y de centro), el ser inmigrante de 1ª generación, la repetición de curso, o los cambios de centro a lo largo del historial académico del alumno. El análisis cuantitativo de los factores de proceso relacionados con la eficacia arroja resultados inconcluyentes, ya que las variables escolares no muestran relaciones significativas con la eficacia, y las variables de los estudiantes presentan algunas relaciones incongruentes, como el comportamiento dispar de variables aparentemente relacionadas (adaptación a la instrucción y apoyo del profesorado o apoyo emocional familiar). Por último, los resultados del estudio de casos a través de técnicas cualitativas coinciden en algunos factores de eficacia (o ineficacia) con la literatura existente, aunque se considera necesaria la ampliación de la muestra de centros.

En la sección de discusión y conclusiones se exploran las implicaciones de los resultados tanto para los centros como para el sistema educativo, así como las limitaciones de los estudios y las futuras líneas de investigación planteadas, como el empleo de nuevas metodologías o la búsqueda de fuentes alternativas de datos.

Palabras clave: PISA, eficacia escolar, evaluación a gran escala, rendimiento académico, análisis multinivel, métodos mixtos

Abstract

The study of school effectiveness has a long history within the field of educational research, although it has been enhanced in the past two decades due to the emergence of large-scale performance assessments such as PISA. These assessments provide the researchers with lots of different variables, such as context (e.g. gender, socioeconomic level, migration status), process (e.g. leadership, educational practices, curriculum management), or product variables (reading, mathematics and science performance), which enable a deep research into academic performance and its determinants.

This research work uses mixed methods, specifically a sequential explanatory mixed method with quantitative predominance, in order to determine two key questions. Firstly, a multilevel analysis of the three competences measured by PISA (reading, math and science) is carried out in order to study the impact of student and school context variables on the performance of students. Based on these analyses, the level of effectiveness of each school is determined, and it is defined as the difference between the expected PISA score according to the school's context characteristics and its actual score. The schools are then ordered by this variable and the ones with the highest and lowest effectiveness scores are selected for the next study phase. This phase intends to find significant relationships between the process variables measured by PISA and school effectiveness. Lastly, a qualitative case study is conducted in a high-effectiveness and a low-effectiveness school within the Autonomous Community of Castilla y León. Through the use of discussion groups with teachers and management teams, we intend to delve into the study of the school-level determinants of effectiveness.

The results from the quantitative phase show the relevance of the impact on performance of factors such as gender, socioeconomic level (both at student and school level), being a 1st generation immigrant, grade repetition or changing schools. The quantitative analysis of the process factors related to the schools' level of effectiveness yields inconclusive results, given that the school-level variables do not show any significant relationships and the student-level variables show some incongruences, such as the disparate behavior of seemingly related variables (adaptation of instruction and teacher support or parental emotional support). Lastly, the qualitative case study yields some results that coincide with the existing literature, although there is a need to extend the school sample in order to obtain more significant results.

The discussion and conclusions section explores the implications of the results both for schools and for the education system, as well as the limitations of the studies and the future research lines proposed, such as the employment of new methodologies or the use of alternative data sources.

Keywords: PISA, school effectiveness, large-scale assessment, academic performance, multilevel analysis, mixed methods



Introducción

Introducción

El alcance de la influencia – o eficacia – de los centros educativos en el desarrollo de los estudiantes ha sido un tema de gran interés para la comunidad investigadora en educación en las últimas décadas. La investigación en eficacia escolar (IEE) surgió como disciplina a partir de la publicación del documento *Equality of Educational Opportunity Study*, coloquialmente conocido como el informe Coleman (Coleman, 1966), cuya principal conclusión fue que las inequidades encontradas en los resultados académicos estaban basadas en características socioeconómicas y demográficas de los estudiantes (raza, nivel socioeconómico familiar, etc.). El análisis realizado para el informe encontró que había pocas diferencias entre las escuelas, cuestionando la posible relevancia de las prácticas educativas llevadas a cabo en los centros para el rendimiento de los estudiantes.

A raíz de esta información surgieron numerosos estudios con un objetivo claro: demostrar que las escuelas sí importan y que pueden marcar la diferencia en el desarrollo académico y social de sus estudiantes (Chapman, Muijs, Reynolds, Sammons, & Teddlie, 2016), impulsando el desarrollo de una nueva línea de investigación para profundizar en el estudio de las prácticas educativas eficaces.

Así, a lo largo de las últimas décadas han surgido investigaciones centradas en medir el nivel de eficacia de los centros de la manera más rigurosa posible, a través del aislamiento o control de aquellas variables contextuales que afectan al rendimiento, pero escapan al control del centro, como el género, el nivel socioeconómico y el estatus migratorio de los estudiantes, o el tamaño y la titularidad del centro (Lizasoain et al., 2013; Milford, Ross & Anderson, 2010; Rodríguez-Santero & Gil-Flores, 2018; Scheerens, Luyten, van den Berg & Glas, 2015; Willms, 2010; Zhu, Kaiser & Cai, 2018). Algunos investigadores tratan también de analizar qué procesos escolares (p. ej. liderazgo, clima, metodologías docentes) se relacionan con la eficacia, y en la historia de la disciplina se han realizado de manera consistente algunas conclusiones sobre las características escolares que promueven unos resultados superiores a los que cabría esperar en función de las características socioeconómicas de su alumnado (ver capítulo 3). Algunas de estas características son un buen clima social y de disciplina, un profesorado con altas expectativas con respecto al rendimiento de los estudiantes, un seguimiento del rendimiento de los estudiantes planificado y basado en criterios

consensuados, un liderazgo comprometido y con funciones definidas, o una alta implicación del profesorado en los proyectos educativos y la formación docente.

Los instrumentos, variables y metodologías empleados para estos estudios han ido cambiando a lo largo del tiempo. Sin embargo, la proliferación en las dos últimas décadas de datos procedentes de evaluaciones a gran escala ha proporcionado a la comunidad investigadora la oportunidad de trabajar con grandes muestras, instrumentos estandarizados y un gran abanico de variables de estudio. Este hecho ha contribuido a la aparición de numerosos trabajos centrados en el estudio del nivel competencial de los estudiantes en matemáticas, ciencias o comprensión lectora, y su relación con variables de contexto (p. ej. género, nivel socioeconómico, estatus migratorio) y de proceso (p. ej. liderazgo, metodologías docentes, clima), así como de la eficacia de los centros educativos.

Además, las especificaciones metodológicas de estas evaluaciones a gran escala han permitido avances significativos en el tipo de análisis de datos realizados. Así, tras el surgimiento de la metodología multinivel en la década de los 80, el siglo XXI trajo consigo nuevas técnicas para el análisis de datos escolares, como los modelos de ecuaciones estructurales (SEM). Aunque el SEM no fue desarrollado específicamente para el análisis de datos educativos, su naturaleza como extensión de la regresión lineal múltiple permite a los investigadores probar el ajuste de un modelo especificado por ellos, así como estudiar las relaciones entre las distintas variables que lo componen (Chapman et al., 2016). Otra técnica de uso reciente, cuyo surgimiento en educación está íntimamente relacionado con la disponibilidad de datos procedentes de evaluaciones a gran escala, es la minería de datos. Esta técnica permite la extracción de patrones o relaciones significativas a partir de bases de datos que contienen gran cantidad de información (Castro y Lizasoain, 2010; Martínez-Abad y Chaparro, 2017).

Las investigaciones más recientes realizadas en nuestro país son un claro ejemplo de este avance en sus planteamientos metodológicos, ya que presentan enfoques rigurosos y adecuados a los objetivos de estudio, y procedimientos sistematizados de análisis de datos. En España destacan algunos grupos de investigación por su interés y trabajo sobre el tema de la eficacia escolar, localizados fundamentalmente en la Universidad del País

Vasco (Evaluación de Programas, Centros y Sistemas Educativos, HEPE¹), la Universidad Autónoma de Madrid (Grupo de Investigación Cambio Educativo para la Justicia Social, GICE²), la Universidad Complutense de Madrid (Medida y Evaluación de Sistemas Educativos³), la Universidad de Sevilla (Métodos de Investigación y Diagnóstico en Orientación⁴) y la Universidad de Salamanca (Grupo de Evaluación Educativa y Orientación, GE2O⁵). Dichos grupos han invertido en los últimos años esfuerzos por esclarecer cuáles son los factores relacionados con el rendimiento de los estudiantes y con la eficacia de los centros, con el objetivo de emitir recomendaciones para la práctica y política educativas, y sus investigaciones serán examinadas a lo largo de este trabajo.

Marco de la investigación

La presente tesis doctoral ha sido desarrollada dentro del Programa de Doctorado Formación en la Sociedad del Conocimiento, surgido en el Instituto Universitario de Ciencias de la Educación (IUCE) de la Universidad de Salamanca.

Este programa adopta un enfoque eminentemente multidisciplinar, buscando la promoción de la Sociedad del Conocimiento a través del desarrollo de los procesos de enseñanza-aprendizaje. Dicho enfoque se apoya en los diversos Grupos de Investigación Reconocidos de la Universidad de Salamanca que conforman el programa, como GRIAL, OCA, VISUALMED, Robótica y Sociedad y E-LECTRA.

De las múltiples líneas de investigación que surgen del programa de doctorado (p. ej. Interacción y eLearning, Investigación-Innovación en Tecnología Educativa, Educación y Sociedad de la Información) este trabajo se enmarca en la destinada al estudio de la Evaluación Educativa y Orientación. Esta línea está apoyada fundamentalmente por los miembros del Grupo de Evaluación Educativa y Orientación (GE2O)⁶, el cual ha desarrollado su trabajo investigador a través de múltiples Proyectos I+D nacionales, concatenados desde el año 2006 hasta la actualidad:

¹ <https://www.science.eus/es/groups/hepe-evaluación-de-programas-centros-y-sistemas-educativos>

² <http://www.gice-uam.es/>

³ <https://www.ucm.es/grupos/grupo/361>

⁴ <http://grupo.us.es/mido/>

⁵ <http://ge2o.usal.es/>

⁶ Integrante del Grupo Consolidado Grupo de Investigación en Interacción y eLearning (GRIAL) y del Instituto de Investigación Instituto Universitario de Ciencias de la Educación (IUCE) de la USAL

- 2006: *Modelo de evaluación y desarrollo de competencias clave implícitas en la ESO: TIC, Lectura y Convivencia Escolar (E-TELECO)* (SEJ2006-10700)
- 2009: *Evaluación de competencias clave y formación de profesorado de educación secundaria: TIC, ALFIN y convivencia escolar (EF-TALCO)* (EDU2009-08753)
- 2012: *Evaluación, Formación e Innovación sobre competencias clave en Educación Secundaria: TIC, Competencia Informacional y Resolución de Conflictos (EFI-CINCO)* (EDU2012-34000)
- 2015: *Evaluación de Impacto del Desarrollo de Competencias Básicas sobre el Rendimiento Académico en Educación Secundaria: Propuesta de Formación e Innovación Docente (EFI-4)* (EDU2015-64524-P)
- 2018: *School Effectiveness for the Improvement of the System (EFI-6)* (PGC2018-099174-B-I00, pendiente de resolución definitiva)

A lo largo de esta evolución, el grupo se ha centrado en diversos aspectos de las competencias clave de los estudiantes de secundaria, siempre con el objetivo de evaluarlas, llevar a cabo acciones de formación del profesorado y propiciar la innovación de la docencia de dichas competencias. Así, desde la evaluación de autopercepción de competencias clave en estudiantes de educación secundaria y la formación del profesorado en estas competencias, el Grupo ha evolucionado hacia el diseño de instrumentos de evaluación objetivos y la formación e innovación en centros educativos, así como la relación de dichas competencias con el rendimiento y la eficacia escolar.

En este último tópico, el de la eficacia escolar, es en el que se desarrolla la presente tesis doctoral, la cual ha sido financiada a través de una ayuda para contratos predoctorales para la formación de doctores (BES-2016-077829) concedida por el Ministerio de Economía y Competitividad en el año 2017 y asociada al Proyecto I+D correspondiente a 2015 (EDU2015-64524-P).

Para concluir esta contextualización del trabajo, cabe mencionar que, con el objetivo de complementar la formación predoctoral, se realizó una estancia de tres meses en el Instituto de Investigación y Desarrollo Educativo (IIDE) de la Universidad Autónoma de Baja California en México. Esta estancia se desarrolló bajo la tutela de Alicia Alelí Chaparro Caso Díaz y Joaquín Caso Niebla, ambos investigadores de referencia en el campo de la eficacia escolar en su país (Caso-Niebla & Hernández-Guzmán, 2010;

Chaparro, González Barbera & Caso Niebla, 2016; González Barbera, Caso Niebla, Díaz López & López Ortega, 2012; Martínez Abad & Chaparro, 2017).

Objetivos, procedimiento y estructura del documento

El objetivo de este trabajo de investigación, por tanto, es detectar aquellas prácticas educativas que favorezcan o dificulten la eficacia de un centro educativo a la hora de promover el rendimiento de sus estudiantes. Para ello, se plantea un procedimiento de métodos mixtos de carácter secuencial con predominancia cuantitativa, el cual se resume a continuación:

- Estudio de la relación de las variables contextuales de estudiantes y centros en las tres competencias evaluadas por PISA a través de la aplicación de modelos multinivel.
- Detección de centros cuyas puntuaciones reales se muestran muy por encima (o por debajo) de lo esperable en función de sus características contextuales, es decir, centros de alta y baja eficacia.
- Estudio de la relación de diversas variables de proceso escolares con el nivel de eficacia de los centros a través de un análisis de regresión logística.
- Estudio de la relación de las variables de proceso escolares con la eficacia de los centros a través de un estudio cualitativo de casos en dos centros identificados como de alta y baja eficacia en la Comunidad Autónoma de Castilla y León.

En cuanto a la estructura del presente documento, el marco teórico de esta tesis recoge y analiza todas aquellas referencias necesarias para dar sentido tanto a los instrumentos y técnicas utilizados como a la disciplina investigadora en la que se enmarca la investigación. Para ello, en el primer capítulo se presenta la evaluación PISA a través de una exploración de sus antecedentes históricos, la definición que se hace de las competencias en su marco analítico y sus características metodológicas. Asimismo, se recogen también algunas de las críticas a esta evaluación más comúnmente encontradas en la literatura.

Con el objetivo de realizar un repaso exhaustivo de los estudios que hacen uso de datos de PISA para la investigación educativa, en el capítulo dos se presenta una revisión sistemática de la literatura que analiza 52 estudios secundarios de PISA realizados entre los años 2005 y 2018. Este capítulo ofrece datos sobre los autores y las revistas científicas prominentes en este campo de investigación, o los países, las competencias y las variables

más frecuentemente analizadas. Además, se recogen también cuestiones metodológicas como el tipo de análisis estadístico realizado, o el tratamiento que se hace de las cuestiones técnicas de PISA relacionadas con su diseño complejo.

El tercer capítulo realiza un repaso de la investigación en eficacia escolar a través de un estudio de sus principales etapas y contribuciones más significativas a lo largo de sus cinco décadas de existencia. Se exploran las conclusiones derivadas de los estudios más relevantes de este campo, y se presentan los modelos integrados de la eficacia que más influencia han tenido en el desarrollo de la disciplina.

En el capítulo cuatro se realiza una introducción al análisis de datos multinivel, desarrollando los principales supuestos y especificaciones de los modelos multinivel, así como los diversos tipos existentes (transversales, de ganancia, de valor añadido, etc.).

Por su parte, el marco empírico comienza en el capítulo quinto con la descripción del método, en el cual se explican las muestras, instrumentos, variables y análisis de datos conducidos en los estudios cuantitativos y cualitativos que componen esta tesis.

El capítulo sexto expone los resultados de los análisis cuantitativos. Este capítulo comienza con la construcción de los modelos multinivel para la exploración de las relaciones entre las variables contextuales y las tres competencias evaluadas por PISA, los cuales derivan en la selección y caracterización de los centros de alta y baja eficacia identificados en España. A partir de esta selección, se presentan los resultados de la regresión logística realizada con las variables de proceso de centros y estudiantes proporcionadas por PISA.

A continuación, en el séptimo capítulo se resumen los hallazgos de la fase cualitativa, presentando los resultados obtenidos en los estudios de caso de los dos centros seleccionados (alta y baja eficacia) en función de las categorías de codificación de la información elegidas.

Por último, el capítulo octavo presenta una discusión sobre la relación de los resultados obtenidos con la literatura existente, así como de sus implicaciones para la práctica y política educativa.



1. El Programa para la Evaluación Internacional de los Alumnos

El Programa Internacional de Evaluación de los Alumnos

El Programa Internacional de Evaluación de los Alumnos (Programme for International Student Assessment, PISA) es una iniciativa impulsada por la OCDE que, desde su primera aplicación en el año 2000, explora el nivel de los estudiantes de 15 años en tres competencias clave: lectura, matemáticas y ciencias⁷. A diferencia de otras evaluaciones internacionales, como TIMSS o PIRLS (Asociación para la Evaluación del Rendimiento Educativo, IEA), que examinan los conocimientos curriculares de los estudiantes, PISA se centra en el desarrollo de aquellas competencias que la OCDE considera necesarias para que los jóvenes que terminan la educación obligatoria puedan participar de manera plena en la sociedad (OECD, 2017a).

En este capítulo se hace un breve repaso a la evolución de este programa de evaluación, se describen sus principales componentes (pruebas competenciales y cuestionarios de contexto), así como los procesos de muestreo y validación de las escalas, y por último se abordan algunas de las principales críticas que se han hecho de esta prueba desde la comunidad de investigadores educativos.

1.1. EVOLUCIÓN

El programa PISA es un exponente claro del trabajo de la OCDE en materia de evaluación y comparación de sistemas educativos. Sin embargo, no es el primer proyecto relacionado con indicadores internacionales de la educación emprendido por esta organización. A continuación, se expone un resumen de la actividad relacionada con la información educativa promovida por la OCDE desde su nacimiento en los años 60 hasta la creación de PISA a finales de los años 90.

1.1.1. Antecedentes: el proyecto INES y *Education at a Glance*

La OCDE fue fundada en 1961 con un claro interés en el desarrollo económico, no encontrándose en sus documentos fundacionales ninguna referencia explícita a la educación (Papadopoulos, 1994). No obstante, su preocupación por la educación y sus

⁷ En cada ciclo de aplicación de PISA, además de estas tres competencias obligatorias, se incluye la evaluación de otras competencias (resolución colaborativa de problemas o competencia financiera, por ejemplo) cuya aplicación es voluntaria para los países participantes.

políticas como posible motor de las economías de las naciones participantes se evidencia ya en 1968 con la creación del Centro para la Investigación y la Innovación Educativa (*Centre for Educational Research and Innovation, CERI*), en el seno del cual surgirían los proyectos descritos a continuación.

Una de las primeras iniciativas relacionadas con la recogida de datos de los sistemas educativos nacionales fue el llamado Libro Verde (OECD, 1967), un manual sobre los factores de inversión eficaz en educación que guiaría las primeras publicaciones sobre indicadores en los años 70 y principios de los 80 (Martens & Balzer, 2007; Papadopoulos, 1994). Sin embargo, a mediados de la década de los 80, las presiones de varios países miembros para generar un sistema de indicadores educativos más fiables y basados en la estadística dieron como resultado la creación del proyecto de Indicadores Internacionales de los Sistemas Educativos (*International Indicators of Educational Systems, INES*), y poco después a la cristalización de la publicación periódica *Education at a Glance*. Lo que hasta ese momento habían sido publicaciones esporádicas centradas fundamentalmente en datos de “entrada” de los sistemas educativos (gasto en educación, tasas de escolarización, cifras de personal docente, etc.) se configuró como una publicación periódica con un complejo sistema de indicadores que proporcionaban información centrada en la eficiencia de los sistemas (Martens, 2007).

El primer informe de *Education at a Glance* fue publicado en 1992 (OECD, 1992), y contenía indicadores relacionados con tres grandes áreas: costes, recursos y procesos escolares (gasto en educación, recursos humanos, participación en la educación, características de la toma de decisiones educativas), contextos de la educación (características demográficas, sociales y económicas de la población estudiantil) y resultados de la educación (rendimiento en lectura, matemáticas y ciencias, número de estudiantes titulados por nivel educativo e indicadores de desempleo). La información presentada en estos indicadores procedía en su inicio de otras fuentes primarias de datos, como datos administrativos de los países participantes u otras encuestas internacionales (OCDE, 1993). De hecho, hasta la creación de PISA los datos sobre el rendimiento de los estudiantes en las tres competencias de interés provenían de evaluaciones como el estudio de la competencia lectora llevado a cabo por la Asociación Internacional para la Evaluación del Rendimiento Educativo (IEA) en 1992 o el estudio *International Evaluation of Educational Progress* (IEAP II), que evaluó la competencia matemática y científica de los estudiantes de diferentes países (OCDE, 1993).

Aunque *Education at a Glance* sigue publicándose actualmente con una periodicidad anual, algunos autores señalan este estudio como uno de los principales precursores de PISA (Addey, 2017; Martens, 2007; Trohler, 2007), ya que además del objetivo común de proporcionar a la comunidad global una medida de la calidad de los sistemas educativos, comparte también la voluntad de informar las políticas educativas de los países participantes (Lingard, Martino & Rezai-Rashti, 2013; Mundy, 2007)

1.1.2.El origen de PISA

A finales de los años 90, y con un voluminoso cuerpo de trabajo sobre indicadores educativos a sus espaldas, la OCDE creó el Programa Internacional de Evaluación de los Estudiantes (*Programme for International Student Assessment*, PISA) con el objetivo de “determinar si el conocimiento y las competencias que [los estudiantes] han adquirido en la infancia los han preparado adecuadamente para la vida adulta” (OECD, 1999, P. 16). Este proyecto permitió dejar de depender de fuentes externas de datos y constituyó efectivamente un mecanismo propio para generar datos estadísticos comparables sobre la calidad y la eficacia de los sistemas educativos (Martens, 2007). Asimismo, también continuó con la tendencia a centrarse cada vez más en indicadores de resultado que permitiesen extraer conclusiones sobre la eficacia y equidad de los sistemas educativos, en este caso a través de la creación de pruebas propias para medir el nivel competencial de los estudiantes en lectura, matemáticas y ciencias.

La responsabilidad del desarrollo de la prueba y el diseño de sus características metodológicas recayó en una Junta de Países Participantes, con representantes de todas las naciones que, entre 1998 y 2000, elaboraron los instrumentos piloto y realizaron pruebas de campo para comprobar su funcionamiento (Adams & Wu, 2003). En 1999 se publicó el documento fundacional del programa, *Measuring Student Knowledge and Skills – A New Framework for Assessment* (OECD, 1999), y en el año 2000 se realizó el primer ciclo de aplicación de la prueba, con la competencia lectora como competencia principal y matemáticas y ciencias como competencias secundarias (tabla 1.1). Además de las pruebas competenciales, se propuso recoger otro tipo de indicadores de resultado educativo que la OCDE calificó como fundamentales para evaluar los sistemas educativos (OECD, 2017a), los llamados resultados no cognitivos, centrados en cuestiones como la motivación, la percepción de la autoeficacia, las actitudes hacia la educación, etc.

Tabla 1.1. Competencias principales y secundarias en PISA, por ciclo.

Ciclo	Competencia principal	Competencias secundarias
2000, 2009 y 2018	Lectura	Matemáticas y ciencias
2003 y 2012	Matemáticas	Lectura y ciencias
2006 y 2015	Ciencias	Lectura y matemáticas

Por otro lado, como se verá en las siguientes secciones del presente capítulo, la información sobre resultados cognitivos y no cognitivos se complementó con la recogida de datos sobre una gran cantidad de variables de contexto socioeconómico y demográfico tanto de los centros como de los estudiantes, así como de variables sobre los procesos educativos, las relaciones entre los miembros de la comunidad educativa, y la organización y gobernabilidad de los centros.

1.2. ESTRUCTURA

La evaluación PISA se compone de dos bloques principales de pruebas: las de nivel competencial y los cuestionarios de contexto.

Las pruebas competenciales de lectura, matemáticas y ciencias se han aplicado en todos los ciclos desde el inicio de PISA en el año 2000, aunque en diferentes proporciones en función de la competencia principal de cada uno de los ciclos. Comenzando por la competencia lectora en el primer ciclo, cada una de las aplicaciones de PISA se ha centrado en una competencia concreta, la cual ocupa aproximadamente dos tercios del total de los ítems (OECD, 2017a), dejando el tercio restante para repartir entre las otras dos competencias. Debido al gran número de ítems generados para cada uno de los dominios evaluados, no todos los estudiantes contestan a todas las preguntas, sino que se genera una serie de cuadernillos compuestos por diferentes ítems. En 2015, los estudiantes dispusieron de dos horas para completar la evaluación, de las cuales una estuvo destinada a ítems relacionados con la competencia científica, y la otra a uno de los otros dominios evaluados (lectura, matemáticas, o resolución colaborativa de problemas en el caso de los países que decidieron participar en su evaluación).

Las pruebas de dominio están compuestas por distintos tipos de ítems en función de la respuesta que requieren: respuesta múltiple, cerrada o abierta.

Por otro lado, los cuestionarios de contexto están destinados a recoger información sobre los llamados resultados no cognitivos (motivación, actitudes hacia el aprendizaje), condiciones individuales (entorno socioeconómico y educativo), y las características

estructurales y procesuales del entorno institucional (recursos y procesos del centro educativo).

1.2.1. Definición de las competencias⁸

1.2.1.1. Competencia científica

La competencia científica fue el dominio principal del ciclo 2015 de PISA, lo cual significa que los ítems relacionados con esta habilidad ocuparon la mayor parte tanto de las pruebas competenciales como de los resultados no cognitivos de los cuestionarios de contexto.

Como resultado de años de trabajo sobre esta competencia, así como del consenso alcanzado entre los educadores de materias relacionadas con la ciencia, la competencia científica se define en el marco de PISA como

la habilidad para enfrentarse a cuestiones relacionadas con la ciencia como un ciudadano reflexivo. Una persona capacitada para la ciencia está dispuesta a participar en un discurso razonado sobre ciencia y tecnología, lo cual requiere habilidades para:

- Explicar fenómenos de manera científica: reconocer, ofrecer y evaluar explicaciones para una variedad de fenómenos naturales y tecnológicos.
- Evaluar y diseñar investigaciones científicas: describir y valorar investigaciones científicas y proponer maneras de abordar las preguntas desde un punto de vista científico.
- Interpretar datos y evidencias de manera científica: analizar y evaluar datos, afirmaciones y argumentos presentados en diversas representaciones y extraer conclusiones científicas apropiadas. (OCDE, 2017a, p.22)

⁸ Dado que el objetivo de esta sección es describir en profundidad las competencias y los mecanismos del ciclo de PISA correspondiente a 2015, las fuentes de referencia utilizadas son reducidas, centrándose fundamentalmente en los documentos *Analytical and Assessment Framework* (OECD, 2017a) y *Technical Report* (OECD, 2017b). Por lo tanto, a no ser que la información presentada provenga de fuentes distintas a estas, o que se trate de una cita textual, figura o tabla tomada de dichos documentos sin realizar modificaciones, no se citarán a lo largo del presente capítulo para evitar reiteraciones que puedan dificultar su lectura.

El dominio de la competencia científica se organiza a través de diversos elementos que ayudan a categorizar los ítems planteados a los participantes de la prueba. Estos elementos, cuyas relaciones se encuentran representadas en la figura 1.1, son los contextos, los conocimientos, las competencias y las actitudes.

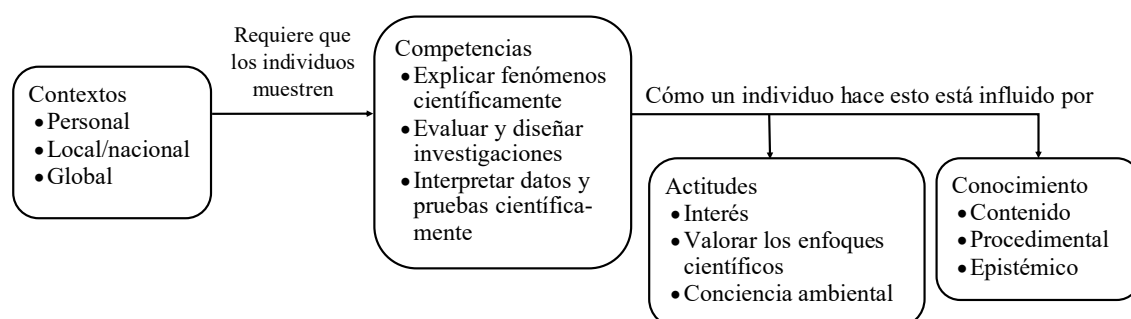


Figura 1.1. Relaciones entre los cuatro elementos que componen el dominio científico en PISA 2015. Adaptado de OECD, 2017a.

Los **contextos** se refieren al ámbito de ocurrencia de las distintas situaciones científicas presentadas en la prueba. Dichos ámbitos no se circunscriben al currículo de educación científica de los países participantes, sino que describen situaciones que pueden enmarcarse en el plano personal, local, nacional o global. Como puede observarse en la tabla 1.2, los contextos pueden concretarse en diversos temas de actualidad científica, como la salud o el medio ambiente.

Tabla 1.2. Contextos y temas de la prueba de competencia científica

Tema	Personal	Local/Nacional	Global
Salud y enfermedad	Mantenimiento de la salud, accidentes, nutrición	Control de las enfermedades, transmisión social, elecciones alimenticias, salud comunitaria	Epidemias, propagación de enfermedades infecciosas
Recursos naturales	Consumo personal de materiales y energía	Mantenimiento de las poblaciones humanas, calidad de vida, seguridad, producción y distribución de comida, suministros de energía	Sistemas naturales renovables y no renovables, crecimiento poblacional, uso sostenible de las especies
Calidad medioambiental	Acciones respetuosas con el medio ambiente, uso y desecho de materiales y dispositivos	Distribución de la población, desecho de residuos, impacto medioambiental	Biodiversidad, sostenibilidad ecológica, control de la polución, producción y pérdida del suelo y la biomasa
Riesgos	Evaluación de riesgos de las elecciones vitales	Cambios rápidos (terremotos, condiciones meteorológicas adversas), cambios lentos y progresivos (erosión de las costas, sedimentación), evaluación de riesgos	Cambio climático, impacto de las comunicaciones modernas

Tema	Personal	Local/Nacional	Global
Fronteras de la ciencia y la tecnología	Aspectos científicos de las aficiones, tecnología personal, música y actividades deportivas	Nuevos materiales, dispositivos y procesos, modificaciones genéticas, tecnología de la salud, transporte	Extinción de las especies, exploración del espacio, origen y estructura del universo

Fuente: OCDE, 2017a, p.26

Las **competencias** científicas que pretenden medir los ítems se dividen en tres bloques o secciones, ya perfilados en la propia definición de la competencia: explicar fenómenos, evaluar y diseñar investigaciones, e interpretar datos y pruebas.

Dentro de la habilidad para explicar fenómenos de manera científica, el objetivo es comprobar si los estudiantes son capaces de:

- Recordar y aplicar los conocimientos científicos apropiados.
- Identificar, utilizar y generar modelos y representaciones explicativas.
- Realizar y justificar predicciones acertadas.
- Ofrecer hipótesis explicativas.
- Explicar las implicaciones potenciales del conocimiento científico para la sociedad.

Por su parte, la capacidad para evaluar y diseñar investigaciones científicas requiere una visión crítica de la ciencia por parte de los estudiantes para poder separar las preguntas o argumentos formulados de acuerdo con los principios científicos de aquellos que han sido alcanzados por otros medios. Esta competencia se concreta en las habilidades para:

- Identificar la cuestión explorada en un determinado estudio científico.
- Distinguir cuestiones que podrían ser investigadas de manera científica.
- Proponer maneras de explorar de manera científica una cuestión determinada.
- Evaluar maneras de explorar de manera científica una cuestión determinada.
- Describir y evaluar cómo la comunidad científica es capaz de garantizar la fiabilidad de los datos y la objetividad y la posibilidad de generalizar sus explicaciones.

Por último, el bloque relativo a la interpretación científica de datos y evidencias pretende valorar la capacidad de los estudiantes para:

- Transformar datos de una representación a otra.

- Analizar e interpretar datos y extraer conclusiones apropiadas.
- Identificar los supuestos, evidencias y razonamientos en textos relacionados con la ciencia.
- Distinguir entre argumentos basados en evidencias y teorías científicas y aquellos basados en otras consideraciones.
- Evaluar argumentos y evidencias científicas procedentes de diversas fuentes (periódicos, Internet, revistas científicas).

Los **conocimientos** alrededor de los cuales se organiza el dominio científico en PISA son fundamentalmente de tres tipos: conocimientos de contenido, de procedimiento y epistemológicos.

Los conocimientos de contenido seleccionados para formar parte de la evaluación se caracterizan por su relevancia para la vida diaria de los estudiantes y su importancia dentro del cuerpo teórico de la ciencia. Aunque no existe una lista exhaustiva de los contenidos que pueden ser considerados en PISA, en general están relacionados con tres tipos de sistemas: los sistemas físicos (estructura y propiedades de la materia, movimiento y fuerzas, energía y sus transformaciones, etc.), los sistemas vivos (células, organismo, salud y anatomía humana, biodiversidad, genética, ecosistemas, etc.) y los sistemas terrestres y espaciales (capas de la Tierra, historia de la Tierra y su posición en el espacio, historia y escala del espacio, etc.).

Los conocimientos procedimentales están relacionados con el método científico y los procesos necesarios para obtener datos que sean válidos y fiables. Estos conocimientos pueden referirse a diversos aspectos del proceso científico, como los distintos tipos de variables y su control, los paradigmas cuantitativo y cualitativo y sus distintas maneras de medición, mecanismos para garantizar la replicabilidad y precisión de los datos, o la naturaleza de los diversos tipos de diseños metodológicos.

Los conocimientos epistemológicos se refieren a los constructos y características básicas que es necesario conocer para comprender el proceso científico. Cuestiones como las diferencias entre teoría e hipótesis, o hechos y observaciones, conforman la base de conocimientos que un estudiante debe poseer para poder explicar no tanto el qué (contenido) ni el cómo (procedimiento), sino el porqué de los procesos científicos. Algunos de los conocimientos epistemológicos evaluados por PISA son: los propósitos y objetivos de la ciencia, sus valores (objetividad, sesgos, etc.), la naturaleza de los

razonamientos, el uso de modelos físicos o abstractos, o la función de las diferentes formas de investigación científica a la hora de establecer conocimientos.

1.2.1.2. Competencia lectora

El marco de PISA 2015 define la competencia lectora como la habilidad para “entender, utilizar, reflexionar y enfrentarse a textos escritos para alcanzar los propios objetivos, desarrollar el conocimiento y el potencial personales, y participar en la sociedad” (OECD, 2017a, p.51). El concepto de competencia lectora va más allá de la propia acción de descodificar una serie de grafemas o leer de manera mecánica, requiriendo del sujeto una participación activa y una aplicación de la lectura significativa y funcional.

Al igual que el científico, el dominio de la lectura se organiza alrededor de varios elementos, como el contexto, las características formales del texto, o los procesos cognitivos necesarios para resolver la tarea.

Por un lado, el tipo de **situación o contexto** indica a qué ámbito se refiere el texto ofrecido como estímulo. Así, los textos enmarcados en un contexto personal se refieren a aquellas lecturas cuyo objetivo es cubrir las necesidades de ocio o los intereses personales de un individuo, así como todas aquellas comunicaciones relacionadas con sus relaciones personales. Por su parte, los textos de carácter público son aquellos de interés social, como documentos oficiales o noticias. Los textos educativos son aquellos diseñados para la instrucción, como por ejemplo los libros de texto. Por último, los textos del ámbito profesional son aquellos relacionados con el mercado de trabajo, como ofertas de empleo.

Por su parte, el texto en sí mismo se divide en función de su formato y su tipo. Los **formatos** pueden ser continuos (frases organizadas en párrafos), no continuos (listas, tablas, gráficos, etc.), mixtos o múltiples (textos de diversa naturaleza, yuxtapuestos pero no necesariamente relacionados de manera directa).

El **tipo** de texto hace referencia a su propósito, y puede ser descriptivo, narrativo, expositivo, argumentativo, instructivo o transaccional.

Por último, el elemento denominado **aspecto** (figura 1.2) se refiere al tipo de proceso cognitivo que el sujeto necesita poner en marcha para poder resolver la tarea que se le plantea. Se proponen cinco procesos principales divididos en tres categorías más amplias.

- **Acceder y obtener:** se trata de la categoría más básica, y se corresponde solamente con el proceso de obtener información, que supone encontrar en el estímulo los datos necesarios para contestar a la pregunta planteada.
- **Integrar e interpretar:** la integración se refiere a la conexión de las distintas informaciones proporcionadas por el texto, mientras que la interpretación supone la construcción de un significado no expresado explícitamente en el estímulo. Ambos elementos son necesarios tanto para el desarrollo de una comprensión global del texto, como para la elaboración de una interpretación, la cual favorecerá una comprensión más profunda del texto y todos sus significados explícitos e implícitos.
- **Reflexionar y valorar:** a través del acceso a las propias experiencias o del uso de criterios no expresados en el texto, las actividades relacionadas con esta categoría requieren que el sujeto compare el estímulo con fuentes externas a él. Este tipo de acciones se pueden llevar a cabo para extraer conclusiones tanto del contenido como del formato del texto presentado.

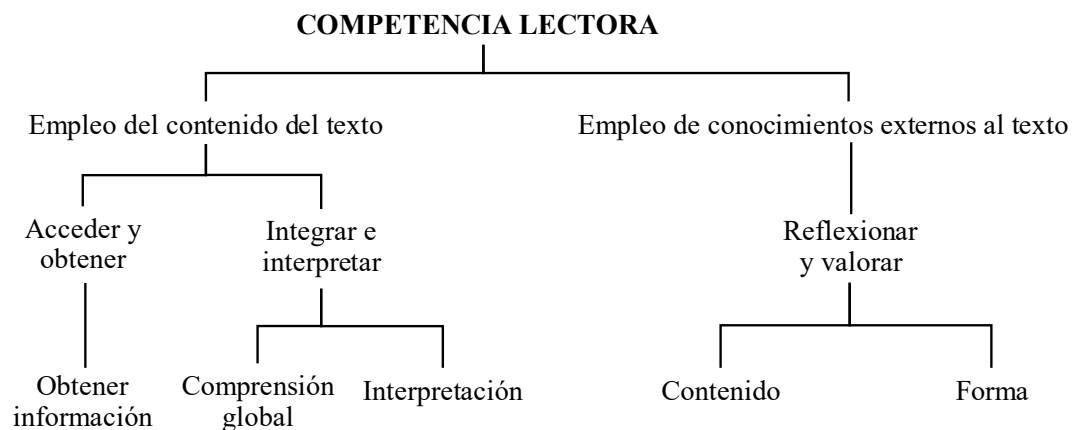


Figura 1.2. Relación entre el marco de la competencia lectora y la subescala de aspecto. Adaptado de OECD, 2017a.

1.2.1.3. Competencia matemática

La competencia matemática es entendida como

la capacidad del individuo para formular, emplear e interpretar las matemáticas en distintos contextos. Incluye el razonamiento matemático y el uso de conceptos, procedimientos, datos y herramientas matemáticas para describir, explicar y predecir fenómenos. Ayuda a los individuos a reconocer el papel que las matemáticas desempeñan

en el mundo y a emitir los juicios y las decisiones bien fundamentadas que son necesarias para los ciudadanos constructivos, comprometidos y reflexivos. (OECD, 2017a, p. 67).

El dominio de la competencia matemática se organiza a través de tres elementos generales: los procesos matemáticos necesarios para resolver los problemas que plantean los ítems, el contenido matemático en el que se basan y el contexto al que se refieren.

PISA propone tres grandes grupos de **procesos** matemáticos (tabla 1.3) que subyacen a las operaciones que los estudiantes deben realizar para resolver con éxito los problemas de evaluación.

El primer proceso es la formulación matemática de situaciones, que se define como la capacidad de los individuos para proponer una estructura matemática que relacione los distintos elementos presentados en el problema y sea capaz de dar una respuesta a la pregunta planteada. Este proceso puede suponer diversas actividades, como identificar variables, reconocer estructuras matemáticas, traducir un problema a un lenguaje o representación matemáticos, o utilizar tecnologías para representar una relación matemática inherente a un problema.

El segundo proceso es el empleo de conceptos, hechos, procedimientos y razonamientos matemáticos para resolver problemas y alcanzar conclusiones basadas en las matemáticas. Las actividades promovidas para evaluar este proceso incluyen la aplicación de normas y algoritmos matemáticos, la realización de diagramas y gráficos para poder extraer datos de ellos, la realización de generalizaciones basadas en los resultados obtenidos, o la reflexión sobre argumentos matemáticos y la justificación de los resultados matemáticos.

El tercer y último proceso está relacionado con la interpretación, aplicación y evaluación de los resultados matemáticos que permitan determinar si son razonables y tienen sentido en el contexto del problema planteado. Algunas de las actividades que se corresponden con este proceso son la comprensión de cómo el mundo afecta a los procesos y resultados matemáticos para poder ajustarlos de manera conveniente, la contextualización de las conclusiones matemáticas, o la identificación de los límites de los conceptos, soluciones y modelos matemáticos.

Tabla 1.3. Relaciones entre los procesos y las habilidades matemáticas fundamentales

	Formulación matemática de situaciones	Empleo de conceptos, hechos, procedimientos y razonamientos matemáticos	Interpretación, aplicación y evaluación de resultados matemáticos
Comunicación	Leer, decodificar y dar sentido a afirmaciones, preguntas, tareas, objetos o imágenes	Articular una solución, mostrar el trabajo implicado en la resolución de un problema y/o resumir y presentar resultados matemáticos intermedios	Construir y comunicar explicaciones y argumentos en el contexto del problema
Matematización	Identificar las variables y estructuras subyacentes en problemas del mundo real y realizar conjeturas	Utilizar la comprensión del contexto para guiar o agilizar la resolución matemática de problemas	Comprender el alcance y los límites de una solución matemática como consecuencia del modelo matemático empleado
Representación	Crear una representación matemática de información del mundo real	Dar sentido, relacionar y utilizar diversas representaciones a la hora de interactuar con un problema	Interpretar soluciones matemáticas en diversos formatos en relación con una situación o uso; comparar o evaluar dos o más representaciones en relación con una situación
Razonamiento y argumentación	Explicar, defender o justificar la representación identificada o propuesta de una situación del mundo real	Explicar, defender u ofrecer una justificación para los procesos y procedimientos utilizados para determinar un resultado o solución matemática. Conectar informaciones para alcanzar una solución matemática, realizar generalizaciones o crear un argumento de varias fases	Reflexionar sobre soluciones matemáticas y crear explicaciones y argumentos que apoyen o refuten la solución a un problema del mundo real
Diseño de estrategias para la resolución de problemas	Seleccionar o desarrollar una estrategia para reformular problemas contextualizados	Activar mecanismos de control efectivos para llevar a cabo procedimientos de resolución de problemas en varias fases	Desarrollar e implementar una estrategia para interpretar, evaluar y validar una solución matemática a un problema contextualizado
Uso de operaciones y lenguajes simbólicos, formales y técnicos	Utilizar variables, símbolos, diagramas y modelos adecuados para representar problemas del mundo real a través de lenguajes simbólicos/formales	Comprender y utilizar constructos formales basados en definiciones, normas y sistemas formales, así como emplear algoritmos	Comprender la relación entre el contexto del problema y la representación de la solución matemática, y usarla para interpretar la solución en su contexto y evaluar la viabilidad y limitaciones de la solución
Uso de herramientas matemáticas	Utilizar herramientas matemáticas para reconocer estructuras matemáticas o representar relaciones matemáticas	Conocer y saber usar de manera adecuada diversas herramientas que pueden ayudar en la implementación de procesos para determinar soluciones matemáticas	Utilizar herramientas matemáticas para determinar hasta qué punto es razonable una solución y cuáles son sus límites y restricciones en función del contexto del problema

Fuente: OECD, 2017a, p. 71.

Para poder resolver con éxito las cuestiones que se plantean en la prueba competencial de matemáticas, los estudiantes también deben dominar distintos tipos de **contenido**, divididos en PISA en cuatro grupos: cambio y relaciones, espacio y forma, cantidades, e incertidumbre y datos. Los contenidos relacionados con el cambio y las relaciones requieren que los estudiantes sean capaces de utilizar modelos y funciones matemáticas que representen adecuadamente las relaciones entre distintas variables y cómo dichas relaciones pueden verse afectadas por cambios puntuales o continuos. La categoría de espacio y forma, por su parte, se refiere a cuestiones presentes en el mundo físico que nos rodea, como patrones, propiedades de los objetos, posiciones y orientaciones, y se basa fundamentalmente en la geometría, aunque también utiliza elementos de otras disciplinas matemáticas. Los contenidos relacionados con las cantidades tienen que ver con las medidas, magnitudes unidades e indicadores de los objetos, situaciones o entidades del mundo. Por último, la categoría de incertidumbre y datos está relacionada con el reconocimiento de la variabilidad y el error de medida de procesos matemáticos como encuestas, previsiones meteorológicas o modelos económicos, y se basa fundamentalmente en las disciplinas de estadística y probabilidad.

Al igual que en los dominios anteriores, también se determinan una serie de **contextos** a los cuales pueden referirse las tareas planteadas en la evaluación de la competencia matemática, personal, ocupacional, social y científica.

1.2.2. Cuestionarios de contexto

Los cuestionarios de contexto aplicados en PISA tienen como objetivo principal recoger una gran cantidad de información complementaria al nivel competencial de los estudiantes. Como puede observarse en la figura 1.3, esta información puede referirse a diversos elementos de carácter contextual, procesual o relacionada con los llamados resultados no cognitivos de los estudiantes.

Toda esta información se recoge a través de diversos instrumentos, diferenciados en función del informante al que van dirigidos. Los cuestionarios generales, enfocados a recoger información y opiniones del alumnado y la dirección de los centros educativos, se aplican de manera obligatoria a todos los participantes.

Antecedentes del alumnado		Procesos			Resultados no cognitivos	
Familia	Educación	Actores	Procesos centrales	Asignación recursos		
Temas ciencia		5. Experiencia científica fuera del centro	1. Cualificación del profesor y conocimientos profesionales	2. Prácticas de enseñanza de ciencias	12. Tiempo de aprendizaje y plan de estudios	4. Resultados relacionados con la ciencia: motivación, interés, creencias...
	Enseñanza y aprendizaje			3. Ambiente de aprendizaje para la ciencia a nivel escolar		
Temas generales	7. Estatus socioeconómico y familiar	9. Itinerarios educativos en la primera infancia	14. Implicación de la familia	13. Ambiente escolar: relaciones, expectativas	16. Recursos	6. Aspiraciones profesionales 10. Comportamiento general y actitudes 11. Disposición para la resolución cooperativa de problemas
	8. Etnicidad y migración		17. Locus de toma de decisiones	19. Evaluación y rendición de cuentas	18. Asignación, selección y elección	
			Políticas del centro			
			Gobernabilidad			

Figura 1.3. Estructura modular de los cuestionarios de contexto en PISA 2015. Adaptado de OECD, 2017a.

Los temas tratados en los cuestionarios de contexto son amplios y diversos, encontrándose cuestiones como las siguientes:

- Datos socioeconómicos y demográficos de alumnado y centros, incluyendo género, estatus migratorio, nivel socioeconómico, recursos educativos y culturales disponibles, titularidad y localización del centro, etc.
- Aspectos de la vida del alumnado relacionados con el aprendizaje (resultados no cognitivos), como su motivación y sus actitudes hacia la educación, su compromiso y disfrute de los distintos dominios cognitivos evaluados, o sus hábitos de estudio y trabajo.

- Factores relacionados con los procesos de enseñanza y aprendizaje, como las metodologías docentes utilizadas, el clima del aula o el apoyo educativo proporcionado por los docentes.
- Políticas y gobernabilidad del centro educativo, como el liderazgo, el desarrollo del currículo, la participación del profesorado en la toma de decisiones, la formación continua del profesorado, etc.

Además de estos dos cuestionarios, PISA ofrece la participación voluntaria en otra serie de cuestionarios que aportan información complementaria sobre diversas cuestiones:

- Cuestionario de familiaridad con la informática, sobre el uso, conocimientos y actitudes del alumnado con respecto a las tecnologías de la información y comunicación.
- Cuestionario sobre la carrera educativa, que pretende recoger información sobre la preparación y aspiraciones del estudiante sobre su futuro académico y laboral.
- Cuestionario para familias, centrado en sus percepciones, actitudes e implicación en la educación de sus hijos, el apoyo que prestan a las cuestiones relacionadas con su aprendizaje, o sus expectativas académicas y laborales futuras.
- Cuestionario para docentes, que fue introducido como novedad en el ciclo de 2015, y se centra en cuestiones relacionadas con las prácticas y metodologías empleadas en clase.

1.3. CARACTERÍSTICAS METODOLÓGICAS

Debido al gran número de países, centros educativos y estudiantes que participan en la prueba, y al volumen de información que pretenden recoger sus instrumentos, el diseño de la evaluación PISA es complejo y presenta una serie de particularidades que se describen a continuación.

1.3.1. Diseño muestral

La población objeto de estudio de la prueba PISA está compuesta por todos aquellos estudiantes de 15 años (entre 15 años y 3 meses y 16 años y 2 meses) matriculados en un curso superior a 6º de educación obligatoria en una institución educativa, sin importar la dedicación al estudio (completa o parcial) o el tipo de programa (académico o profesional).

Para seleccionar la muestra de cada uno de los países se llevó a cabo un muestreo estratificado⁹ de dos etapas. En la primera etapa se seleccionaron los centros educativos, cuya probabilidad de ser escogidos era proporcional a su tamaño, calculado en función del número de estudiantes de 15 años matriculados en dicho centro. La segunda etapa consistió en conformar una muestra con estudiantes de los centros previamente seleccionados. En los países que decidieron aplicar la prueba a través de medios informáticos, se seleccionaron 42 estudiantes al azar en los centros con más de 42 estudiantes elegibles, y en los centros con menos de 42 estudiantes de 15 años, todos los estudiantes fueron seleccionados para participar en la prueba. En el caso de países que aplicasen la prueba en papel, el número de corte para la selección de estudiantes fue 35.

Existen algunas excepciones en el proceso de muestreo de estudiantes, fundamentalmente relacionadas con necesidades educativas especiales o insuficiencias en el conocimiento de la lengua vehicular del sistema educativo, aplicables también a los centros educativos que solamente atiendan a este tipo de estudiantado.

En general, se seleccionaron al menos 150 centros educativos de cada país, y alrededor de 35-42 estudiantes por cada centro, dando como resultado una muestra mínima de entre 4500 y 5250 estudiantes, a menos que la población estudiantil de un país fuese inferior a esa cifra, en cuyo caso toda la población formaría parte de la muestra. En el caso de España, además de la muestra nacional, en el ciclo de 2015 las Comunidades Autónomas decidieron ampliar la muestra para poder convertirse en unidades comparables a nivel internacional, por lo que la muestra española se compuso de 32.330 estudiantes y 976 centros educativos (Ministerio de Educación, Cultura y Deporte, 2016).

Pesos muestrales

Debido al proceso de muestreo utilizado en las pruebas PISA, no todos los estudiantes de un país tienen la misma probabilidad de ser seleccionados para formar parte de la muestra. Los motivos de esta desigualdad pueden ser de diversa naturaleza, como el sobre o submuestreo de ciertos sectores de la población, la falta de respuesta de algunos centros o estudiantes, o la información errónea acerca del tamaño de los centros en el momento de realizar la muestra. Esta cuestión genera la necesidad de utilizar los pesos muestrales

⁹ Cada país escogió su variable de estratificación de preferencia. En el caso de España, las variables de estratificación explícita fueron la Comunidad Autónoma y la financiación, así como el modelo lingüístico para los centros del País Vasco. No hubo variables de estratificación implícita.

a la hora de realizar cualquier tipo de análisis para poder asegurar la representatividad de los datos obtenidos con respecto de la población general.

Los pesos muestrales de los estudiantes se calculan en función del peso base del centro educativo, el peso base intra-escuela y cinco factores de ajuste, y puede expresarse a través de la siguiente ecuación (1.1):

$$W_{ij} = t_{2ij}f_{1i}f_{2ij}f_{1ij}^A t_{1i}w_{2ij}w_{1i} \quad (1.1)$$

Donde:

w_{1i} es el peso base del centro escolar y representa el recíproco de la probabilidad de inclusión del centro i en la muestra

w_{2ij} es el peso base intra-escuela y representa el recíproco de la probabilidad de seleccionar al estudiante j dentro del centro seleccionado i

f_{1i} es un factor de ajuste cuyo objetivo es compensar la no participación de otros centros de similar naturaleza al centro i

f_{1ij}^A es un factor de ajuste destinado a compensar los centros de algunos países participantes en las que solo se seleccionaron para la muestra aquellos estudiantes de 15 años que estaban matriculados en el curso modal (10º curso, 4º de ESO en España)

f_{2ij} es un factor de ajuste destinado a compensar la no participación de los estudiantes dentro de la misma celda de no-respuesta de centro y estrato específico y, en los casos en los que la muestra lo permite, dentro de la misma categoría de rendimiento alto o bajo y de género

t_{1i} es un factor de recorte del peso base del centro escolar, utilizado para reducir valores inesperadamente altos de w_{1i}

t_{2ij} es un factor de recorte del peso del estudiante, utilizado para reducir los pesos de los estudiantes con valores inesperadamente altos en el producto de todos los componentes anteriores

Por otro lado, con el objetivo de estimar la varianza muestral de las estimaciones de los parámetros, PISA utiliza una técnica de remuestreo conocida como replicación repetida equilibrada (*balanced repeated replication*, BRR), en concreto emplea la variación de Fay. Este método se basa en el emparejamiento de centros en función de las variables de estratificación explícitas e implícitas utilizadas durante el proceso de muestreo y la asignación sistemática de distintos pesos a cada uno de estos centros

siguiendo una matriz de Hadamard de orden 80. Este método de remuestreo permite obtener como resultado 80 réplicas que le confieren a los análisis una mayor precisión estadística (Lee & Forthofer, 2005).

1.3.2. Modelos de Teoría de Respuesta al Ítem y generación de Valores Plausibles de rendimiento

En el último ciclo de PISA se han implementado algunos cambios en relación con los modelos utilizados para calibrar los ítems de las escalas. Hasta 2012 se utilizó una combinación del modelo de Rasch (1961) y el Modelo de Crédito Parcial, pero, debido a una serie de críticas relacionadas con las posibles limitaciones del modelo de Rasch a la hora de analizar la complejidad inherente a los datos de PISA (Kreiner & Christensen, 2014; Oliveri & von Davier, 2011), en el ciclo de 2015 se cambió de estrategia analítica, y se comenzó a utilizar el modelo logístico de dos parámetros (Birnbaum, 1968) para respuestas dicotómicas, y el modelo generalizado de crédito parcial (Muraki, 1992) para ítems con más de dos categorías de respuesta.

Para asegurar que el modelo de Teoría de Respuesta al Ítem (TRI) utilizado para el análisis se ajusta adecuadamente a los datos observados, se realizan diversos análisis, como el de funcionamiento diferencial de los ítems (DIF), cuyo objetivo es reducir los sesgos debidos a cuestiones lingüísticas y/o culturales, o el modelo de invarianza métrica, que pretende estimar las diferencias de medida de los ítems en función de la modalidad de aplicación de la prueba (papel u ordenador).

Otra de las características derivadas de los modelos de análisis de los datos es la generación de diversos valores plausibles para estimar el nivel de competencia de los sujetos participantes. A través de una combinación del modelo TRI y un modelo de regresión latente, se generan una serie de valores plausibles para aumentar la precisión de las estimaciones de la distribución multivariante del nivel competencial. Estos valores podrían definirse como “el rango de habilidades que un estudiante podría tener razonablemente, dadas sus respuestas a los ítems” (Wu, 2005, p. 116), y son particularmente útiles en el análisis de datos con diseños complejos como PISA porque facilitan el abordaje de los problemas relacionados con sesgos en la estimación de los parámetros poblacionales, permiten que los análisis secundarios sean realizados a través de herramientas estadísticas normales teniendo en cuenta los errores de medida y facilitan también la computación de los errores típicos de estimación en diseños muestrales

complejos (Wu, 2005). En 2015 se generaron 10 valores plausibles para cada sujeto en cada competencia.

Debido a la aplicación de las pruebas competenciales a través de cuadernillos incompletos de ítems (ver sección 2), las estimaciones puntuales del desempeño de cada uno de los participantes podrían llevar a estimaciones ampliamente sesgadas de los parámetros poblacionales (Wingersky, Kaplan & Beaton, 1987), por lo que se hace necesario un método para poder estimar estos parámetros de manera más ajustada. Los valores plausibles permiten esta estimación de la distribución del nivel competencial de la población, permitiendo también la consideración de la incertidumbre o error al nivel individual.

1.4. CRÍTICAS A PISA

A lo largo de sus casi veinte años de historia, PISA ha generado algunas dudas y críticas en la comunidad científica en torno a distintos temas, como las competencias que mide, sus características metodológicas o los usos sociales y políticos que se hacen de sus resultados. En esta sección se exploran las principales críticas expresadas por los investigadores.

Evaluación de la calidad del sistema educativo

Jornet (2016) señala que la OCDE toma el nivel competencial del alumnado como una medida directa de la calidad de los sistemas educativos evaluados, obviando otro tipo de resultados más relacionados con el desarrollo integral del ser humano. Este autor recomienda cautela a la hora de conceptualizar PISA como un estudio evaluativo, y aconseja la incorporación de resultados de carácter psico-socio-afectivo para solventar esta limitación.

Universalidad de las competencias

Una de las cuestiones que se plantean algunos investigadores (Fernández-Cano, 2016; Jornet, 2016; Sjøberg, 2015) es hasta qué punto las pruebas PISA se acercan de manera fidedigna a las habilidades que presentan los estudiantes para enfrentarse a los retos de la vida moderna, lo cual constituye el objetivo principal de la OCDE con respecto a la aplicación de esta evaluación. En primer lugar, teniendo en cuenta que PISA se aplica en más de 70 economías alrededor del mundo, se plantea si los retos a los que se enfrentan los jóvenes como ciudadanos del siglo XXI son equivalentes independientemente del

contexto que les rodea, es decir, ¿son las competencias necesarias para el futuro iguales para un adolescente en Japón, Perú, Kazajstán o Noruega? La negativa de estos autores a esta pregunta podría traducirse en una necesidad de revisar el concepto de “competencias para la vida” que plantea PISA.

Otra de las consideraciones que se hacen sobre este punto es el formato de aplicación de las pruebas. Por ejemplo, Sjøberg (2015) encuentra cuestionable la relación entre la situación de la prueba (participar en una prueba estandarizada durante dos horas y media) y las situaciones de la vida real, en el sentido que el único contexto en el que estas dos situaciones coinciden es en el ámbito académico, no siendo así los resultados generalizables a otras situaciones.

Competencias en contexto

Una de las principales cuestiones que caracterizan la prueba PISA y que la distinguen de otras evaluaciones a gran escala, como TIMMS o PIRLS, es que se basa en un marco competencial propio, no en conocimientos reflejados en los currículos nacionales. Esta cuestión, que puede entenderse como una ventaja a la hora de proporcionar comparabilidad internacional, también puede tener implicaciones negativas. Aunque no sea de manera explícita, las cuestiones planteadas en la prueba podrían asemejarse más a unos currículos nacionales que a otros, cuestión que influiría en la interpretación de los resultados, pero que no se analiza en el marco de la evaluación PISA (Jornet, 2016). Desde otro enfoque diferente, Carabaña (2015, 2016) propone que, al no formar parte del currículo, los ítems de PISA no evalúan aprendizajes académicos sino aptitudes o competencias generales que no dependen de la escolarización o de las acciones de los docentes, dificultando de esta manera la extracción de conclusiones relacionadas con el impacto de los centros educativos en el rendimiento de los estudiantes.

Por otro lado, algunos autores observan también una cierta desconexión entre los estímulos presentados en las pruebas competenciales y los lenguajes propios de las disciplinas que evalúan, principalmente las matemáticas y la ciencia, no utilizando, por ejemplo, un gran número de representaciones simbólicas o visuales que se utilizan para expresar conocimiento en estos dominios (Yore, Anderson & Chiu, 2010), lo cual alejaría dichos estímulos de los usos reales de esta competencia.

Plurilingüismo y traducciones

Los manuales técnicos de PISA proporcionan detalladas descripciones de los procesos implicados en la traducción de los cuestionarios a los diversos idiomas de los países participantes, y en el caso de países plurilingües, a todas las lenguas cooficiales. Sin embargo, las versiones que difieren del original en inglés o francés crean ciertas controversias en algunos países, alegando los autores que el lenguaje que se utiliza no es tan natural como en los textos originales, lo cual puede dificultar la motivación e interés de los participantes para relacionarse con ellos (Arffman, 2010; Sjøberg, 2010), o la mayor cercanía de los elementos cognitivos a las culturas occidentales participantes en la prueba (Fernández-Cano, 2016). Otros autores hacen esta misma observación en relación con los constructos culturales, sociales y económicos (Caro, Sandoval-Hernández & Lüdtke, 2013).

Estos sesgos lingüísticos y culturales se encuentran asociados de manera más relevante a los países asiáticos y de Oriente Medio (Grisay, De Jong, Gebhardt, Berezner & Halleux-Monseur, 2007).

Falta de datos clave para el análisis educativo

A la hora de interpretar los datos de PISA, muchos autores apuntan a ciertas dificultades derivadas de la configuración técnica de PISA que impiden el establecimiento de relaciones causales como conclusión de sus investigaciones (ver sección 2.2.2). La principal es la falta de datos longitudinales (Carabaña, 2015; Han, 2018; Rutkowski, Rutkowski, Wild & Burroughs, 2017), o en su defecto de datos sobre el rendimiento previo de los estudiantes (Dumay & Dupriez, 2014; Willms, 2010), lo cual supone la imposibilidad de establecer modelos de valor añadido que muestren una información más ajustada de los efectos del centro educativo.

Otra falta clave de información que acusa la comunidad investigadora son los datos a nivel de aula (Lafontaine, Baye, Vieluf & Monseur, 2015; Sandilands, Barclay McKeown, Lyons-Thomas & Ercikan, 2014; Scheerens, Luyten, van den Berg & Glas, 2015). Como se ha reconocido en diversos estudios, las características de los docentes y de los procesos de enseñanza-aprendizaje llevados a cabo en el aula son altamente relevantes para el estudio de los resultados de los estudiantes (Chapman et al., 2016; Scheerens & Bosker, 1997). Sin embargo, aunque en 2015 PISA ha introducido

información del profesorado, esta no supone una equivalencia con el nivel de aula, ya que los alumnos no se encuentran identificados en función del docente que les atiende.

Análisis de datos

En este capítulo se ha presentado la enorme variedad de las variables recogidas por los instrumentos de PISA. Sin embargo, como apuntan De la Orden y Jornet (2012), la propia OCDE realiza un análisis un tanto deficitario de los datos, simplemente reportando la relación entre el rendimiento de los estudiantes y algunas de las variables sociodemográficas, pero obviando muchas de las otras variables recogidas.

En ese sentido, Jornet (2016) recomienda profundizar en estudios de corte explicativo que favorezcan una mayor comprensión de los factores asociados al rendimiento, bien por parte de la OCDE, o por parte de los institutos nacionales de evaluación educativa o los propios investigadores educativos.

Usos de PISA: influencia y convergencia en las políticas educativas

Desde su primera aplicación, los resultados de PISA han generado diversas reacciones en los gobiernos de los países participantes en función de su desempeño general. Algunos autores denominan como “PISA *shock*” a la tendencia de los gobiernos a modificar sus políticas educativas como reacción a los resultados de PISA, habitualmente basándose en los resultados descriptivos iniciales y no en los análisis secundarios más profundos que se realizan de los datos (Wiseman, 2013). Sin embargo, estos resultados pueden llevar a la realización de inferencias equivocadas sobre la calidad de los sistemas educativos (Jornet, 2016)

Aunque no todos los análisis que se hacen de este efecto son negativos – algunos autores simplemente lo reportan como una consecuencia del aumento en información y comunicación entre naciones, y otros incluso le restan importancia a la influencia real de PISA en las políticas nacionales (Yore, Anderson & Chiu, 2010) – existen algunas corrientes que detectan un empleo de los datos de la evaluación para poner en marcha políticas que no necesariamente reportan beneficios al sistema educativo del país, yendo incluso contra los principios y valores de los propios sistemas educativos (Dolins & Krogh, 2010; Sjøberg, 2015). A este respecto, Greger (2012) ejemplifica a través de los casos alemán y checo cómo desde dos reacciones opuestas a los resultados iniciales de PISA, con numerosos cambios de políticas en el caso alemán y una escasa reacción en el

checo, ninguno de los dos sistemas educativos ha intentado abordar una de sus prácticas educativas más claramente problemáticas: la segregación o selectividad de estudiantes.

En las últimas décadas, se han advertido en la literatura científica ciertas tendencias a la convergencia en materia de política educativa, principalmente a causa del aumento de la comunicación transnacional y el intercambio de prácticas educativas (Jakobi & Teltemann, 2011). Uno de los ámbitos en los que se detecta esta tendencia son las políticas de evaluación de los estudiantes, centros, profesores y las prácticas de rendición de cuentas, en las que se muestra un descenso de la variabilidad en las respuestas de los países a lo largo de las aplicaciones de PISA (Teltemann & Jude, en prensa). Otros ejemplos son la extensión del tiempo de instrucción en los países en vías de desarrollo para acercarse al tiempo dedicado a la educación en los países desarrollados, el aumento de atención individualizada que proporcionan los docentes a los estudiantes, o las políticas dirigidas a la mejora de la equidad educativa y la reducción de diferencias de resultado entre escuelas (Wiseman, 2013).

La figura 1.4 presenta una síntesis de las críticas presentadas en esta sección, agrupadas por categorías temáticas.

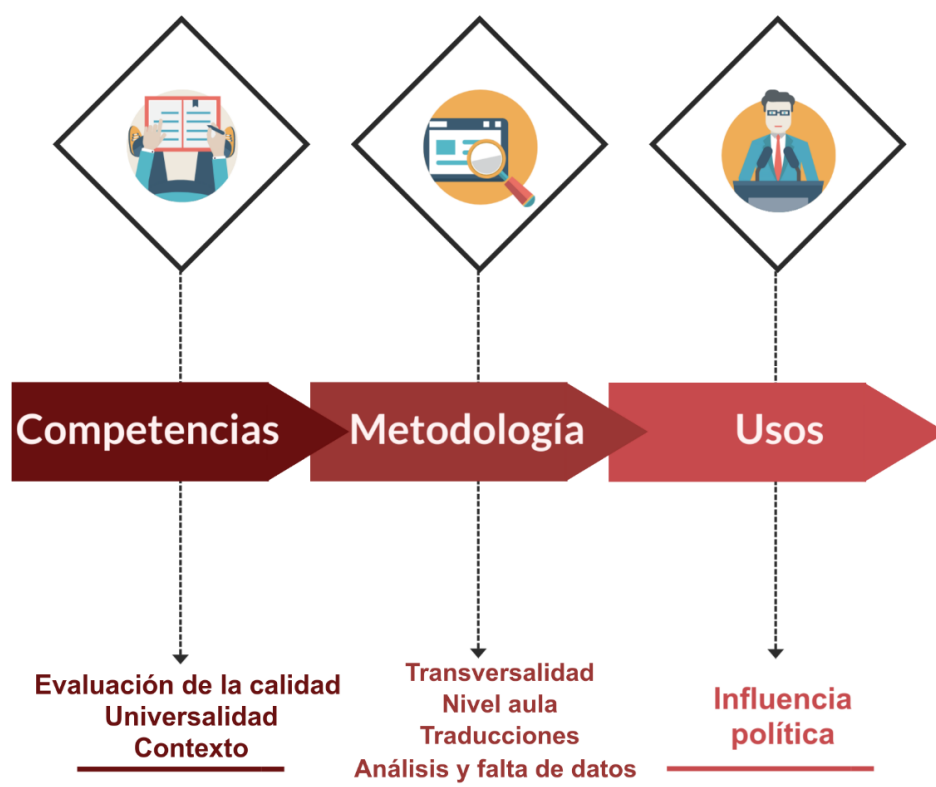
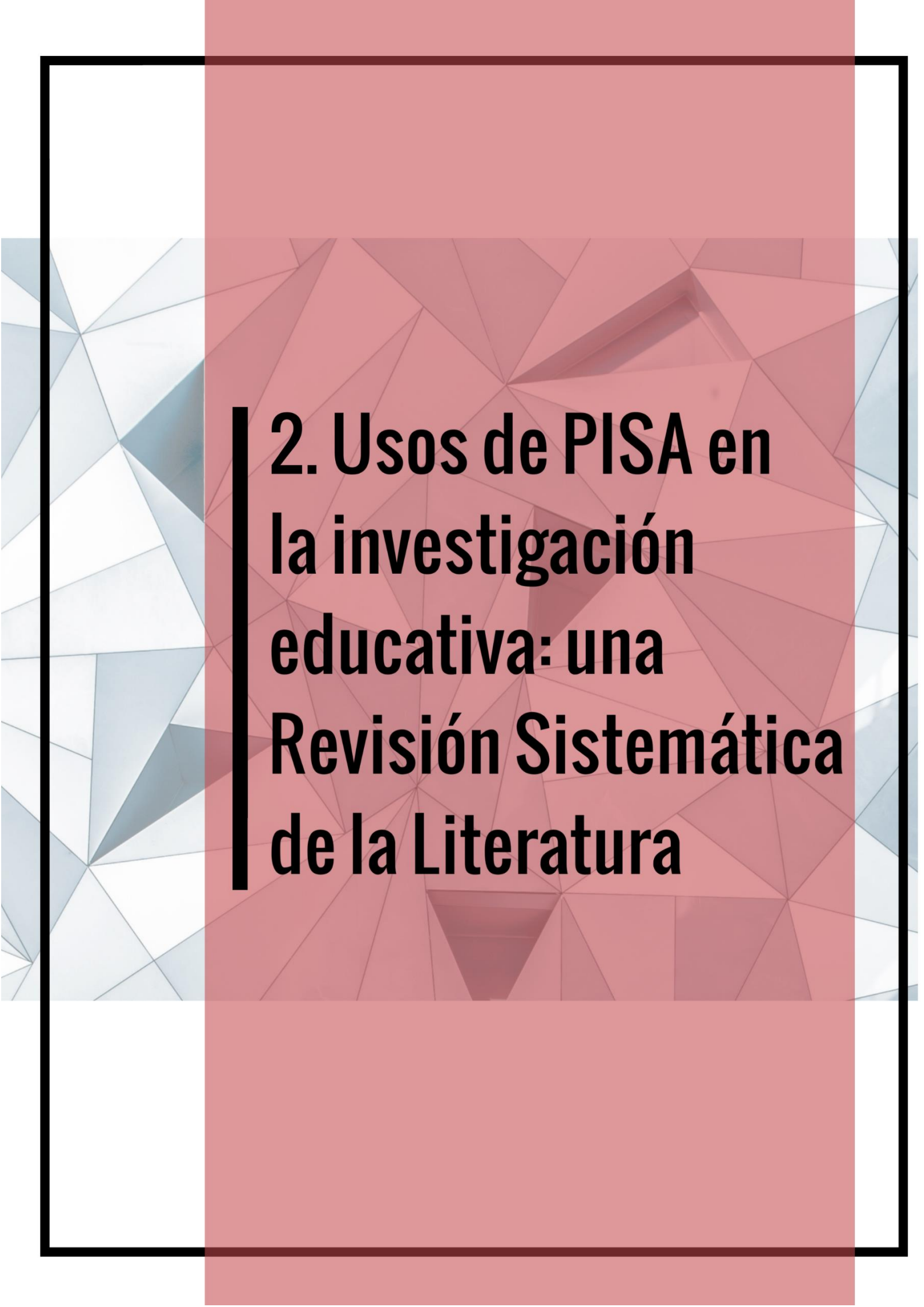


Figura 1.4. Síntesis de las principales críticas a PISA

En el marco de la presente tesis doctoral, algunas de estas críticas se presentan abiertamente como limitaciones claras del trabajo, como la falta de longitudinalidad y de datos del nivel de aula. Otras, fundamentalmente aquellas referidas a las cuestiones metodológicas de la prueba, pretenden abordarse desde una perspectiva metodológica robusta que asegure todas las garantías posibles con respecto al tratamiento de los datos (consideración de la estructura jerárquica, uso de los valores plausibles, incorporación de los pesos muestrales, etc.).

Aquellas consideraciones que tienen que ver con la diversidad lingüística, curricular y contextual de los países participantes se juzgan menos relevantes para el presente trabajo teniendo en cuenta que el país de estudio forma parte de la cultura occidental predominante en la prueba PISA. España pertenece culturalmente al grupo “latino-europeo”, y forma junto con el resto de los países europeos y occidentales la región “meta-occidental” (Gupta, Hanges & Dorfman, 2002) debido a sus similitudes, por lo que su ajuste a la visión cultural presentada por la OCDE a través de PISA es relativamente alto, y menos proclive a la generación de sesgos.

Por último, con respecto a las críticas referidas a los usos políticos de PISA y las convergencias políticas en materia de educación que parecen estar produciéndose en los últimos años en algunos países participantes, creemos que la existencia de análisis secundarios por parte de investigadores independientes tanto de la OCDE como de los gobiernos nacionales puede contribuir a una difusión de resultados más objetivos sobre la calidad de los sistemas educativos y proporcionar a las personas a cargo de las decisiones políticas y educativas a todos los niveles el acceso a una información más imparcial sobre el impacto de los factores en el rendimiento y las buenas prácticas educativas. Por otro lado, estos análisis secundarios también abordarían la crítica del análisis deficitario de los datos que apuntan De la Orden y Jornet (2012) en su trabajo, ya que, como puede observarse en el capítulo 2, existe en la literatura científica un análisis exhaustivo de un gran número de variables de diversa naturaleza.



2. Usos de PISA en la investigación educativa: una Revisión Sistemática de la Literatura

Usos de PISA en la investigación educativa: una Revisión Sistemática de la Literatura

Con el objetivo de conocer cómo se aborda el análisis de los datos de PISA para el estudio del rendimiento escolar en los diferentes países, se ha realizado un estudio de *mapping* (Kitchenham, Budgen, & Brereton, 2011) y una revisión sistemática de la literatura (*Systematic Literature Review*, SLR) (Kitchenham y Charters (2007), dos metodologías de revisión que priorizan la obtención y análisis de la información de manera intencional y planificada siguiendo un proceso organizado y reproducible.

Estas metodologías proponen la formulación de una serie de preguntas de investigación previas que se pretenden responder mediante la búsqueda y extracción de información de la literatura existente, siguiendo un proceso estructurado y diseñado para optimizar el análisis de grandes cantidades de documentos. Adicionalmente, la realización de revisiones sistemáticas implica el registro de cada paso dado para garantizar la transparencia y reproducibilidad de la investigación.

La revisión de la literatura realizada en el presente capítulo comienza con un estudio de *mapping*, realizado con el objetivo de responder preguntas de investigación más generales que sirvan para ofrecer un panorama general del número de investigaciones realizadas sobre el estudio de la eficacia escolar a través de los datos de PISA, la cual permite identificar cuestiones básicas como la evolución del interés en la comunidad investigadora, los autores más relevantes en esta materia y las áreas susceptibles de ser exploradas.

Por otro lado, también se ha realizado una SLR para responder preguntas de investigación más complejas, referidas a la metodología de análisis, las variables y las limitaciones más habituales. La respuesta a estas preguntas requiere de un análisis más profundo de los artículos que el que permiten los estudios de *mapping*, por lo que se trata de dos metodologías de revisión de la literatura complementarias, una destinada a ofrecer una panorámica general del estado de la investigación y la otra a resolver interrogantes más profundos sobre el contenido de las producciones científicas.

Para la realización de esta revisión se ha seguido la metodología propuesta por Kitchenham y Charters (2007) y Cruz-Benito, García-Peñalvo y Therón (2019). En consecuencia, este capítulo se divide en tres secciones. En la primera sección se presentan

la metodología, incluyendo las preguntas de investigación, los criterios de inclusión, los repositorios y cadena de búsqueda empleados y los resultados de las diferentes fases del proceso de selección de las publicaciones. En la segunda sección se ofrecen los resultados obtenidos con el objetivo de responder las preguntas de investigación. Por último, en la sección tres se cierra este capítulo analizando las implicaciones y limitaciones derivadas de la revisión realizada.

2.1. METODOLOGÍA

A continuación, en esta sección se describirá la metodología de investigación empleada y el proceso seguido para realizar la búsqueda y selección de los artículos incluidos en esta revisión de la literatura.

2.1.1. Preguntas de investigación

El primer paso a la hora desarrollar cualquier revisión sistemática implica la formulación de las preguntas de investigación, cuya función es orientar y dar sentido al estudio (Kitchenham & Charters, 2007). Así, se formularon dos grupos de preguntas, el primero para ser respondidas mediante el estudio de mapping y el segundo para ser respondidas mediante la SLR.

Teniendo en cuenta las características ya descritas de los estudios de *mapping*, para esta revisión se formularon preguntas de carácter más global que pudieran aportar una visión general sobre el estado y desarrollo de los estudios centrados en el análisis de la eficacia escolar a partir de los datos de PISA. Para ello se plantean las siguientes preguntas de *mapping*:

PM1. ¿Cómo ha sido la evolución de la producción científica sobre análisis secundarios de PISA desde su inicio hasta el presente?

PM2. ¿Quiénes son los autores de referencia en el campo de los análisis secundarios de los datos de PISA?

PM3. ¿Cuáles son las fuentes bibliográficas que han demostrado un mayor interés por la publicación de análisis secundarios de los datos de PISA?

PM4. ¿Qué países se analizan con mayor frecuencia? En caso de que se agrupen varios países para su análisis ¿Qué criterio de agrupación se sigue?

PM5. ¿Qué competencias se utilizan con mayor frecuencia como variable criterio?
¿Con qué frecuencia la competencia o competencias evaluadas en un estudio son la competencia principal del ciclo al que se refieren los datos?

PM6. ¿Qué ciclos de PISA han sido los más analizados?

PM7. ¿Qué metodologías son las más empleadas para el análisis de datos secundarios de PISA?

Por otro lado, para la SLR se propusieron preguntas que buscan profundizar más en el contenido de las publicaciones, centradas en las siguientes cuestiones referidas a variables, metodología y limitaciones:

PI1. ¿Cuáles son las variables predictoras más comúnmente utilizadas?

PI2. ¿Con qué frecuencia tienen en cuenta los autores la estructura jerárquica de los datos de PISA en el diseño de sus estudios?

PI3. ¿Con qué frecuencia utilizan los autores los valores plausibles en los cálculos relacionados con la variable criterio?

PI4. ¿Con qué frecuencia tienen en cuenta los autores el uso de los pesos muestrales en la estimación de parámetros? ¿Cuál es el método más habitual de incorporación de los pesos muestrales?

PI5. ¿Cuáles son las limitaciones del uso de PISA más comúnmente reportadas en la literatura científica?

2.1.2. Criterios selección

Una vez formuladas las preguntas de investigación, se diseñaron los criterios de inclusión de publicaciones que garantizaran la selección de los artículos que fueran relevantes para la presente revisión:

Criterios de inclusión (CI):

CI1. Los trabajos de investigación están relacionados con el análisis estadístico de los datos de PISA.

CI2. Los trabajos de investigación están publicados en revistas científicas indexadas en bases de datos de impacto.

CI3. Los artículos de investigación utilizan como variable criterio el rendimiento.

CI4. Los trabajos de investigación están escritos en inglés.

CI5. Los trabajos de investigación han sido publicados a partir del año 2000.

CI6. Los trabajos de investigación han sido publicados tras ser sometidos a un proceso de revisión de pares.

Para que una publicación sea considerada para análisis dentro de esta revisión, ha de cumplir con todos los criterios de inclusión.

Adicionalmente, se elaboraron también una serie de criterios de calidad que permitieran asegurar la validez de los trabajos de investigación que hubiesen pasado el proceso de inclusión/exclusión.

En total se formularon 10 preguntas (tabla 2.1) con tres opciones de respuesta: sí, no y parcialmente. Estas diez preguntas se aplicaron a todos los trabajos otorgando 1 punto por cada pregunta respondida de manera afirmativa, 0,5 puntos por cada pregunta respondida de manera parcial y 0 puntos por cada pregunta respondida de manera negativa. Finalmente, se estableció el valor de corte 7,5 quedando eliminados de la selección final todos aquellos artículos que no llegaran a dicha puntuación.

Tabla 2.1. Criterios de calidad

Pregunta
1. ¿Existen preguntas de investigación (u objetivos) cuantitativas o cualitativas claramente identificadas?
2. ¿Permiten los datos disponibles abordar las preguntas y objetivos planteados?
3. ¿Utiliza el estudio una muestra completa de PISA (es decir, un muestreo realizado por expertos de PISA)?
4. ¿Son las variables apropiadas (origen claro, validez conocida, instrumento estandarizado) y están claramente definidas?
5. ¿Se explican los métodos de manera clara?
6. ¿Son los métodos utilizados adecuados para contestar a las preguntas/objetivos?
7. ¿Se da respuesta a todas las preguntas/objetivos de investigación?
8. ¿Se basan las conclusiones en los datos expuestos en la investigación?
9. ¿Se hace un uso correcto de los valores plausibles?
10. ¿Se hace un uso correcto de los pesos muestrales?

Para aplicar los criterios de calidad de la forma más objetiva se elaboró una rúbrica de evaluación para asignar las puntuaciones (Tabla 2.2).

Tabla 2.2. Rúbrica para la evaluación de las preguntas de calidad

Pregunta	No	Parcial	Sí
1. ¿Existen preguntas de investigación (u objetivos) cuantitativas o cualitativas claramente identificadas?	No hay objetivos/preguntas claras, ni se percibe un intento de explicarlas en el texto	Las preguntas/objetivos no se especifican como tal, pero son claramente deducibles del texto	Las preguntas/objetivos están claramente identificados en el texto
2. ¿Permiten los datos disponibles abordar las preguntas y objetivos planteados?	Los datos son insuficientes para abordar las preguntas/objetivos	Los datos permiten contestar las preguntas/objetivos de manera parcial	Los datos permiten contestar de manera adecuada a todas las preguntas/objetivos
3. ¿Utiliza el estudio una muestra completa de PISA (es decir, un muestreo realizado por expertos de PISA)?	Utiliza una submuestra ad hoc y no queda claro si es representativa de la población estudiada	Utiliza una submuestra ad hoc pero su composición y representatividad están claramente explicadas	Utiliza una muestra seleccionada originalmente por los expertos de PISA
4. ¿Son las variables apropiadas (origen claro, validez conocida, instrumento estandarizado) y están claramente definidas?	El origen de las variables no es claro/válido/fiable, y/o las variables no se encuentran claramente definidas	El origen de las variables es claro, válido y fiable pero su definición no es clara	El origen de las variables es claro, válido y fiable y están claramente definidas
5. ¿Se explican los métodos de manera clara?	El método no se explica de manera clara en el texto	Algunos elementos del método no se explican con claridad	El método se explica claramente en el texto
6. ¿Son los métodos utilizados adecuados para contestar a las preguntas/objetivos?	El método no produce los resultados necesarios para contestar a las preguntas	El método permite contestar a las preguntas, aunque no es el método óptimo	El método es el más adecuado para responder a las preguntas planteadas
7. ¿Se da respuesta a todas las preguntas/objetivos de investigación?	Al menos dos de las preguntas no se contestan de manera adecuada	Se contestan todas las preguntas excepto una	Se contestan todas las preguntas
8. ¿Se basan las conclusiones en los datos expuestos en la investigación?	Las conclusiones no se basan en los datos presentados en la investigación	Las conclusiones están más o menos basadas en los resultados, pero los autores realizan afirmaciones no basadas en los datos	Todas las conclusiones están basadas en los datos presentados en la investigación
9. ¿Se hace un uso correcto de los valores plausibles (VP)?	No se menciona el uso de VP / No se usan los VP / Hay un uso incorrecto de los VP	El uso de VP se menciona pero no se explica claramente	El uso de VP es adecuado y está correctamente explicado
10. ¿Se hace un uso correcto de los pesos muestrales (PM)?	No se menciona el uso de PM / No se usan los PM / Hay un uso incorrecto de los PM	El uso de PM se menciona pero no se explica claramente	El uso de PM es adecuado y está correctamente explicado

2.1.3. Cadena de búsqueda

El último paso antes de llevar a cabo el proceso de selección de artículos consiste en determinar las bases de datos que se van a utilizar y la cadena de búsqueda más eficaz para localizar las publicaciones.

Para seleccionar las bases de datos se estableció que estas debían cumplir cuatro requisitos básicos: permitir el acceso a los artículos desde cuentas institucionales o inscripciones personales, admitir el uso de expresiones lógicas para construir la cadena de búsqueda, admitir búsquedas en todo el texto de los trabajos o en campos específicos y ser de relevancia en el campo. Atendiendo a estos criterios se seleccionaron para la búsqueda las bases de datos electrónicas: Web of Science (WoS) y Scopus (García Sánchez, Therón & Gómez-Isla, 2019).

Así mismo se elaboró una cadena de búsqueda adaptada a cada base de datos tras probar diferentes alternativas para asegurar que la cadena seleccionada resultaba adecuada. De esta manera en WoS se utilizó la cadena de búsqueda: (*PISA AND "secondary analysis"*) mientras que en Scopus se emplearon los mismos términos adaptados a las características de la base de datos: *TITLE-ABS-KEY (PISA AND "secondary analysis")*.

2.1.4. Proceso de selección

El proceso de búsqueda y selección de los artículos se llevó a cabo en dos etapas, una primera desarrollada en junio de 2018 y la segunda en febrero de 2019, cuyo objetivo fue actualizar la base de datos de manera que contuviera todas las publicaciones aparecidas hasta diciembre de 2018.

El proceso de selección estuvo dividido en tres fases. La primera fase (identificación) estuvo destinada a recopilar los resultados obtenidos en las dos fuentes seleccionadas y elaborar una base de datos empleando una hoja de cálculo donde se incluyeron todos los artículos¹⁰. Adicionalmente esta base de datos también se empleó para registrar todo el proceso de selección realizado en las fases siguientes.

¹⁰ La base de datos con todos los artículos incluidos y excluidos del análisis, así como el procedimiento de evaluación de la calidad y recopilación de la información puede consultarse en https://docs.google.com/spreadsheets/d/1IdlBypxDVe_bum94Qa53VDmwNOB41tLBGSDO6IBLpk/edit?usp=sharing.

Una vez registrados todos los documentos encontrados en los dos repositorios, se eliminaron los elementos duplicados como paso previo a aplicar los criterios de exclusión e inclusión. En esta fase se localizaron un total de 424 publicaciones entre los dos repositorios, 36 en WoS y 388 en Scopus. Tras la eliminación de los duplicados, el número final de artículos quedó reducido a 388.

En la segunda fase se revisó el título, resumen y palabras clave de las publicaciones para aplicar los criterios de inclusión, lo que permitió un primer filtro de los artículos que claramente no cumplían con los mismos. En caso de que la información de estos apartados no fuese suficiente para tomar una decisión, se optó por incluir los artículos en el siguiente paso y aplicar de nuevo los criterios de inclusión sobre el texto completo, junto con las preguntas de calidad. En esta fase de filtro se descartaron 313 registros, pasando 67 a la siguiente fase.

Por último, en la fase tres se revisó el texto completo de los 67 artículos aplicando los criterios de calidad y eliminando aquellos que no alcanzaran una puntuación suficiente. En consecuencia, 15 de los artículos fueron eliminados de la lista final en esta fase, siendo seleccionados 52 artículos para realizar el proceso de extracción de información y síntesis de resultados (figura 2.1). La referencia de los 52 artículos puede consultarse en el Anexo I.

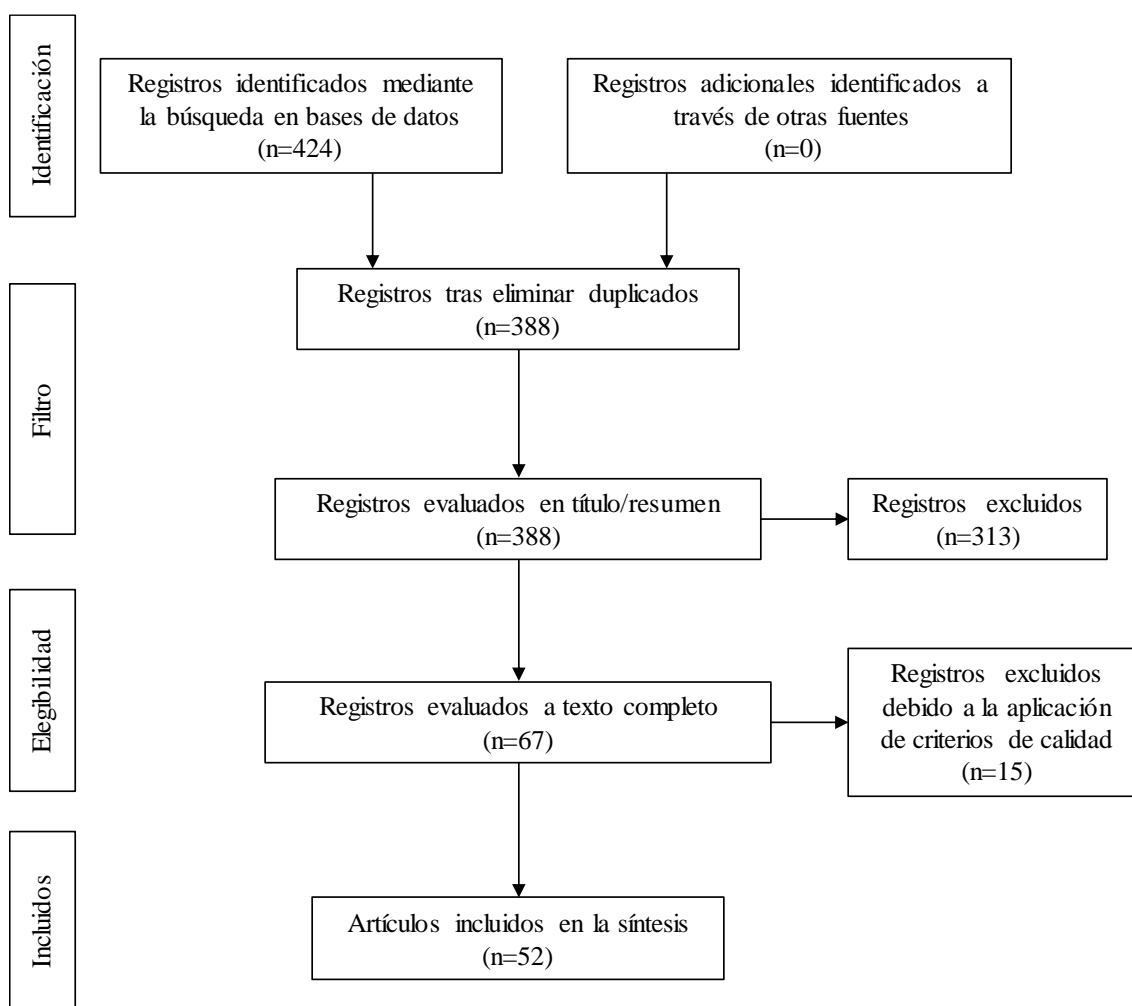


Figura 2.1. Proceso de selección.

Los 52 artículos resultantes de la selección final fueron sometidos a un procedimiento de vaciado de información, a través del cual se extrajeron de cada uno de ellos los datos necesarios para contestar a las preguntas de *mapping* y SLR planteadas.

2.2. RESULTADOS

A continuación, se da respuesta a las preguntas planteadas a través del análisis de los 52 artículos seleccionados. De esta manera, la presente sección se estructura en función de dichas preguntas, exponiéndose primero los resultados del estudio de *mapping* y después los de la SLR.

2.2.1. Resultados del estudio *mapping*

La primera pregunta del estudio de *mapping* está centrada en la evolución de volumen de producción científica relacionada con el análisis secundario de PISA. Para responder

dicha pregunta se ha realizado un análisis del número de publicaciones por año (figura 2.2).

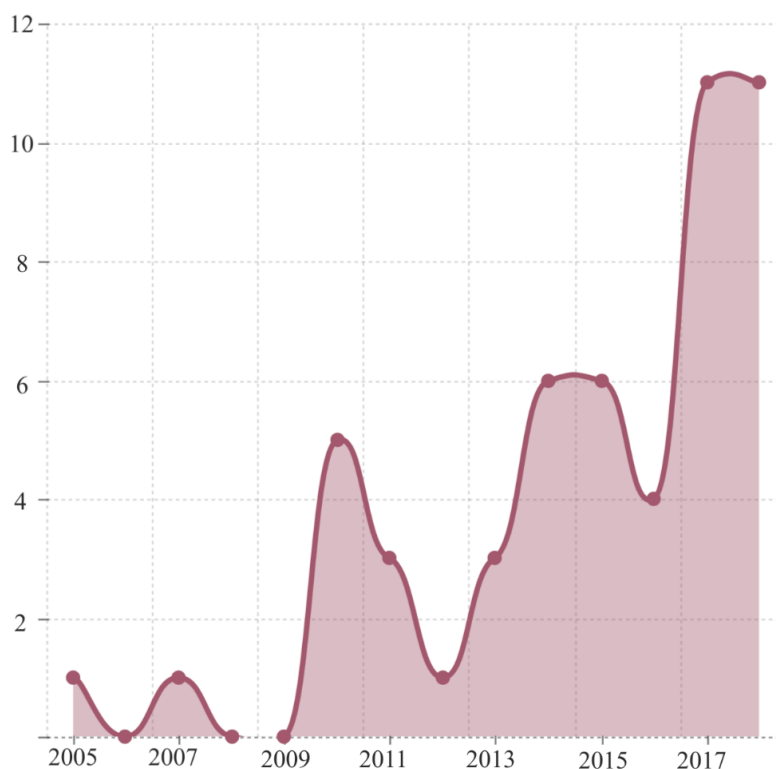


Figura 2.2. Número de publicaciones por año.

Como se puede ver en la figura 2.2, la primera publicación de relevancia que difunde los resultados de un análisis secundario de PISA es la realizada en el año 2005 por Güzel y Berberoğlu (2005). No obstante, este objeto de investigación no llama la atención de la comunidad científica hasta el año 2010, y hasta dicha fecha no se encuentra más que una publicación de relevancia (Maslowski, Scheerens y Luyten, 2007).

A lo largo de la última década existe un interés creciente en la realización de estudios secundarios de PISA, alcanzándose el número máximo de publicaciones por año en 2017 y 2018.

En cuanto a los autores de referencia en el campo (PM2), se constata que existe una gran variedad de autores, en total participaron 143 autores diferentes en las 52 publicaciones analizadas. En la mayoría de los casos se trata de autores que participaron en una sola obra, destacando McConney y Perry con 5 y 4 publicaciones respectivamente (tabla 2.3). Cabe señalar que las aportaciones realizadas por estos autores, cuyas publicaciones son conjuntas, consisten en su totalidad de análisis puramente descriptivos que no hacen uso de los pesos muestrales ni los valores plausibles, utilizando en su lugar

la media de todos los valores plausibles como variable criterio. Adicionalmente, varias de sus publicaciones presentan exactamente los mismos análisis con cambios mínimos, como la competencia analizada, el ciclo de PISA o el país de referencia. Por lo tanto, a pesar de que estos artículos hayan pasado los criterios de calidad establecidos para la presente revisión, se considera que sus aportaciones a la disciplina de investigación no son particularmente relevantes.

Tabla 1.3 Número de publicaciones por autor/a

Autor/a	Nº de publicaciones
McConney, A.	5
Perry, L.B.	4
Van de Vijver, F.J.R.	3
Acosta S.T., Cheema, J.R., Cheung K.-C., Ercikan K., Gil-Flores J., He J., Hsu H.-Y., Liu X., Luyten H., Lyons-Thomas J., Mak S.-K., Sandilands D., Scheerens J., Sit P.-S., Soh K.-C.	2

Al igual que sucede con los autores, también encontramos una gran diversidad en cuanto a las revistas empleadas para la difusión de las investigaciones, donde encontramos 33 revistas diferentes. Entre estos medios destacan las revistas *Studies in Educational Evaluation* y *Teachers College Record*, con cuatro publicaciones cada una, siendo además la primera la revista con el índice SJR más alto (1.92). En la Tabla 2.4 se puede consultar el nombre de las revistas con más de una publicación junto con el número de artículos, su índice SJR de acuerdo con SCIMAGO y el cuartil en el que se encuentran.

Tabla 2.4. Fuente de las publicaciones

Revista	Nº de pub.	SJR	Cuartil
<i>Teachers College Record</i>	4	1	Q1
<i>Studies in Educational Evaluation</i>	4	0.64	Q2
<i>International Journal of Science and Mathematics Education</i>	3	0.7	Q1
<i>Educational Research and Evaluation</i>	3	0.57	Q2
<i>Asia-Pacific Education Researcher</i>	3	0.42	Q2
Revista de Educación	3	0.35	Q3
<i>Computers and Education</i>	2	2.32	Q1
<i>School Effectiveness and School Improvement</i>	2	1.51	Q1
<i>International Journal of Educational Development</i>	2	0.79	Q1
<i>Journal of Educational Computing Research</i>	2	0.79	Q1
<i>Educational Studies</i>	2	0.44	Q2

Si se fija la atención en los países en los que se ha desarrollado un mayor número de investigaciones (PM4), se puede ver que los contextos más explorados son Australia (21), Alemania (17), Dinamarca, España, Corea del Sur y Estados Unidos (16), y Bélgica, Canadá, Finlandia, Países Bajos y Noruega (15). Como puede observarse en el mapa de la figura 2.3, las regiones más estudiadas son Norteamérica, Europa y Australia.

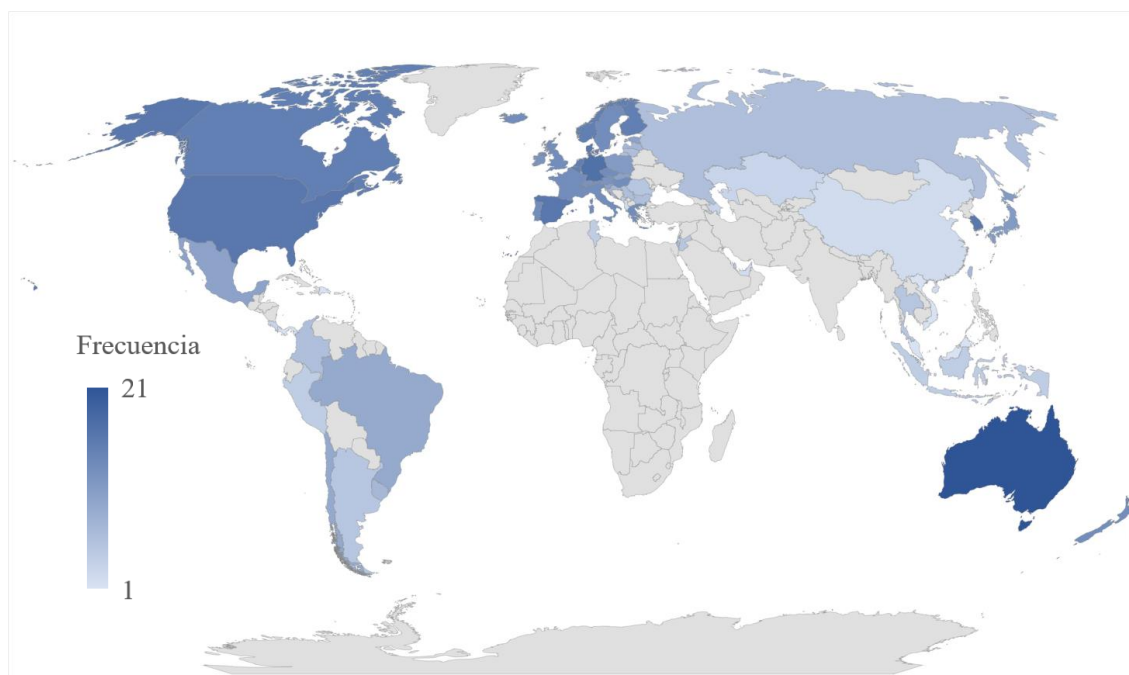


Figura 2.3. Número de publicaciones por país

La mayoría de los estudios (31) incluyen varios países para su investigación, utilizando diversos criterios de selección en función de sus objetivos. Algunos introducen los datos de todos los países disponibles en su correspondiente ciclo de PISA (Chiu, 2015; Sebastian & Huang, 2016; Täht, Must, Peets & Kattel, 2014; Willms, 2010). Otros se interesan por la comparación de los países que pertenecen a organizaciones territoriales, como la OCDE (Condrón, 2011; Han, 2018; Maslowski et al., 2007) o la Unión Europea (Rodríguez-Santero & Gil-Flores, 2018). Por último, hay algunos estudios que agrupan países en función de su cercanía geográfica o cultural, como los países del continente americano (Milford, Ross & Anderson, 2010), del norte de Europa (Sortkær & Reimer, 2018) o algunos países de la región de Asia Oriental, como Taiwan, Hong Kong, Japón, Corea del Sur, Macao y China (Chiao & Chiu, 2018; Cheung, Sit, Soh, Jeong & Mak, 2013; Kim & Law, 2012).

Por su parte, cabe destacar también aquellos países que más veces han sido investigados en solitario, que son Australia, con 4 estudios (Gabriel, Signolet & Westwell,

2017; McConney & Perry, 2010a, 2010b; Perry & McConney, 2010), Estados Unidos, también con 4 estudios (Barnard-Brakk, Lan & Yang, 2018; Liu & Whitford, 2011; Kim, 2017; Rutkowski, Rutkowski, Wild & Burroughs, 2017), y España, con 3 estudios (Cordero, 2015; Gil-Flores & García-Gómez, 2017; Tourón, López-González, Lizasoain, García & Navarro, 2018).

La pregunta de *mapping* número 5 está centrada en las competencias más empleadas como variables criterio a lo largo de los ciclos (tabla 2.5). En dicha tabla se puede comprobar cómo generalmente se produce un incremento en el número de investigaciones que emplean una determinada competencia coincidiendo con el año en que dicha competencia es designada como principal en el ciclo de PISA. Así mismo se constata que la competencia en matemáticas es utilizada como variable criterio en un mayor número de investigaciones, siendo además la única que se utiliza en solitario como variable criterio en ciclos en los que no es la competencia principal (Kim & Law, 2012; Chiu, 2015; Zhu, Kaiser & Cai, 2018).

Tabla 2.5. Competencias analizadas como variable criterio en función del ciclo de PISA

Ciclo de PISA	Lectura	Matemáticas	Ciencias	Competencia principal
2000	3	2	1	Lectura
2003	2	4	2	Matemáticas
2006	4	7	14	Ciencias
2009	8	4	3	Lectura
2012	3	20	4	Matemáticas
2015	3	4	7	Ciencias
Total	23	41	31	

En cuanto a los ciclos de PISA más comúnmente analizados (PM6), podemos comprobar que el más analizado es el del año 2012, seguido del ciclo de 2006, aunque todos los ciclos han sido analizados en al menos una publicación en los últimos dos años. Cabe destacar además que en los últimos años han surgido una serie de estudios multiciclo (Zhou, Fan, Wei X y Tai, 2017; He, Jan de Vijver y Kulikova, 2017; Arikan, van de Vijver y Yagmur, 2017; Hwang, Choi, Bae y Shin, 2018) que analizan la evolución histórica de las puntuaciones a lo largo de diversos ciclos.

Tabla 2.6. Número de publicaciones por ciclo de PISA

Año de publicación	Ciclo de PISA					
	2000	2003	2006	2009	2012	2015
2005	1					
2006						
2007	1					
2008						
2009						
2010		2	3			
2011			3			
2012			1			
2013			2	1		
2014			5	1		
2015				4	2	
2016				1	3	
2017	1	2	1	2	9	3
2018				1	6	5
Total	3	4	15	10	20	8

Por último, para conocer las metodologías más empleadas para el análisis de datos en las publicaciones se ha empleado la clasificación propuesta por Liou y Hung (2015), que clasifican las metodologías de análisis en tres categorías:

- Básicas: Análisis descriptivos, correlación, ANOVA, Chi-cuadrado, prueba t.
- Intermedia: Regresión múltiple, comparación múltiple post hoc.
- Avanzada: Correlación canónica, análisis factorial, análisis de discriminación, análisis *cluster*, modelos de ecuaciones estructurales (SEM), regresión logística, modelos multinivel y teoría de respuesta al ítem (TRI).

Esta clasificación fue ampliada incluyendo el uso de *big data* y *machine learning* (BD/ML) dentro de los métodos avanzados de análisis (tabla 2.7).

Como se puede observar en la tabla 2.7, los 52 artículos combinan análisis básicos, especialmente descriptivos, con análisis más avanzados. De entre estas últimas, destaca el uso de modelos multinivel en 22 de las 52 publicaciones, seguido del empleo de los modelos de ecuaciones estructurales, que se utilizan en 10 de los 52 estudios.

Tabla 2.7. Número de publicaciones por metodología empleada

Nivel	Técnica	n
Básico	Descriptivo	52
	Correlacional	16
	ANOVA	1
	Chi-cuadrado	2
	Prueba t	6
Intermedio	Regresión múltiple	9
	Comparación múltiple post hoc	-
Avanzado	Correlación canónica	1
	Análisis factorial	8
	Análisis de discriminación	-
	Análisis <i>cluster</i>	2
	SEM	10
	Regresión logística	6
	Modelos multinivel	22
	TRI	1
	BD/ML	3
Otros	4	

Para finalizar la parte del *mapping*, en la tabla 2.8 se ofrecen las principales conclusiones extraídas del análisis de cada una de las preguntas planteadas.

Tabla 2.8. Resumen de las respuestas a las preguntas del *mapping*

Pregunta	Respuesta
PM1	La primera investigación se publica en 2005, existe un interés creciente en este tema de investigación, especialmente a lo largo de la última década.
PM2	Los autores más activos son McConney y Perry.
PM3	Las revistas que han sido empleadas para difundir un mayor número de investigaciones son <i>Studies in Educational Evaluation</i> y <i>Teachers College Record</i> .
PM4	Los países más estudiados son Australia (21); Alemania (17); Dinamarca, España, Corea del Sur y Estados Unidos (16); y Bélgica, Canadá, Finlandia, Países Bajos y Noruega (15).
PM5	La competencia en matemáticas es la que atrae el interés de un mayor número de investigaciones
PM6	Los ciclos de PISA más analizados son los del 2012 y 2006, aunque todos los ciclos han sido analizados en los últimos dos años.
PM7	Los modelos multinivel son la metodología de análisis avanzado más empleada.

2.2.2. Resultados de la SLR

La primera pregunta de la SLR está centrada en las variables predictoras más utilizadas en el estudio del nivel competencial de los estudiantes participantes en PISA. El gran

número de variables disponibles en los cuestionarios de contexto de PISA, así como el elevado volumen de artículos incluidos en la presente SLR, dificulta la tarea de enumerar todas las variables utilizadas, por lo que en esta sección se informa sólo de aquellas variables con una frecuencia de uso más alta.

En cuanto a las variables referidas a los alumnos, las más utilizadas son aquellas que tienen que ver con su contexto socioeconómico y demográfico, como el género, el nivel socioeconómico (ESCS), o el estatus migratorio. Otras variables de interés son aquellas relacionadas con el uso de las TIC, y algunas variables relacionadas con las tres competencias evaluadas por PISA, como la autoeficacia percibida, el interés o las actitudes mostradas por los estudiantes hacia dichas materias (tabla 2.9)

Tabla 2.9. Variables predictoras a nivel de alumno más estudiadas

Tema	Variable	n
Contexto	ESCS	32
	Género	21
	Estatus migratorio	14
	Lengua hablada en el hogar	7
Materia	Matemáticas - autoeficacia	6
	Ciencias - interés	5
	Ciencias - autoeficacia	4
	Ciencias - valor	4
	Ciencias - disfrute	4
TIC	Uso de las TIC	9

Por su parte, las variables a nivel de centro que más se utilizan (tabla 2.10) también son las de carácter contextual, destacando el ESCS agregado por centros, el tamaño del centro, su titularidad y los recursos educativos de los que dispone. Entre las variables que informan sobre los procesos organizativos y educativos del centro, las más utilizadas son la medida de la autonomía escolar y la información sobre el liderazgo educativo ejercido en el centro.

Tabla 2.10. Variables predictoras a nivel de centro más estudiadas

Tema	Variable	n
Contexto	ESCS	14
	Tamaño de centro	12
	Titularidad	10
	Recursos educativos	10
	Ratio profesor/alumnos	7
Proceso	Autonomía escolar	6
	Liderazgo educativo	4

Por último, algunos estudios realizan comparaciones entre diversos países, y para ello introducen variables a nivel de país en el análisis. Las variables nacionales más utilizadas son el Producto Interior Bruto (PIB) per cápita (6 estudios), el coeficiente Gini, que mide la desigualdad de los ingresos de los ciudadanos de un país (3 estudios), y el gasto educativo por alumno y el Índice de Desarrollo Humano, ambos incluidos en dos estudios cada uno.

En segundo lugar, la PI2 hace referencia a la frecuencia con la que los autores tienen en cuenta la estructura jerárquica de los datos. A este respecto existe una fuerte división en las investigaciones consultadas dado que en 26 de las 52 publicaciones no se ha tenido en cuenta esta estructura, frente a 22 casos en los que sí. De los cuatro casos restantes en uno se ha tenido la estructura jerárquica parcialmente en cuenta (Smith, Cheema, Kumi-Yeboah, Warrican y Alleyne, 2009) y en los tres casos restantes no se especifica con claridad (Gabriel, Signolet y Westwell, 2012; Chiao y Chiu, 2012; Zhou, Fan, Wei y Tai 2017).

Pasando a la pregunta de investigación PI3 (¿Con qué frecuencia utilizan los autores los valores plausibles en los cálculos relacionados con la variable criterio?), tras revisar las publicaciones se comprueba que en la mayoría de los casos (35) se hace un uso adecuado de los valores plausibles utilizando estos de manera simultánea.

Sin embargo, en las 17 investigaciones restantes el uso de estos valores no se reporta o se realiza de manera incorrecta. De estas, en dos casos se ha optado por utilizar solamente uno de los diez valores plausibles (Gabriel, Signolet y Westwell, 2017; Cheema, 2017), mientras que en cinco (McConney y Perry, 2010a; McConney y Perry 2010b; Perry, McConney, 2010; Perry y McConney, 2013; Woods-McConney, Oliver, McConney, Maory Schibeci, 2013) se ha calculado la media entre los valores, lo que ocasiona la pérdida de la variabilidad (cabe destacar que las cinco publicaciones participa McConney como autor). En los diez casos restantes no se ha explicado claramente cuál de los dos métodos anteriores se ha utilizado, aunque resulta claro que no han tratado los valores plausibles de manera simultánea.

En cuarto lugar, la pregunta de investigación PI4 hace referencia al tratamiento de los pesos muestrales. En este caso, sorprende que en 23 de las 52 publicaciones no se incluye ningún tipo de peso muestral, o al menos no se reporta debidamente su uso. Entre las investigaciones que sí incluyen estos pesos, la mayoría (10) utiliza el peso de los

estudiantes con las 80 réplicas repetidas balanceadas (ver sección 1.3.1), mientras que en 8 ocasiones se utilizan los pesos de los estudiantes sin incluir las réplicas. El peso de los centros solamente es incluido en tres de las investigaciones (Milford, Ross y Anderson, 2010; Sebastian y Huang, 2016; Eickelmann, Gerick y Koop, 2016), en todos los casos combinados con los pesos de los estudiantes.

En las 8 investigaciones restantes se indica que se incluyen los pesos de los estudiantes, pero no se especifica la manera en la que se lleva a cabo este proceso.

Por último, la pregunta de investigación 5 se centra en las limitaciones teóricas o metodológicas del uso de los datos de PISA que más frecuentemente se reportan en la literatura. El objetivo de esta pregunta es complementar la sección de críticas a PISA presentada en el capítulo anterior (sección 1.4) con cuestiones metodológicas más concretas y basadas en datos. De los 52 estudios, 21 no reportan ninguna limitación a la hora de tratar con estos datos. Las limitaciones más mencionadas en los 31 artículos restantes son el carácter transversal de la prueba, que no permite evaluar el progreso de los estudiantes ni tener en cuenta su nivel de rendimiento previo (16 estudios) y la imposibilidad de establecer relaciones causales entre las variables de estudio, señalada en 15 investigaciones.

La validez de las respuestas a los cuestionarios de contexto es cuestionada por seis estudios (Cheema, 2017; Chi, Liu, Wang & Won, 2018; Dumay & Dupriez, 2014; Petko, Cantieni & Prasse, 2016; Sandilands, Barclay, Lyons-Thomas & Ercikan, 2014; Tourón et al., 2018) debido a la dependencia de autoinformes. Los autores de dichos estudios aconsejan cautela a la hora de operar con variables extraídas de estos instrumentos.

Otra limitación reportada con frecuencia es la falta de información relevante para los análisis que los investigadores se proponen llevar a cabo. La primera es la falta de información a nivel de aula, el cual se considera un nivel fundamental para poder sacar conclusiones relacionadas con las prácticas educativas que influyen en el rendimiento (Eickelmann et al., 2016; Lafontaine, Baye, Bieluf & Monseur, 2015; Luu & Freeman, 2011; Sandilands et al., 2014; Scheerens, Luyten, van den Berg & Glas, 2015). Otros autores acusan la falta de variables de interés para sus investigaciones, como por ejemplo datos sobre el manejo de TIC por parte del profesorado (Hu, Gong, Lai & Leung, 2017), la edad de llegada al país de los estudiantes inmigrantes de primera generación (Arikan et al., 2016), la creatividad relacionada con las matemáticas (Sebastian & Huang, 2016)

o el nivel de competencia en lengua escrita dentro de la evaluación de la competencia lectora (Ercikan et al., 2015).

Por último, algunos autores incluyen como limitación los sesgos metodológicos inherentemente relacionados con las evaluaciones a gran escala, los cuales pueden abordarse a través del uso de los valores plausibles y los pesos muestrales (Acosta & Hsu, 2014a, 2014b).

Para finalizar esta sección se recoge un resumen de las respuestas a las cinco preguntas de investigación planteadas en esta SLR (tabla 2.11).

Tabla 2.11. Resumen de las respuestas a las preguntas de la SLR

Pregunta	Respuesta
PI1	Las variables predictoras más utilizadas son de carácter contextual, tanto a nivel de alumno (ESCS, género, estatus migratorio) como de centro (ESCS, tamaño, titularidad).
PI2	En la mitad de las publicaciones (26) no se ha tenido en cuenta la estructura jerárquica de los datos.
PI3	En la mayoría de las investigaciones (35) se hace un uso adecuado de los valores plausibles, en el resto de los casos se utiliza sólo uno de los valores o se utiliza la media entre los 10.
PI4	En casi la mitad de las investigaciones no se incluyen los pesos muestrales. En las que sí se incluyen lo más habitual es utilizar los pesos de los estudiantes con las 80 réplicas repetidas balanceadas
PI5	Las limitaciones más frecuentemente mencionadas son la naturaleza transversal de PISA y la imposibilidad de establecer relaciones causales entre las variables predictoras y criterio.

2.3. CONCLUSIONES

Antes de exponer las principales conclusiones extraídas del análisis de los artículos incluidos en este estudio, resulta necesario discutir algunas limitaciones derivadas del proceso. En primer lugar, a pesar de que los repositorios seleccionados para llevar a cabo la revisión son los más representativos en investigación, no recogen todos los artículos publicados sobre el tema de interés, por lo que quizá podría ampliarse la búsqueda incluyendo otros repositorios más comprensivos como Google Scholar.

Por otro lado, también es necesario abordar dos sesgos relacionados con el proceso de revisión sistemática. En primer lugar, la subjetividad del revisor puede afectar a la selección de artículos y la aplicación de los criterios de calidad. Sin embargo, el hecho de

exponer de manera pública todo el procedimiento de búsqueda y selección dota de transparencia al proceso y garantiza su reproducibilidad. En segundo lugar, cabe considerar también el sesgo de publicación, el cual puede producir una sobrerrepresentación de artículos con resultados positivos (Egger, Smith, Schneider, & Minder, 1997).

De la revisión de la literatura realizada pueden extraerse conclusiones de diversa naturaleza. En primer lugar, se visualiza claramente un aumento del interés de la comunidad investigadora por realizar estudios secundarios con los datos de PISA en la última década, y en especial en los últimos cinco años (figura 2.2). Sin embargo, no se detectan autores especialmente interesados en la materia, ya que la mayoría de los autores presentan sólo una o dos publicaciones, encontrándose sólo tres que superan las dos publicaciones.

Por otro lado, pueden comentarse varias prácticas investigadoras relacionadas con la metodología empleada para realizar estos análisis. Aunque existe una variedad amplia de técnicas estadísticas para el análisis de los datos de PISA, destaca el uso de los modelos multinivel como técnica avanzada más utilizada (42% de los estudios), seguida por los modelos de ecuaciones estructurales, utilizados en un 19% de las investigaciones. Este hallazgo cobra sentido si tenemos en cuenta que la estructura jerárquica de los datos de PISA es una de las principales cuestiones de diseño que hay que tener en cuenta a la hora de escoger la técnica adecuada para abordar su análisis, lo cual conduce a otra consideración relevante. Sólo la mitad de las investigaciones (26) han tenido en cuenta la estructura anidada de los datos, a pesar de las numerosas advertencias recogidas en la literatura sobre la no consideración de esta característica, que puede resultar en errores tanto de estimación como de interpretación de los resultados (Snijders & Bosker, 2012).

Otras consideraciones metodológicas de gran relevancia en el trabajo con datos procedentes de evaluaciones a gran escala, y de PISA particularmente, es el tratamiento de los valores plausibles y los pesos muestrales. En la mayoría de las investigaciones (67%) se detecta un uso adecuado de los valores plausibles, consistente en la aplicación de la técnica estadística con cada uno de ellos y el promedio de sus resultados. Sin embargo, en algunos estudios se trabaja con el promedio de los valores, o sólo con uno de ellos, lo cual puede conllevar una computación incorrecta de los errores de medida, así como de los errores típicos de estimación en diseños muestrales complejos (Wu, 2005). El uso de los pesos muestrales presenta más división entre las investigaciones

seleccionadas, ya que en el 44% de ellas no se utilizan (o no se reporta su uso), pudiendo incurrir por tanto en una incorrecta representatividad de la muestra (ver sección 1.3.1).

Estas cuestiones evidencian que existe cierta falta de rigor científico en el campo de los análisis secundarios de PISA, y es necesario un mayor esfuerzo de los investigadores por cumplir con los requerimientos metodológicos derivados del uso de evaluaciones a gran escala con diseños complejos.

En relación con las variables más frecuentemente investigadas, las variables contextuales son las que se utilizan con más frecuencia tanto a nivel de estudiante como de centro, mostrando una menor explotación de las variables relacionadas con aspectos no cognitivos de los estudiantes o cuestiones pedagógicas y organizativas de los centros. Por otro lado, existen multitud de estudios (31) que incluyen varios países en su análisis (algunos de ellos comparan hasta 65 países), y sin embargo llama la atención la falta de análisis de variables que exploren las características de los sistemas educativos. En la mayoría de los estudios que incluyen variables a nivel de país, estas se relacionan con cuestiones económicas (PIB per cápita, gasto en educación, etc.) o con indicadores de igualdad, pero pocos utilizan indicadores descriptivos del sistema educativo, como la comprensividad, la formación inicial del profesorado, la segregación de estudiantes en función de su rendimiento académico o de su nivel socioeconómico (Chiu, 2015), o las políticas de elección de centro educativo (Scheerens et al., 2015). Estas cuestiones podrían aportar información valiosa acerca de las políticas de los sistemas educativos y su relación con la existencia de brechas de rendimiento en función del género, nivel socioeconómico o estatus migratorio de los estudiantes, y podrían suponer una guía en relación con políticas que fomenten la equidad.

Por último, el análisis de las limitaciones mencionadas en los estudios revisados revela algunos de los principales inconvenientes de PISA derivados de su estructura. El principal es la falta de datos longitudinales o de rendimiento previo, la cual dificulta el establecimiento de relaciones causales entre las variables de estudio y el nivel competencial de los estudiantes. Otras cuestiones mencionadas por los autores son la falta de datos a nivel de aula y la cuestionable validez de la información recogida en los cuestionarios de contexto a causa de la dependencia de autoinformes. La sección de conclusiones y discusión del presente trabajo abordará todas estas limitaciones, así como sus implicaciones con respecto a los resultados obtenidos.



3. La investigación sobre eficacia escolar

La investigación sobre eficacia escolar

Los últimos cincuenta años han supuesto el auge de una nueva temática de interés dentro de la investigación en educación: la eficacia educativa. El estudio de este fenómeno se ha realizado desde diversas tradiciones y perspectivas investigadoras que pueden resumirse principalmente en cinco (Scheerens y Bosker, 1997):

- **Investigación sobre la igualdad de oportunidades educativas y el papel de la escuela:** se centra en las limitaciones que el entorno socioeconómico del estudiante y el centro impone sobre el nivel educativo que este es capaz de alcanzar.
- **Estudios económicos sobre las funciones productivas de la educación:** el interés de esta disciplina es investigar cómo influyen en el producto de la educación (resultados académicos) los recursos invertidos en ella.
- **Evaluación de programas compensatorios:** esta perspectiva trata de evaluar hasta qué punto el cambio del entorno escolar que suponen los programas compensatorios contribuye a una mejora en el rendimiento de los grupos de estudiantes desfavorecidos a los que van dirigidos.
- **Estudios sobre escuelas eficaces y la evaluación de programas de mejora escolar:** centrados en el análisis de las prácticas organizativas y pedagógicas de las escuelas en las que se ha encontrado un nivel inusual de eficacia (una diferencia relevante entre el rendimiento real de los estudiantes y aquel que se espera de ellos en función de su entorno socioeconómico).
- **Estudios sobre la eficacia de los profesores, las clases y los procesos instructivos:** de manera similar a la anterior, esta perspectiva se centra en el estudio de prácticas asociadas a la eficacia, pero se centra en el nivel de aula (prácticas docentes, metodologías didácticas, etc.).

Cada una de estas perspectivas de investigación tiene sus características propias, como las principales variables de estudio que utilizan, la disciplina desde la que se suele adoptar este enfoque, o el principal tipo de estudio que se lleva a cabo (tabla 3.1).

Tabla 3.1. Características de los tipos de investigación en eficacia educativa.

Corriente	Tipo de variable independiente	Tipo de variable dependiente	Disciplina	Principal tipo de estudio
Igualdad de oportunidades	Estatus socioeconómico, CI de los estudiantes, recursos materiales de la escuela	Nivel educativo alcanzado por los estudiantes	Sociología	Encuesta
Funciones productivas	Recursos materiales de la escuela	Rendimiento	Economía	Encuesta
Evaluación de programas compensatorios	Currículo específico	Rendimiento	Pedagogía interdisciplinaria	Cuasi-experimental
Eficacia escolar	Características procesuales de la escuela	Rendimiento	Pedagogía interdisciplinaria	Estudio de casos
Instrucción eficaz	Características de los docentes, la instrucción, la organización de las clases	Rendimiento	Psicología educativa	Experimental, observación

Fuente: Scheerens y Bosker (1997, p. 140)

Aunque estas perspectivas no son excluyentes, y pueden encontrarse estudios que combinan dos o más de ellas, este capítulo va a centrarse principalmente en dos: el enfoque de la eficacia escolar, ya que es el que guiará la parte de la tesis relacionada con el estudio de la eficacia, y los estudios sobre la eficacia de los profesores, debido a que es el enfoque que más comúnmente se combina con la eficacia escolar, con el objetivo de añadir los procesos del nivel de aula.

La investigación en eficacia escolar (IEE) se define en la literatura como:

Conjunto de estudios empíricos que tienen por objeto la estimación de la magnitud de los efectos escolares y el análisis de sus propiedades científicas, y/o el estudio de los factores escolares, de aula y de contexto que caracterizan una escuela eficaz, sea cual sea el enfoque metodológico utilizado para conseguirlo (Murillo, 2005, p. 22).

Teniendo esto en cuenta, cabe también definir el concepto de escuela eficaz, que según el mismo autor es aquella escuela que “consigue un desarrollo integral de todos y cada uno de sus alumnos mayor de lo que sería previsible, teniendo en cuenta su

rendimiento previo y el nivel social, económico y cultural de las familias” (Murillo, 2005, p. 30). Esta definición establece tres conceptos fundamentales que sirven como pilares para la investigación relacionada con este tema (Murillo, 2003):

- **Equidad:** reflejada en la búsqueda del desarrollo de todos los estudiantes, independientemente de las características socioeconómicas de su entorno de procedencia o de la escuela a la que pertenecen.
- **Desarrollo integral:** además del desempeño académico, la IEE busca también la promoción de los valores y el bienestar estudiantil.
- **Valor añadido:** concepto que supone la inclusión de factores contextuales en el análisis del rendimiento académico.

En este capítulo se presentará brevemente la evolución de la eficacia escolar como campo de investigación educativa, así como los principales estudios realizados en el contexto internacional y nacional.

3.1. EVOLUCIÓN DE LA DISCIPLINA

La historia de la investigación en eficacia escolar, desde su inicio hasta nuestros días, puede dividirse en cuatro fases (figura 3.1) diferenciadas en función de la década y las tendencias en cuanto a los enfoques teóricos, temas de interés y metodologías utilizadas. Esta sección realiza un recorrido por dichas fases, destacando aquellos trabajos de mayor relevancia para la evolución y el asentamiento de la disciplina.

IEE

1966 - ACTUALIDAD

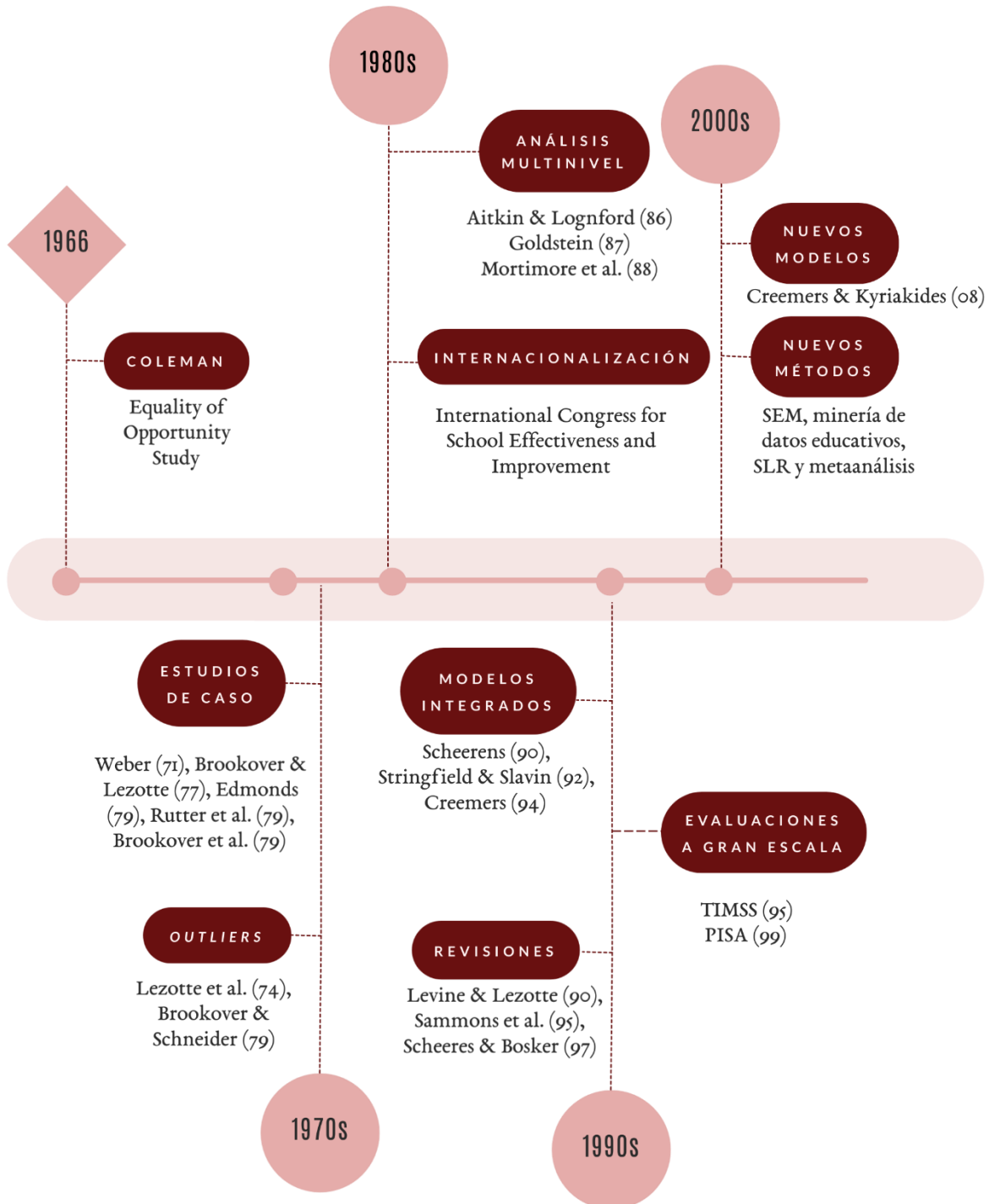


Figura 3.1. Hitos principales de la investigación en eficacia educativa.

3.1.1. Años 70: Los orígenes de la IEE

Los orígenes del estudio de la eficacia escolar se remontan a la publicación del documento *Equality of Educational Opportunity Study*, coloquialmente conocido como el informe Coleman (Coleman, 1966). El objetivo de este estudio, en el que participaron alrededor de 4000 escuelas y 600000 estudiantes, era analizar la igualdad de oportunidades educativas para los estudiantes estadounidenses. Su principal conclusión fue que las inequidades encontradas en los resultados académicos estaban basadas en características socioeconómicas y demográficas de los estudiantes (raza, nivel socioeconómico familiar, etc.). El análisis realizado para el informe encontró que había pocas diferencias entre las escuelas, e incluso que las diferencias existentes (disponibilidad de algunos recursos, cualificación del personal docente, entre otras) no suponían una diferencia apreciable en el rendimiento de los estudiantes (Hill, 2017). Unos años más tarde, los resultados de este informe fueron respaldados por el reanálisis de los datos llevado a cabo por Jencks et al. (1972), que llegó a las mismas conclusiones.

Estos hallazgos sobre el exiguo efecto de las escuelas sobre el rendimiento de los estudiantes en comparación con la incidencia de los factores económicos y sociales generaron un fuerte impacto en la comunidad política y científica (Chapman, Muijs, Reynolds, Sammons, & Teddlie, 2016), impulsando el desarrollo de una nueva línea de investigación para profundizar en su estudio.

Uno de los primeros ejemplos de trabajos de investigación asociados a la línea de la eficacia escolar es el llevado a cabo por Weber (1971). El estudio de este autor estuvo motivado por su observación de que, en general, las escuelas situadas en el centro de las ciudades grandes y con una población estudiantil eminentemente pobre producían aprendizajes de la lectura más bajos que otras escuelas, pero que había algunas escuelas urbanas con estas características que igualaban o incluso superaban en nivel lector a otras escuelas más favorecidas. Las explicaciones que hasta el momento se habían dado para estas disparidades, como la financiación de los centros (Sexton, 1969), la inteligencia de los estudiantes (Jensen, 1969) o incluso la del propio Coleman, no satisficieron del todo a Weber, que decidió emprender un estudio propio en cuatro escuelas pobres de ciudad cuyas puntuaciones en lectura superaban sus expectativas. La hipótesis y propósitos del estudio reflejan adecuadamente los principios de la investigación en eficacia escolar:

Accordingly, I developed a hypothesis: that several inner-city public schools exist in the United States where reading achievement in the early grades is far higher than in most inner-city schools, specifically, is at the national average or higher. A study to investigate this hypothesis would have two purposes. If the hypothesis proved correct, the study would show that inner-city children can be taught reading well, and it might discover some common factors in the success of the good programs. (Weber, 1971, p. 2)

Los criterios de selección de escuelas para este estudio fueron el nivel lector de los estudiantes y su nivel socioeconómico, con el objetivo de estudiar en profundidad aquellas escuelas que probasen su eficacia en la enseñanza de la lectura. Una vez seleccionadas, Weber realizó un proceso de observación y entrevistas para estudiar la incidencia de factores como la organización del currículo, las metodologías docentes, las prácticas de liderazgo o el clima escolar, introduciendo de esta manera factores de proceso en un campo que hasta ese momento había considerado principalmente factores de tipo contextual.

En esta primera fase de la IEE surgieron algunos otros estudios, clasificados por Purkey y Smith (1983) como estudios de casos atípicos – *outliers* – y estudios de caso. Los análisis de caso, como el llevado a cabo por Weber, pretendían extraer conclusiones sobre factores escolares asociados a la eficacia a través del estudio del rendimiento académico (Brookover & Lezotte, 1977; Edmonds, 1979), o también de otro tipo de factores de resultado, como la asistencia y el comportamiento (Rutter, Maughan, Mortimore & Ouston, 1979) o el autoconcepto y la autonomía (Brookover, Beady, Flood, Schewitzer & Wisenbaker, 1979).

Brookover y Lezotte (1977) realizaron un estudio de caso en ocho escuelas de educación primaria en las que se observaron puntuaciones crecientes en las habilidades lectoras de los estudiantes, y se identificaron algunos factores que diferenciaban dichas escuelas de aquellas con una tendencia contraria en las puntuaciones en lectura (tabla 3.2).

Por su parte, tras una serie de estudios sobre factores de eficacia escolar a lo largo de la década de los 70, Edmonds (1979) publicó una investigación basada en cinco escuelas primarias en las que el rendimiento de los estudiantes había probado ser independiente de su nivel socioeconómico, obteniendo rendimientos medios superiores a otras escuelas con

la misma composición social. Sus principales conclusiones sobre los factores de proceso asociados a la eficacia se resumen en cinco características, que pueden observarse en la tabla 3.2.

Uno de los estudios más influyentes de esta época en Reino Unido fue el llevado a cabo por Rutter et al. (1979), titulado *Fifteen thousand hours* (15.000 horas) debido al tiempo estimado que un estudiante pasa en la escuela a lo largo de su vida académica. Durante tres años, el equipo de Rutter estudió los resultados de los estudiantes y el funcionamiento de 12 escuelas urbanas de educación secundaria de Londres caracterizadas por atender a estudiantes de entornos desfavorecidos. Los indicadores de resultado del alumnado recogidos fueron la asistencia, el nivel de logro académico y el comportamiento (incluyendo la actividad delictiva). Aunque estos indicadores fueron ajustados para reflejar las variaciones en algunos factores de entrada de los estudiantes (nivel previo de razonamiento verbal, ocupación de los padres), los autores percibieron diferencias entre las escuelas que no podían ser explicadas por sus características físicas ni por sus recursos. Tras examinar la relación entre estas diferencias y los factores de proceso de las escuelas, el estudio concluyó que las escuelas que mejores resultados producían parecían promover en mayor medida la participación de los estudiantes tanto dentro de las clases como fuera de ellas, y presentaban unas expectativas más positivas acerca del rendimiento de los alumnos, proporcionándoles un entorno más agradable y motivante (Reynolds, Hargreaves & Blackstone, 1980).

Otro estudio destacable dentro de esta década es el de Brookover et al. (1979), centrado en estudiar las percepciones de estudiantes, docentes y equipo directivo acerca del clima social del centro en escuelas de educación primaria, incluyendo cuestiones como el autoconcepto de los estudiantes, las expectativas de los docentes, o la orientación del centro hacia los procesos de mejora del rendimiento.

Como puede observarse en la tabla 3.2, los principales factores que emergieron como relevantes de manera consistente en esta época fueron el clima de disciplina entre el alumnado, el liderazgo fuerte del equipo directivo, las altas expectativas de los docentes con respecto a los estudiantes, las políticas educativas orientadas al rendimiento y el tiempo dedicado a las tareas (Scheerens & Bosker, 1997).

Tabla 3.2. Principales conclusiones de los estudios de la década de los 70, agrupados por tema.

Factor	Weber (1971)	Brookover & Lezzote (1977)	Edmonds (1979)	Rutter et al. (1979)	Brookover et al. (1979)
Objetivos de la escuela	-	Énfasis en los objetivos de aprendizaje de lectura y matemáticas	Énfasis en la adquisición de competencias básicas	-	Compromiso con el rendimiento de los estudiantes
Expectativas	Altas expectativas	Creencia por parte de los profesores de que los estudiantes adquirirán las competencias básicas	Expectativas positivas del rendimiento del alumnado que impiden que los alumnos rindan por debajo del mínimo establecido	Actitud general de expectativas sobre el rendimiento de los estudiantes unida a prácticas que demuestren estas expectativas	Expectativas del rendimiento de los estudiantes
		Altas expectativas del logro educativo de los estudiantes			
Organización del personal y el horario	Más personal dedicado a la enseñanza de la lectura	Más tiempo empleado en la instrucción de la lectura	-	-	Más tiempo dedicado a la instrucción
Clima	Ambiente de orden, determinación y placer en la lectura	-	Ambiente disciplinado, pero no rígido	-	Buen clima social
Evaluación	Evaluación detallada del progreso de los estudiantes	-	Monitorización frecuente del progreso estudiantil, desde controles sobre el aprendizaje realizado en clase hasta pruebas estandarizadas	-	-
Liderazgo	Fuerte liderazgo	-	Fuerte liderazgo administrativo que aúna todos los elementos de la educación eficaz	-	Liderazgo del equipo directivo
Prácticas educativas	Individualización del aprendizaje	Programas de educación compensatoria en los que el profesorado se implique en la detección de necesidades	-	-	Prácticas de refuerzo positivo del aprendizaje

Fuente: elaboración propia

Los estudios basados en *outliers* (Brookover & Schneider, 1975; Lezzote, Edmonds & Ratner, 1974) compartían algunas características entre ellos, como el uso de regresión para la selección de escuelas o el estudio de un número parecido de escuelas primarias (24 y 20 respectivamente), y también coincidieron en algunos factores relevantes como el liderazgo, las expectativas de los docentes o la disciplina. Sin embargo, la calidad metodológica de estos estudios y sus conclusiones varían de manera considerable (Purkey & Smith, 1983).

Como puede observarse, esta primera fase de la IEE se caracterizó fundamentalmente por la proliferación de estudios empíricos de los factores de eficacia escolar como reacción a las conclusiones del informe Coleman.

3.1.2. Años 80: Consolidación de la disciplina

En los años 80 comenzaron a surgir técnicas estadísticas más complejas que se adaptaban mejor a las características de los datos procedentes de los centros educativos, y permitían realizar análisis más ajustados a realidad. Ejemplo de este fenómeno son los trabajos de Aitkin y Longford (1986), Raudenbush y Bryk (1986) o Goldstein (1987), que impulsaron el uso de los modelos multinivel en el marco de la IEE al considerar que su tratamiento diferencial de la variabilidad en los distintos niveles de análisis (estudiantes y centros) supondría unos resultados menos sesgados.

El estudio de Aitkin y Longford (1986), en el que examinan datos de contexto y resultado de 18 escuelas, concluye con el asentamiento de las bases para el uso de modelos multinivel en la investigación educativa: la necesidad de contar con datos de rendimiento y contexto de los estudiantes, así como de información de los centros y las administraciones educativas; el uso explícito de una estructura multinivel para incorporar los componentes de varianza a cada nivel de análisis (estudiante, centro y administración educativa); y el análisis de las interacciones entre las variables explicativas de los distintos niveles y la variación aleatoria de las escuelas en los coeficientes de regresión de las variables de nivel 1.

Un reflejo claro de estos avances metodológicos en la disciplina es el trabajo de Mortimore, Sammons, Stoll, Lewis & Ecob (1988, 1989), que examina los efectos de la escuela en el progreso académico de los estudiantes a través de técnicas de análisis multinivel. Estos autores recogieron datos sobre variables de contexto (edad, nivel socioeconómico, género, raza, rendimiento previo), proceso (políticas educativas y

organizativas a nivel de centro y aula, comportamiento de los docentes, opiniones de las familias, etc.) y producto (cognitivo – lectura y matemáticas – y no cognitivo – comportamiento, actitudes hacia la escuela, asistencia, etc.) en una muestra de 50 escuelas primarias a lo largo de cuatro cursos escolares. La principal variable criterio de este estudio fue el progreso académico observado en los estudiantes participantes a lo largo de tres años, y uno de los principales hallazgos fue el alto porcentaje de varianza de dicho progreso explicada por la escuela. Por ejemplo, de un 30% total de varianza entre las puntuaciones en lectura de los estudiantes, más de un 23% era debida a las características de la escuela y sólo un 6% era atribuible a las características contextuales de los estudiantes (Mortimore et al., 1989). En cuanto a las conclusiones sobre los factores escolares que contribuyen al progreso académico, se encontraron 12 cuestiones principales: liderazgo con objetivos claros, compromiso del director y los docentes, consistencia entre docentes, sesiones estructuradas y centradas en un tema concreto, enseñanza exigente a nivel intelectual, ambiente centrado en el trabajo, amplia comunicación entre estudiantes y docentes, registros escritos del progreso de los estudiantes, implicación de las familias y clima escolar positivo (Mortimore et al., 1989). Como puede observarse, la mayoría de estos resultados están en línea con los estudios anteriores.

Por otro lado, en esta fase comienza un proceso de internacionalización de la disciplina a través de la creación de una asociación de carácter internacional (*International Congress for School Effectiveness and Improvement*, ICSEI) que unió a los principales investigadores de la materia. La creación de este organismo dio paso a otros dos de los principales hitos en la internacionalización de la IEE: el establecimiento de un congreso internacional anual (ICSEI) en 1988, y la publicación del primer número de la revista científica *School Effectiveness and School Improvement* en 1990. Ambos continúan funcionando actualmente como punto de encuentro y difusión de los resultados de investigación de este campo.

3.1.3. Años 90: Los modelos integrados

La década de los 90 se caracterizó por un interés en las diferencias entre escuelas en función de los factores de proceso (Chapman et al., 2016), así como por una proliferación de revisiones sobre el estado de la cuestión en el campo de la IEE (Levine & Lezotte, 1990; Sammons, Hillman & Mortimore, 1995; Scheerens & Bosker, 1997; Scheerens & Creemers, 1996). Las conclusiones de estas revisiones acerca de los factores más

comúnmente relacionados con la eficacia no difieren en gran medida de las presentadas en la década de los 80 (Chapman et al., 2016), aunque sí son recogidas y catalogadas de manera más sistemática. La conclusión novedosa que sí subrayan algunos de estos trabajos es la gran importancia del nivel de aula en el rendimiento y la eficacia, por encima del efecto de la escuela, derivada de trabajos tan relevantes dentro del área como el de Creemers (1994) o Teddlie & Stringfield (1993).

Así como en las fases anteriores se advierte una predominancia de trabajos procedentes de la tradición anglosajona, cabe destacar en esta década el esfuerzo dedicado por los investigadores de los Países Bajos al estudio de los factores asociados con la eficacia educativa en su país. En su revisión del estado de la cuestión sobre la IEE en este país europeo, Scheerens y Creemers (1996) detectaron una serie de factores que parecían repetirse de manera consistente en los 29 estudios sobre educación primaria y los 13 estudios sobre educación secundaria. La tabla 3.3. resume estos factores, indicando en cuántos estudios aparece cada uno y el tipo de correlación presentada con las variables de producto estudiadas. Estas variables se controlaron a través de la inclusión de factores de contexto en todos los casos.

Tabla 3.3. Factores escolares relacionados con variables de producto en los estudios neerlandeses.

Factor	Educación primaria		Educación secundaria	
	Asociación positiva	Asociación negativa	Asociación positiva	Asociación negativa
Enseñanza/feedback estructurados	5		1	
Experiencia docente	3	1		1
Liderazgo instruccional		2	1	
Clima de disciplina	2		3	1
Evaluación de los estudiantes	5			
Diferenciación		2		
Enseñanza a toda la clase	3			
Orientación al rendimiento	4		4	
Estabilidad/cooperación del equipo		3		3
Tiempo de aprendizaje/deberes	4		4	
Otras variables	16		8	
Varianza media entre escuelas		9		13,5

Fuente: Scheerens & Creemers (1996, p.187)

Tal como apuntan los autores, los resultados de esta revisión de la literatura parecen no apoyar claramente los hallazgos de la disciplina de la eficacia escolar concluidos por otros autores, al menos para el contexto neerlandés (Scheerens & Creemers, 1996).

Otra de las aportaciones principales de esta fase fue el surgimiento de modelos integrados en los que se relaciona el contexto, los factores de entrada (inputs), los procesos y los resultados (outputs) del sistema educativo, organizados en una estructura multinivel (Scheerens y Bosker, 1997).

Un primer ejemplo de este tipo de modelos se encuentra en el estudio de Scheerens (1990), el cual se centra en el análisis de los factores de proceso de los centros educativos y los docentes que se encontraban positivamente relacionados con el rendimiento de los estudiantes.

El modelo de eficacia escolar propuesto por este autor (figura 3.2) combina distintos tipos de factores, cuyas interrelaciones afectan al rendimiento de los estudiantes. En este estudio, el rendimiento no se concibe como una medida bruta, sino que se ajusta en función del rendimiento previo, la inteligencia y el nivel socioeconómico de los estudiantes (Scheerens, 1990). Además, como puede observarse en la figura 3.2, se tiene en cuenta la influencia de factores de todos los niveles (entorno, centro educativo, aula y alumno), así como de distintos tipos de factores, incluyendo variables de entrada y de proceso.

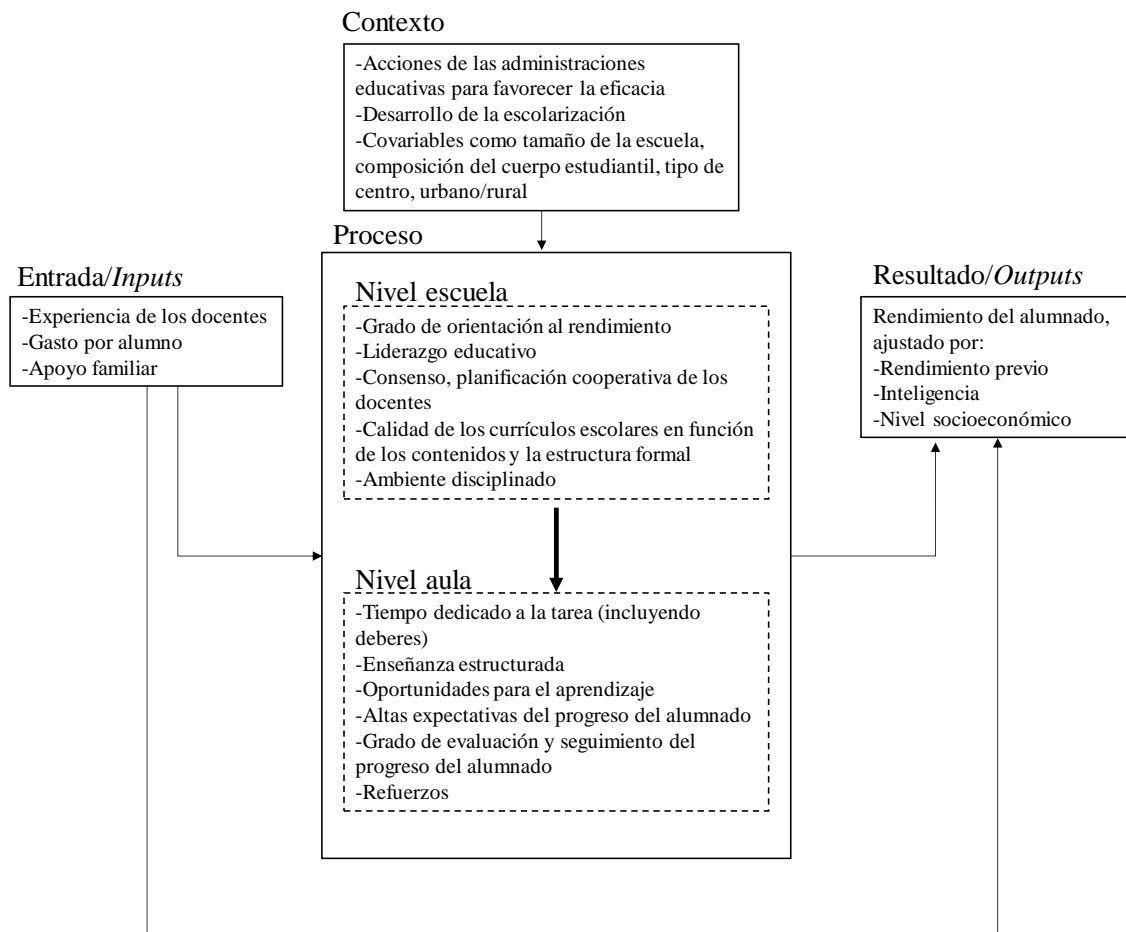


Figura 3.2. Modelo de eficacia escolar de Scheerens (1990)

Otro ejemplo de modelo integrado fue el presentado por Stringfield y Slavin (1992), denominado QAIT/MACRO (figura 3.3). Este modelo está organizado en cuatro niveles (estudiantes, profesionales que interactúan directamente con los estudiantes, escuelas y niveles superiores a la escuela, como la comunidad o las autoridades educativas) y centra su atención en la mejora del rendimiento de los estudiantes, permitiendo además el estudio de otros resultados deseables derivados de la escolarización.

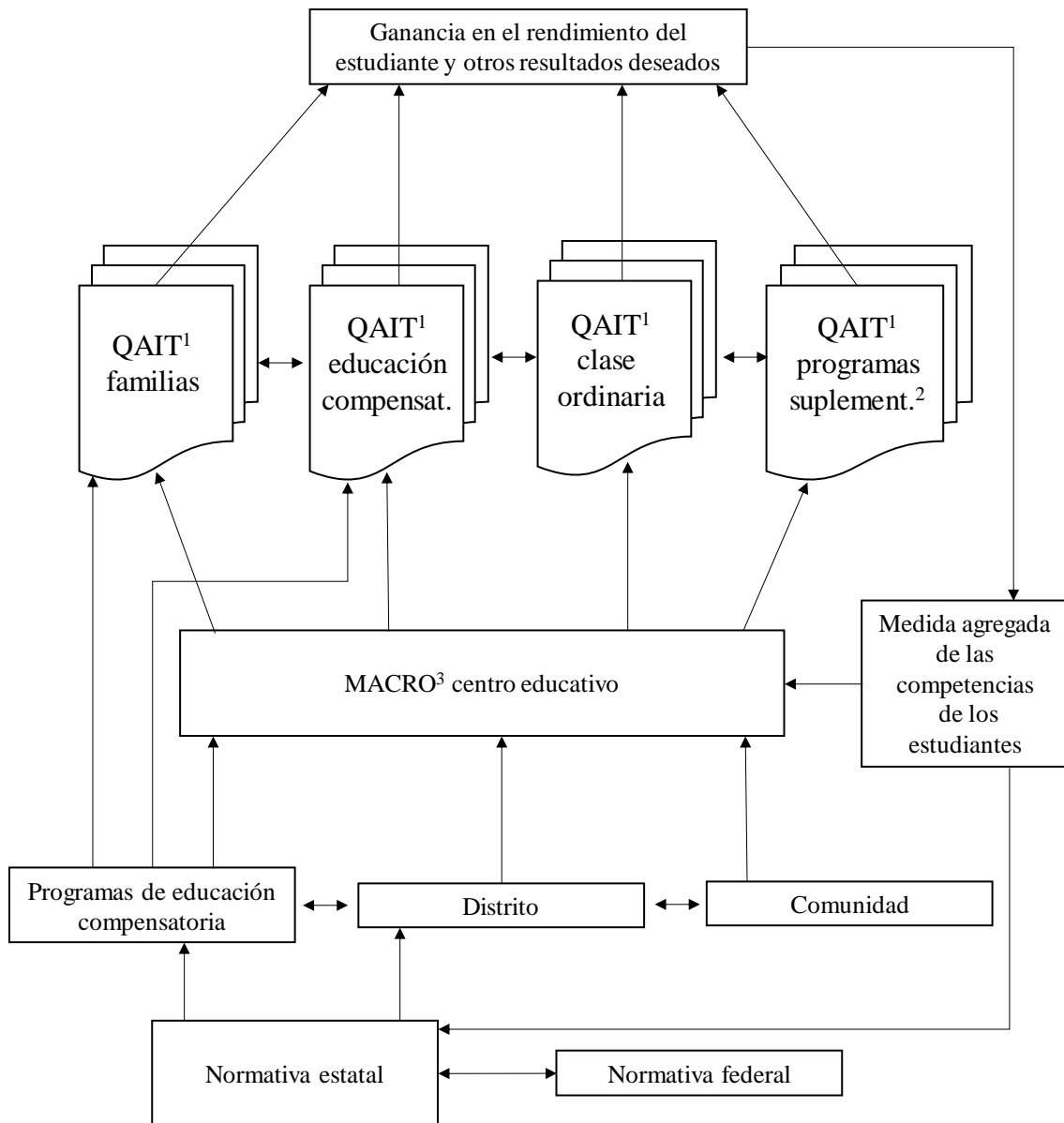


Figura 3.3. Modelo QAIT/MACRO, adaptado de Stringfield y Slavin (1992).

Notas: ¹QAIT: calidad, pertinencia, incentivos, tiempo de instrucción.²educación especial, educación bilingüe, etc. ³MACRO: objetivos significativos, atención a las funciones académicas, coordinación, contratación y formación, organización.

El acrónimo QAIT, referido a los agentes y programas encargados de la instrucción de los estudiantes, se concreta en los siguientes elementos:

- *Quality* (calidad): oportunidades de aprendizaje que se ofrecen a los estudiantes, tiempo dedicado a las tareas y principios docentes fundamentales (provisión de *feedback* de calidad, planteamiento de preguntas a los estudiantes, orientación del aprendizaje, etc.).

- *Appropriateness* (pertinencia): grado de adaptación del nivel de la instrucción a los conocimientos previos de los estudiantes.
- *Incentive* (incentivos): estimulación adecuada de la motivación por el aprendizaje.
- *Time on task* (tiempo de instrucción): aprovechamiento del tiempo lectivo para tareas de aprendizaje.

Por otro lado, las siglas MACRO, relacionadas con el funcionamiento del centro educativo, se refieren a:

- *Meaningful goals* (objetivos significativos): objetivos de centro claramente definidos y compartidos por todo el personal.
- *Attention to academic functioning* (atención a las funciones académicas): actividades de liderazgo educativo.
- *Coordination* (coordinación): relaciones organizativas entre los distintos miembros del centro educativo.
- *Recruitment and training* (contratación y formación): métodos de contratación del personal y orientación al desarrollo profesional de los docentes.
- *Organization* (organización): estructuración del día lectivo y empleo de personal de apoyo.

Una de las ideas novedosas que introduce es que la enseñanza efectiva no es responsabilidad únicamente de los docentes, sino que las familias y los programas educativos también cumplen funciones a este nivel. Por otro lado, también introduce un elemento dinámico, al indicar que los efectos de la docencia se acumulan a lo largo del tiempo (Scheerens & Bosker, 1997). A pesar de estas características, este modelo no tuvo excesiva repercusión en otros trabajos de la disciplina, posiblemente por su especificidad con respecto al sistema educativo estadounidense (Murillo, 2008).

Por último, en esta época se formula también el modelo comprensivo de la eficacia educativa de Creemers (1994) (figura 3.4). En este modelo, también de naturaleza multinivel, los niveles superiores establecen las condiciones necesarias para el funcionamiento de los inferiores, siendo así los productos de la educación el resultado de los efectos combinados de todos los niveles (Creemers & Kyriakides, 2008).

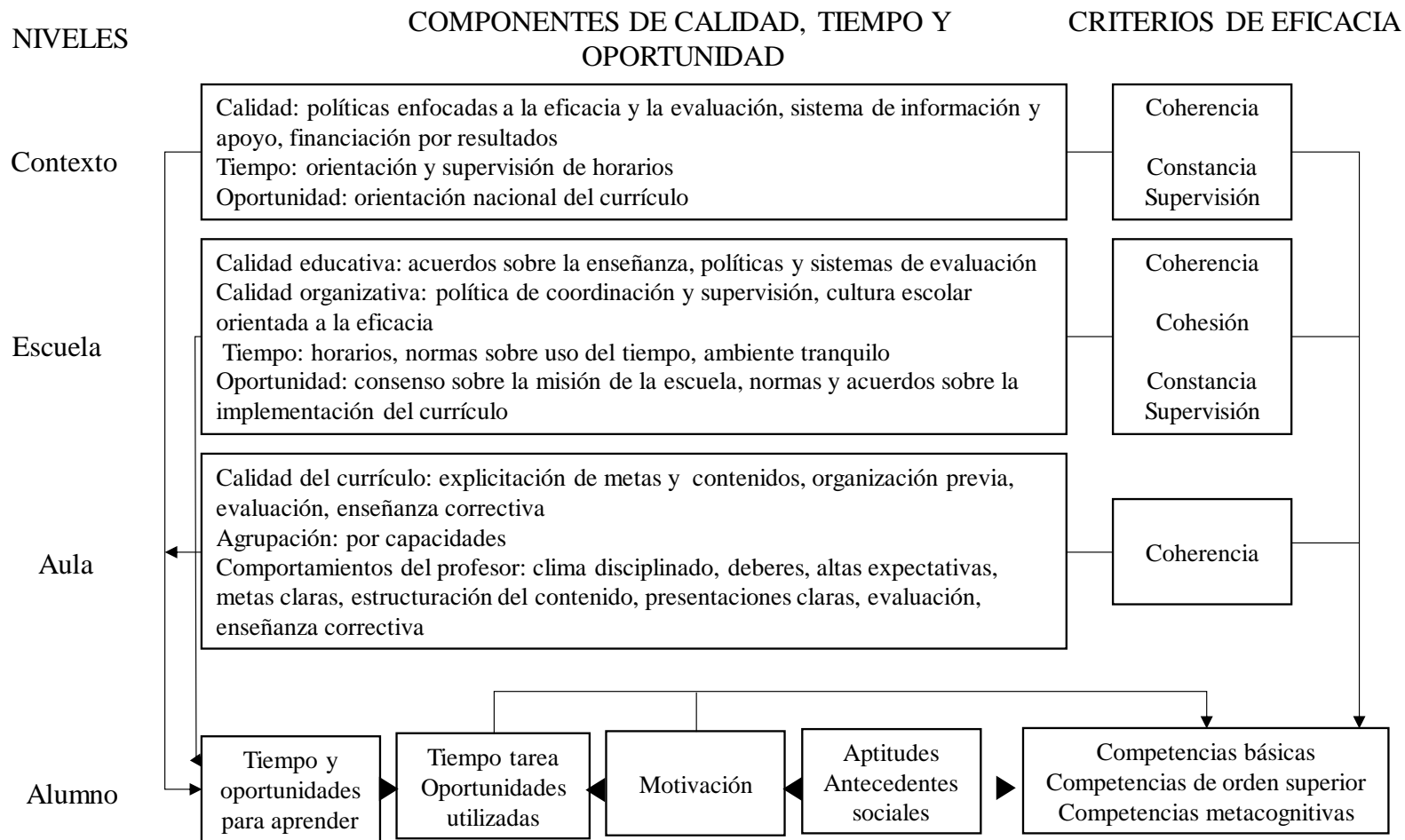


Figura 3.4. Modelo de Creemers (Muñoz-Repiso et al., 2000, p.18)

Una de las características principales de este modelo es que sitúa un mayor énfasis en el nivel del aula, confiriéndole mayor importancia en cuanto al efecto que tiene sobre los resultados de los estudiantes.

En esta fase se continuó trabajando en la apertura de la IEE en diversos sentidos. Por un lado, siguiendo el proceso que se comenzó a finales de los años 80 con la creación del ICSEI, se continuó con la internacionalización de la disciplina, lo cual produjo oportunidades para el contacto entre investigadores de diversos países, el establecimiento de estudios conjuntos y la mezcla de enfoques.

Otro de los hitos destacables de esta década fue la aparición de las evaluaciones de rendimiento a gran escala. La creación del TIMSS en 1995 y de PISA en 1999 abrió todo un nuevo abanico de posibilidades para los investigadores educativos, que comenzaron a disponer de grandes cantidades de datos relacionados tanto con el rendimiento o nivel competencial de los estudiantes como de su entorno demográfico y socioeconómico. Este hecho facilitó a su vez un avance metodológico con la inclusión de métodos mixtos que fusionaban el trabajo con datos procedentes de evaluaciones a gran escala con análisis profundos de caso (Chapman et al., 2016).

3.1.4. Siglo XXI: Nuevos modelos y métodos

Por último, la fase actual, iniciada hacia finales de la década del 2000 y todavía en desarrollo, se caracteriza por el surgimiento de nuevos modelos de eficacia y nuevas metodologías para su estudio.

Tras el surgimiento de la metodología multinivel en la década de los 80, el siglo XXI trajo consigo nuevos desarrollos en el ámbito del análisis de datos en la IEE. Una de las nuevas metodologías planteadas para el análisis de datos escolares fueron los modelos de ecuaciones estructurales (SEM). Aunque el SEM no fue desarrollado específicamente para el análisis de datos educativos, su naturaleza como extensión de la regresión lineal múltiple permite a los investigadores probar el ajuste de un modelo especificado por ellos, así como estudiar las relaciones entre las distintas variables que lo componen (Chapman et al., 2016). Además, esta metodología posibilita la introducción de variables de tipo latente (análisis factoriales confirmatorios), es decir, constructos que no pueden medirse directamente, pero que pueden estudiarse a través del análisis de otras variables observables (Castro y Lizasoain, 2012).

Otra técnica de uso reciente, cuyo surgimiento en educación está íntimamente relacionado con la disponibilidad de datos procedentes de evaluaciones a gran escala, es la minería de datos. Esta técnica permite la extracción de patrones o relaciones significativas a partir de bases de datos que contienen gran cantidad de información (Castro y Lizasoain, 2010; Martínez-Abad y Chaparro, 2017).

Por otro lado, los últimos años han visto un crecimiento en el número de revisiones sistemáticas de la literatura y de metaanálisis, que tienen por objetivo el resumen sistemático de la información producida por una disciplina determinada. Aunque las revisiones sistemáticas de la literatura resultan útiles para componer un estado de la cuestión de manera metódica, los metaanálisis se muestran especialmente ventajosos a la hora de resumir y analizar los hallazgos sobre la magnitud del efecto de un fenómeno en concreto (Chapman et al., 2016).

Además de las novedades en materia de métodos de análisis, en la época más reciente también han surgido nuevos modelos que aportan enfoques innovadores al estudio de la eficacia educativa. Un gran ejemplo de estas nuevas corrientes es el modelo dinámico de la eficacia educativa formulado por Creemers y Kyriakides (2008) (figura 3.5). Este modelo, que pertenece a la categoría de modelos multinivel integrados, como otros que se han mencionado anteriormente, tiene en cuenta resultados del aprendizaje de los estudiantes que no se circunscriben al ámbito cognitivo, sino que incluyen aspectos afectivos, psicomotores y nuevos aprendizajes.

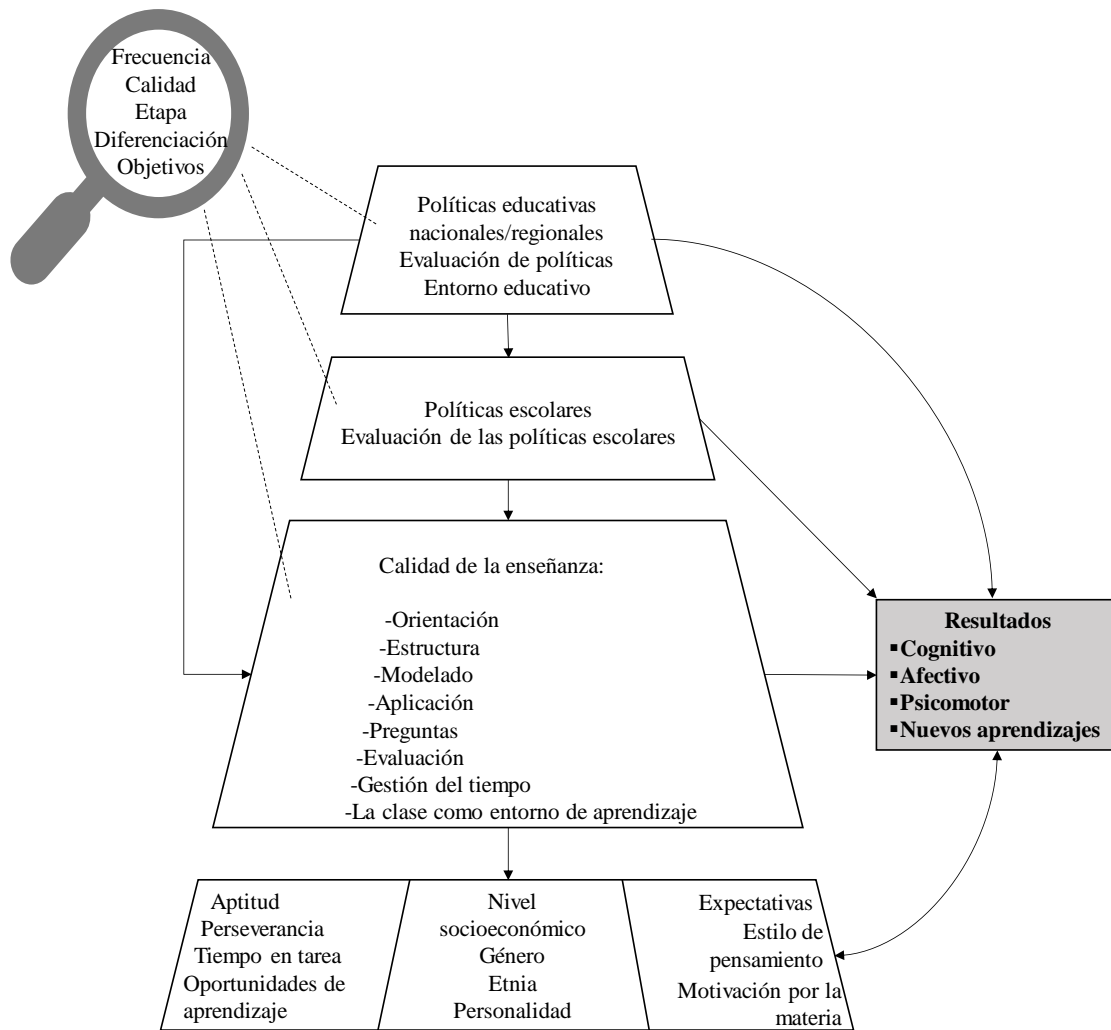


Figura 3.5. Modelo dinámico de eficacia educativa. Adaptado de Creemers & Kyriakides (2008).

En el nivel del modelo relativo al estudiante se incluyen fundamentalmente dos tipos de variables. Por un lado, se encuentran aquellos factores de contexto que han emergido como relevantes en el enfoque sociológico de la IEE, y que dentro de este nivel son los factores más estables a lo largo del tiempo (nivel socioeconómico, género, etnia, etc.). Por otro lado, los factores más relacionados con la perspectiva psicológica de la IEE se relacionan con la motivación, las expectativas el estilo de pensamiento del estudiante, o con cuestiones relativas a las tareas de aprendizaje (tiempo dedicado a la tarea, oportunidades de aprendizaje), elementos que tienden a cambiar con el paso del estudiante por el sistema educativo (Creemers & Kyriakides, 2008).

Por su parte, el nivel de aula presenta ocho factores que determinan el comportamiento de los docentes dentro del aula y que están relacionados con la enseñanza eficaz:

- **Orientación:** información proporcionada por el docente que ayuda a los estudiantes a comprender los objetivos o la razón de ser de una tarea o lección, promoviendo su compromiso con la misma.
- **Estructura:** organización de los contenidos presentados a los estudiantes, de manera que se produzca una introducción de los objetivos/contenidos, una enfatización de las ideas principales, y un repaso al final para afianzar los aprendizajes, con el objetivo de facilitar la comprensión y memorización.
- **Preguntas:** número, dificultad y tipo (producto o proceso) de las preguntas que se plantean al alumnado para implicarles en la marcha de la clase.
- **Modelado:** tareas que ayudan al alumnado a emplear estrategias de resolución de problemas o a generar estrategias propias, lo cual promueve el desarrollo de habilidades metacognitivas de regulación del propio aprendizaje.
- **Aplicación:** oportunidades para poner en práctica los aprendizajes realizados en el aula, con el objetivo de promover su significatividad.
- **El aula como entorno de aprendizaje:** este factor se compone de cinco elementos relacionados con el ambiente del aula: interacción docente-estudiantes, interacción entre estudiantes, tratamiento de estudiantes por parte del docente, competición entre estudiantes, y desorden en el aula.
- **Gestión del tiempo:** maximización del tiempo dedicado verdaderamente al aprendizaje a través de una gestión eficiente del aula.
- **Evaluación:** desarrollo del seguimiento de los estudiantes, especialmente a través de tareas de evaluación formativa, con un doble objetivo: identificar las necesidades del alumnado y valorar su propia práctica docente.

En el nivel escolar del modelo, el ámbito de políticas educativas se centra en dos elementos, la docencia y la creación de un entorno de aprendizaje en el centro educativo. Además, el modelo también identifica como factores determinantes de la eficacia aquellas acciones llevadas a cabo con el objetivo de mejorar dichos elementos, y la evaluación que se realiza de las políticas escolares. Dentro de las políticas relacionadas con la docencia, se tienen en cuenta aspectos como la cantidad, la calidad, o la provisión de oportunidades de aprendizaje.

Por último, el nivel de sistema educativo es, según los autores, el menos desarrollado debido a la falta de estudios que demuestren la relación entre los factores del sistema y el desempeño de los estudiantes (Creemers & Kyriakides, 2008). En este nivel se encuentran las políticas nacionales y regionales sobre educación que tengan consecuencias para la mejora de la docencia o el entorno de aprendizaje de las escuelas, así como la evaluación de dichas políticas. Además, incluye dos aspectos referidos al contexto más amplio de la educación, que son el apoyo proporcionado por diversos agentes (empresas, universidades, investigadores, etc.) a las escuelas, ya sea a través de financiación o información valiosa, y las expectativas de los distintos agentes sociales sobre los resultados que las escuelas deberían producir.

Como puede observarse, hay algunos elementos habitualmente presentes en los modelos integrados que no se encuentran en el modelo dinámico, como los conocimientos del profesorado, el liderazgo del equipo directivo o la estructura del sistema educativo. A este respecto, Creemers y Kyriakides (2008) destacan que su modelo pretende reflejar solamente aquellos factores que pueden tener un impacto directo sobre el aprendizaje de los estudiantes, eliminando de esta manera otras características cuya influencia en el rendimiento es indirecta o no está probada a través de los estudios de la IEE.

3.2. LA IEE EN ESPAÑA

Como se ha visto a lo largo de este capítulo, muchas de las investigaciones reseñables relacionadas con la eficacia escolar se han realizado en los países anglosajones. Sin embargo, en España también existe una tradición investigadora en eficacia escolar, con grandes exponentes dentro de la disciplina. Esta sección presenta las investigaciones más importantes realizadas en nuestro país.

3.2.1. Siglo XX

Los primeros trabajos realizados dentro de la línea de la IEE son resumidos por Murillo en su revisión del estado de la cuestión de la eficacia escolar en España (Murillo, 2003). Este autor destaca diez trabajos principales elaborados entre 1979 y 2000 que suponen los cimientos de la IEE en España. Como trabajos fundacionales de la disciplina en el ámbito nacional se encuentran las tesis doctorales de Millán en 1978 y Fuentes en 1986 (ambos citados por Murillo, 2003). El primero se trata de una descripción narrativa de los modelos y enfoques de la IEE, mientras que el segundo constituye el primer estudio empírico nacional sobre eficacia escolar.

A principios de los años 90 hubo un aumento de trabajos empíricos sobre este tema, fundamentalmente a través de tesis doctorales (Barrueco, 1992; García Durán, 1991; Rodríguez Díez, 1990; Rodríguez Gómez, 1991, todos citados en Murillo, 2003). Todos estos estudios partían de objetivos similares, basados en el control del impacto de las variables de contexto o entrada en el rendimiento de los estudiantes para poder estudiar los factores de proceso escolar relacionados con la eficacia. Asimismo, también extrajeron algunas conclusiones comunes, como la importancia de la participación de la comunidad educativa, el clima del centro o el funcionamiento de los órganos de gestión (ver tabla 3.4).

En 1995, el Centro de Investigación y Documentación Educativa (CIDE) emprende, en colaboración con la OCDE, un estudio sobre la eficacia con una doble vertiente: una revisión de la normativa nacional y un estudio empírico. El objetivo de la parte empírica de este trabajo fue “conocer la respuesta de los centros educativos a las reformas tendentes a mejorar la eficacia en la gestión de recursos y la visión que de esta realidad se tiene a nivel local” (Muñoz-Repiso et al., 1995, p.22). El estudio se llevó a cabo en 24 centros educativos (12 de cada etapa), donde se realizaron entrevistas a inspectores, equipos directivos, y docentes, familias y estudiantes implicados en el Consejo Escolar, obteniéndose más de 500 variables de estudio. Las conclusiones principales del trabajo dieron como resultado un modelo de eficacia para cada una de las etapas educativas estudiadas (primaria y secundaria). La figura 3.6 muestra el modelo de eficacia para escuelas primarias, que representa las relaciones entre las principales variables de contexto o entrada, de proceso y de producto. El modelo relativo a la educación secundaria no se reproduce aquí debido a que sus relaciones presentan menos claridad, coherencia y parsimonia que el modelo de primaria, encontrándose pocas relaciones triangulares entre los elementos (Muñoz-Repiso et al., 1995).

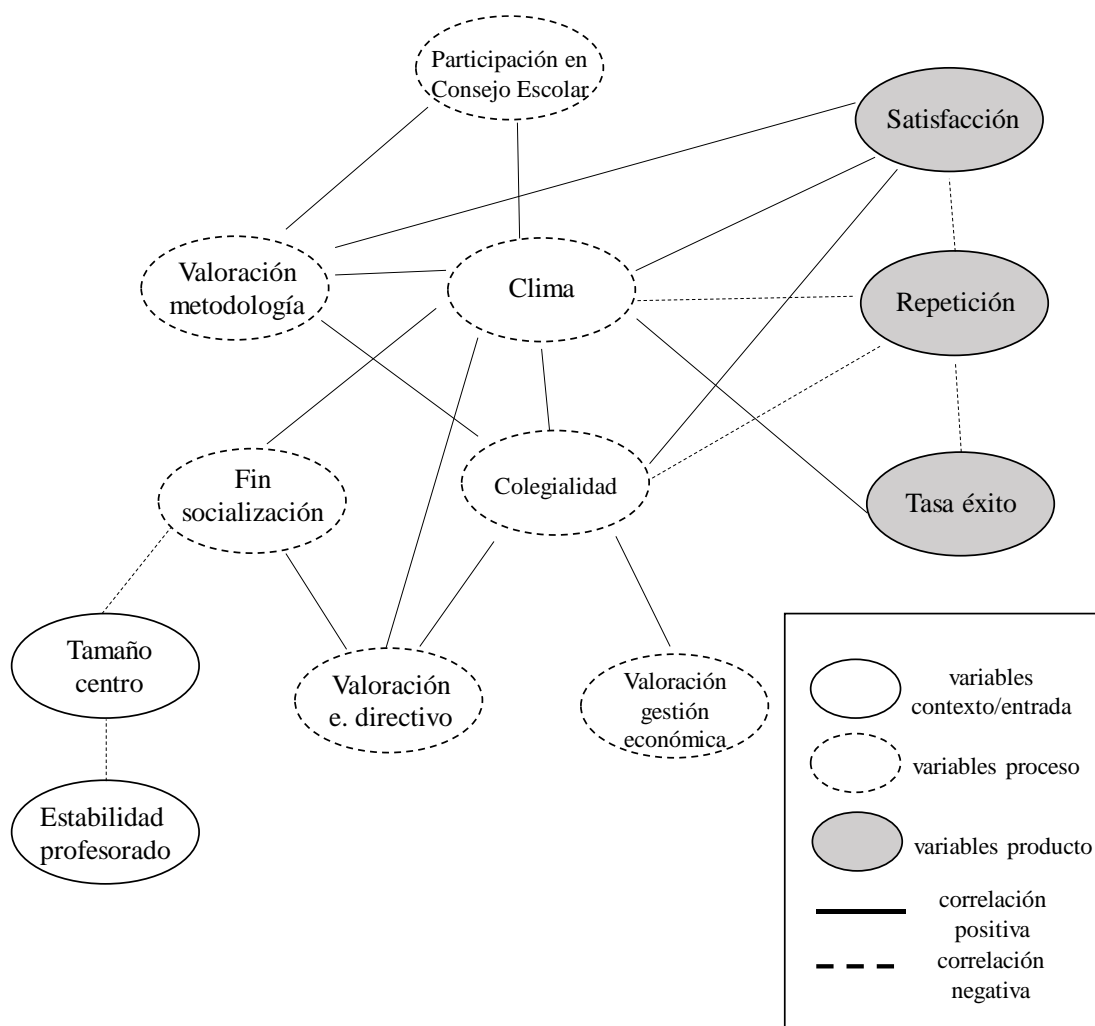


Figura 3.6. Modelo de eficacia de los centros de primaria, Muñoz-Repiso et al. (1995, p.118)

Este mismo organismo condujo en el año 2000 una nueva investigación sobre la eficacia en las escuelas de educación obligatoria, enmarcada igualmente en una colaboración europea, y esta vez centrada en el estudio de programas de mejora de la eficacia escolar llevados a cabo en centros educativos (Muñoz-Repiso et al., 2000). A través de una recogida de datos tanto cualitativos (revisión documental, entrevistas) como cuantitativos (cuestionarios), los investigadores de este proyecto llevaron a cabo un estudio en profundidad de cinco centros educativos en los cuales se desarrollaban experiencias exitosas de mejora de la eficacia, con el objetivo de analizarlos y detectar factores en común que se pudiesen extrapolar a otros centros. Sus principales conclusiones pueden consultarse en la tabla 3.4.

Por último, cabe destacar otros dos estudios empíricos que han contribuido a la tradición de la IEE en España. Uno de ellos es el de Castejón (1994, 1996), que estudió

los factores escolares asociados a la eficacia, calculada en función de la diferencia entre las puntuaciones esperadas y las obtenidas por los estudiantes. El otro es la tesis doctoral de González Galán (2000, citado por Murillo, 2003), que profundiza en el estudio del clima escolar en relación con la eficacia, concluyendo que el clima de trabajo existente en el centro es un fuerte predictor del nivel de eficacia.

Como puede apreciarse en la tabla 3.4, casi todos los trabajos presentan similitudes en cuanto a su objetivo (identificar factores de proceso relacionados con la eficacia de los centros educativos) y el tipo de variables que utilizan. Asimismo, también comparten algunas limitaciones que caracterizan las investigaciones de eficacia escolar llevadas a cabo en el siglo XX. La primera es la falta de validez en la variable criterio de rendimiento, la cual se basa en calificaciones escolares difícilmente comparables entre centros, en lugar de consistir en mediciones estandarizadas. La segunda es la utilización de metodologías de análisis de datos que no tienen en cuenta la estructura jerárquica de los datos, en los que los estudiantes se encuentran agrupados por escuelas, y por lo tanto hay que tener en cuenta la variabilidad de ambos niveles para no cometer errores de estimación de los parámetros. Como apunta Murillo (2003), una metodología de análisis más adecuada sería la aplicación de modelos multinivel. Estas limitaciones se abordan en los estudios realizados con posterioridad, que muestran un mayor rigor metodológico en la recogida y el análisis de los datos.

Tabla 3.4. Características de los estudios de IEE más relevantes del siglo XX en España

Estudio	Muestra	V. criterio	Variables predictoras	Metodología	Conclusiones asociadas al rendimiento o eficacia
Fuentes (1986)	91 centros EGB	Calificaciones	Factores de contexto del alumno, cuestiones de proceso sobre el funcionamiento del centro	Cuantitativa	Origen social del alumnado, actuación docente, funcionamiento y clima del centro
Rodríguez (1990)	-	Calificaciones y otras (repetidores, satisfacción, absentismo...)	Entrada (características de alumnos y docentes, recursos materiales y humanos, tipo de centro) Proceso (modelo organizativo, clima y procesos de aula)	Cuantitativa	Dotación del centro y funcionamiento del Equipo Directivo. Clima de centro influye en rendimiento
García Durán (1991)	14 centros rurales de EGB <i>outliers</i> en variable criterio	Calificaciones	Origen social alumnos (controlada) Consejo escolar, plan de centro, equipo directivo, organización profesorado, org. alumnado, org. Trabajo docente, espacios y recursos, servicios complementarios, clima, relaciones con familias y comunidad, evaluación de la labor escolar	Cuantitativa	Revisión del plan de centro, funcionamiento de equipo directivo y claustro, orientación, producción de recursos didácticos, evaluación contrastada, cooperación con familias y servicios externos, altas expectativas sobre el centro
Rodríguez Gómez (1991)	-	Calificaciones	Clase social alumnos (controlada) Características profesorado, clima organizacional y de centro, relaciones con comunidad, funcionamiento de órganos directivos	Cuantitativa	Docentes: satisfacción, altas expectativas, cooperación, innovación, adaptación de la instrucción. Centro: participación de todos los agentes, clima positivo, funcionamiento adecuado de claustro y liderazgo
Barrueco (1992)	-	Calificaciones	Formación del alumnado, satisfacción de familias, docentes y alumnos	Cuantitativa	Naturaleza de la comunidad educativa, autocontrol participativo de la comunidad educativa, gestión técnica, cualificación docente del profesorado, planificación didáctica

Estudio	Muestra	V. criterio	Variables predictoras	Metodología	Conclusiones asociadas al rendimiento o eficacia
Castejón (1994, 1996)	24 centros de Secundaria	Calificaciones	Alumnos: nivel socioeconómico, inteligencia, rendimiento previo, autoconcepto (controladas), percepción del centro. Docentes: formación y experiencia, motivación, satisfacción, percepción del centro.	Cuantitativa	Percepción de los alumnos de los procesos educativos Satisfacción docente, percepción de profesores sobre equipo directivo
CIDE (1995)	20 centros primaria y secundaria	Tasa de éxito, repetidores, satisfacción	542 variables de contexto, proceso y producto	Cualitativa	Clima de centro (primaria), clima escolar, trabajo en equipo, valoración equipo directivo, participación de los miembros del consejo escolar, acuerdo entre agentes sobre la socialización como un fin deseable (secundaria)
González Galán (2000)	-	Eficacia escolar	Variables relacionadas con el clima escolar	Cuantitativa	El clima de trabajo como factor fundamental que influye en la eficacia escolar
CIDE (2000)	Cinco centros	Experiencia exitosa de mejora de la eficacia escolar	Variables de proceso organizativas y pedagógicas de los centros educativos.	Mixta	Cultura escolar (compromiso, proyecto, valores, cohesión), dirección (gestión de conflictos, planificación, movilización de docentes), organización (flexibilidad), planificación del proceso de mejora (consensuada, flexible, detallada), procesos enseñanza-aprendizaje, estímulo externo para emprender el programa, experiencias previas de innovación

Fuente: elaboración propia a partir de Murillo (2003)

3.2.2. Siglo XXI

Las investigaciones más recientes realizadas en nuestro país demuestran un gran avance en cuanto a sus planteamientos metodológicos, adoptando enfoques más rigurosos y adecuados a los objetivos de estudio, y procedimientos más sistematizados que en la fase anterior.

En las dos últimas décadas surgen principalmente dos grandes trabajos dedicados al estudio de la eficacia escolar en nuestro país, ambos con un enfoque metodológico mixto que aporta por un lado la robustez de los resultados cuantitativos y por otro lado la gran profundidad proporcionada por los análisis cualitativos. Estos estudios son la Investigación Iberoamericana sobre Eficacia Escolar, coordinada por el profesor Javier Murillo y el proyecto sobre Eficacia Escolar en el País Vasco, liderado por el profesor Luis Lizasoain Hernández.

El primero de estos trabajos fue la investigación iberoamericana de Murillo (2007) y llevada a cabo en casi 100 escuelas de primaria de 9 países (Bolivia, Chile, Colombia, Cuba, Ecuador, España, Panamá, Perú y Venezuela). Para evaluar los factores relacionados con la eficacia escolar, se recogió información de una gran cantidad de factores de contexto (nivel socioeconómico, características personales, rendimiento previo, etc.), proceso (comportamiento de los estudiantes, relaciones entre los miembros de la comunidad educativa, metodologías docentes, organización social y pedagógica, clima institucional, estilo directivo, etc.) y de producto (autoconcepto y rendimiento en matemáticas y lengua). Los instrumentos utilizados para la recolección de información fueron tanto de tipo cuantitativo (pruebas estandarizadas para medir el rendimiento y cuestionarios), como de tipo cualitativo (entrevistas, grupos focales, observación en aula).

El proyecto se divide en dos partes en función de la metodología de análisis de datos empleada. En primer lugar, con el objetivo de determinar el nivel de eficacia de los centros, se llevaron a cabo una serie de modelos multinivel de cuatro niveles (estudiante, aula, centro y país). Estos modelos permitieron estudiar los factores de contexto asociados al rendimiento de los estudiantes, y estimar sus niveles de logro, que se define como el nivel de rendimiento una vez aislados los efectos de los factores socioeconómicos y personales (Murillo, 2007). Cabe destacar que, al incorporar en los modelos el nivel de rendimiento previo del estudiante, los resultados de este análisis reflejan el aprendizaje realizado por los alumnos en un curso escolar, pudiendo entonces referirnos al efecto escolar en dicho aprendizaje como su *valor añadido*. Los resultados de esta fase

relacionados con la magnitud de los efectos se resumen en la tabla 3.5, que evidencia dos cuestiones fundamentales: el gran aporte del nivel de aula a la explicación de la variabilidad en las puntuaciones de los estudiantes, y la notable diferencia existente entre el rendimiento de matemáticas y el de lengua.

Tabla 3.5. Porcentaje de varianza del logro académico explicado por el aula, la escuela y el país para el rendimiento de matemáticas y lengua

Nivel	Magnitud del efecto	
	Matemáticas	Lengua
Aula	22.02%	11.61%
Escuela	10.70%	8.45%
País	15.52%	1.44%

Fuente: Murillo (2007, p.124)

Una vez determinada la eficacia de los centros, se procedió a realizar estudios cualitativos de los centros que habían probado ser los más y menos eficaces dentro de sus países, para estudiar los factores de proceso que podrían estar relacionados con el valor añadido aportado por dichos centros. A través de los análisis de caso se detectaron 10 áreas generales en las que se percibían diferencias de funcionamiento entre los centros más y menos eficaces (tabla 3.6)

Tabla 3.6. Resumen de resultados de los estudios de caso en el estudio iberoamericano

Categoría	Alta eficacia	Baja eficacia
Misión de la escuela	Formación integral de los estudiantes Preocupación por el papel social de la escuela	Transmisión de conocimientos El proyecto educativo es sólo un requisito formal
Implicación y compromiso docente	Fuerte implicación de los docentes Trabajo en equipo	Falta de compromiso Falta de trabajo en equipo
Dirección escolar	Dirección implicada y comprometida con la mejora de la escuela Preocupada por el desarrollo de los docentes Liderazgo positivo Capacidad técnica Estilo centrado en cuestiones pedagógicas	Dirección no comprometida Estilo burocrático Falta de liderazgo Falta de preparación
Clima de escuela	Cordial y afectivo Buenas relaciones Se muestra alegría	Ambiente tenso y desconfianza División entre docentes
Clima de aula	Relaciones afectivas hacia los alumnos	Relaciones de poder y dominación Uso de castigo físico
Expectativas	Altas expectativas hacia los alumnos Altas expectativas hacia los docentes y la dirección	Bajas expectativas hacia los alumnos
Instalaciones	Adecuadas, bien mantenidas, limpias	Inadecuadas, mal mantenidas, desaseadas
Metodología docente	Sesiones planificadas, estructuradas, con objetivos claros	Expositiva Repetitiva

	Centrada en competencias básicas Actividades variadas y participativas Utilización de recursos didácticos	No utilización de recursos didácticos
Gestión del tiempo	Pocas ausencias de los docentes Clases iniciadas con puntualidad Ausencia de interrupciones en el aula Organización flexible del tiempo	Pérdida de muchos días lectivos Inicio de las clases con retraso Mala gestión del tiempo en el aula
Organización del aula	Distribución heterogénea de los alumnos	Situación frontal de las mesas Agrupamiento homogéneo

Fuente: Murillo (2007, p.253)

Tras haber analizado toda la información recogida sobre características de entrada, proceso y producto a todos los niveles, los investigadores participantes en este proyecto elaboraron un modelo que resumen tanto los factores implicados en la mejora del logro de los estudiantes, como sus relaciones entre ellos (figura 3.7).

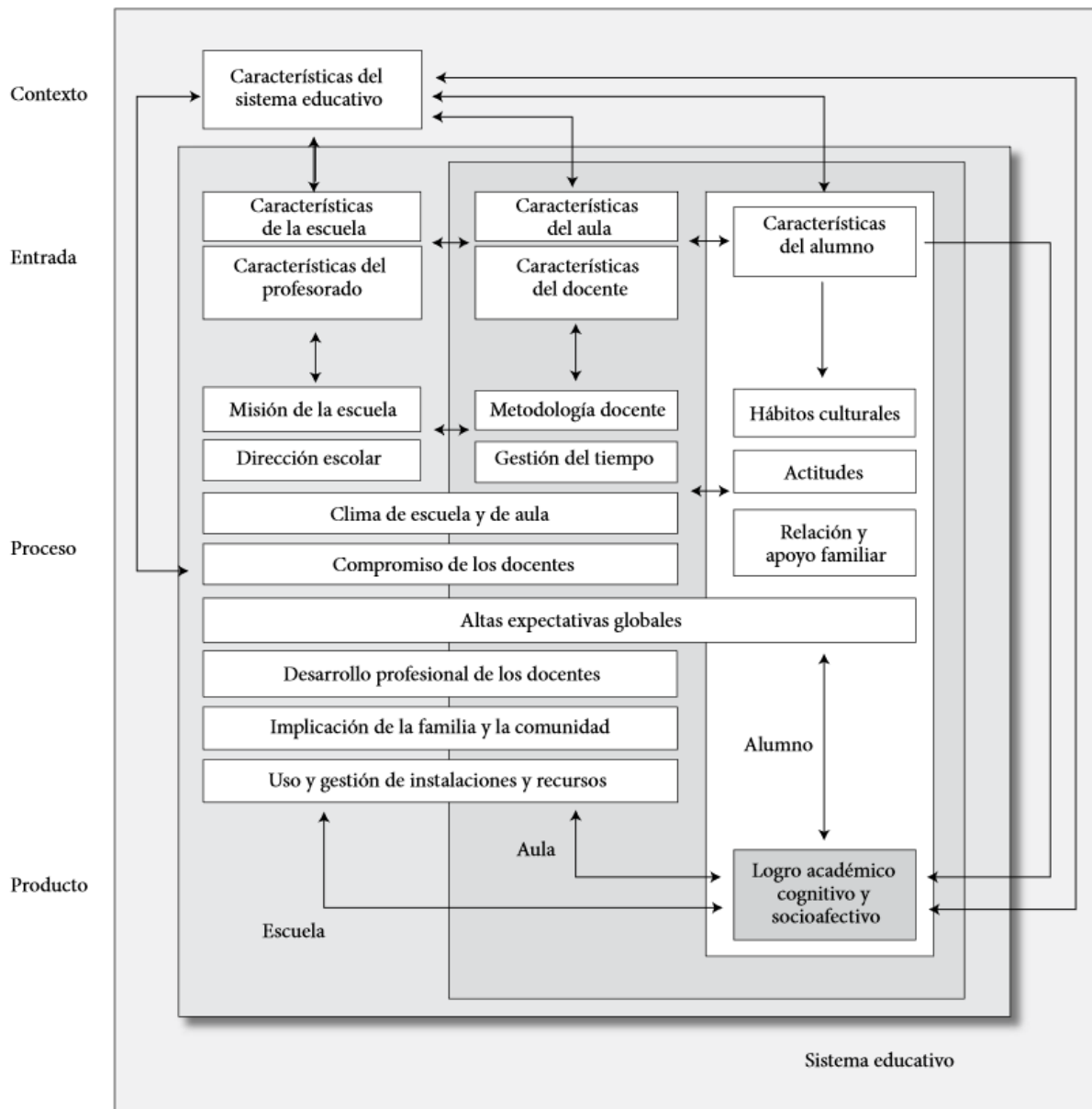


Figura 3.7. Modelo Iberoamericano de Eficacia Escolar, Murillo (2007, p.278)

El segundo de los principales estudios relacionados con la eficacia escolar llevados a cabo en nuestro país es el proyecto sobre la **eficacia escolar en el País Vasco** realizado por un equipo de investigadores de la Universidad del País Vasco (UPV/EHU) en colaboración con el Instituto de Evaluación e Investigación Educativa (ISEI/IVEI), con el profesor Luis Lizasoain Hernández como investigador principal. El objetivo general de este proyecto es “Conocer la verdadera aportación de la escuela al conjunto de aprendizajes y logro de competencias que se miden en las pruebas de la Evaluación Diagnóstica para después elaborar un catálogo de buenas prácticas” (Lizasoain et al., 2013).

El proyecto se compuso de dos fases claramente diferenciadas. En primer lugar, se llevó a cabo una fase de caracterización de los centros de alto valor añadido. Para ello, se construyeron modelos multinivel a partir de los datos de las Evaluaciones Diagnósticas de 2009 y 2010 referidos tanto al nivel competencial de los estudiantes como a las características de su entorno. Los resultados de este análisis sirvieron para seleccionar aquellos centros que mostrasen una mayor diferencia entre sus puntuaciones reales (brutas) y las puntuaciones estimadas en función de las variables contextuales que resultaron ser significativas en el análisis multinivel (residuo o puntuación ajustada), seleccionándose finalmente 32 centros educativos de primaria y secundaria (Joaristi, Lizasoain & Azpillaga, 2014).

Una vez seleccionados los centros, comenzó el proceso de recogida de información cualitativa (entrevistas) sobre los centros seleccionados para poder analizar sus procesos organizativos y pedagógicos. Los informantes de dichas entrevistas fueron los equipos directivos de los centros educativos, los inspectores encargados de los centros, y los asesores de los servicios zonales de atención a los centros educativos (*Berritzeguneak*).

El análisis de las 90 entrevistas que se realizaron entre estos informantes en los 32 centros del estudio dio como resultado una serie de buenas prácticas asociadas a la alta eficacia, divididas en 14 categorías (tabla 3.7).

Tabla 3.7. Resumen de las buenas prácticas detectadas en el estudio de las escuelas de alta eficacia, primera fase del estudio UPV/ISEI.

Categoría	Prácticas
Proyectos, planes y formación	Participación activa en actividades y cursos de formación Compromiso con la finalización de los proyectos Demanda y continuidad de formación y constancia con nuevos proyectos
Metodologías y materiales de enseñanza	Predominancia de actividades docentes innovadoras, aunque un 33% de los centros presentan metodología tradicional. Inconcluyente.
Atención a la diversidad	Especial atención a alumnado con bajo rendimiento o repetidor, a los criterios de agrupación de alumnos, al alumnado con NEE y al alumnado de origen inmigrante Implicación de todo el profesorado Recursos y programas específicos según las necesidades del centro Protocolos de actuación y acogida para alumnado migrante
Seguimiento del alumnado, atención individualizada, orientación y tutoría	Seguimiento exhaustivo del alumnado Recogida altamente sistematizada de información (detección precoz e intervención temprana) Relación estrecha y sistemática con las familias Atención individualizada a través de adaptación de la instrucción Plan de acción tutorial elaborado en estrecha colaboración docentes-departamento de orientación Peso importante del profesorado consultor y orientador dentro del centro (seguimiento, convivencia, detección de problemas, refuerzo y NEE)
Evaluación del alumnado	Mínimos de evaluación claramente definidos, acordados por el claustro y públicos Realización de evaluaciones externas y/o iniciales
Gestión del tiempo	Puntualidad, respeto al horario y aprovechamiento del tiempo lectivo
Liderazgo y equipo directivo (ED)	Estabilidad del ED y capacidad de gestión Equipo docente estable que participa en el proyecto de centro y tiene sentimiento de pertenencia Clima y expectativas positivas Reconocimiento del trabajo docente por parte del ED Compromiso del ED con el desarrollo integral del alumnado Liderazgo del ED en la innovación y compromiso con la formación
Modelos de gestión y organización	Alto grado de reparto de responsabilidades con clara asignación de roles Organización sistematizada (sistemas de gestión de calidad, por ejemplo) Cuidado, protección y acompañamiento al profesorado
Coordinación	Especial atención a la coordinación con otros centros, redes e instituciones socio-sanitarias Planificación previa de calendarios de reuniones de coordinación Herramientas telemáticas para gestión de la coordinación
Implicación y pertenencia al centro	Alto nivel de implicación, disponibilidad, compromiso y entrega al trabajo Fuerte sentimiento de pertenencia al centro por parte del profesorado
Evaluación de docentes, centro, programas y actividades, uso de la Evaluación Diagnóstica	Cultura evaluadora orientada a la mejora de procesos y resultados Valoración positiva de las Evaluaciones Diagnósticas como oportunidad de mejora Resultados de las evaluaciones utilizados para incrementar la atención individualizada a los estudiantes, mejorar procedimientos de evaluación del aprendizaje, u orientación del trabajo por competencias
Clima y convivencia	Clima positivo y actitud proactiva hacia el mismo (formación docente, iniciativas, etc.) Existencia de plan de convivencia Protocolos eficaces de resolución de conflictos basados en medidas educativas
Familia, comunidad y capital social	Presencia consistente de familias en el centro, con un alto grado de implicación en actividades pedagógicas y formación

Fuente: Lizasoain et al. (2013).

En segundo lugar, y con el objetivo de mejorar la validez interna del trabajo, se emprendió una segunda fase en la que se realizaron profundos estudios de caso de centros concretos, esta vez añadiendo centros de baja eficacia para poder comparar los resultados. En esta segunda fase se contó con 9 centros de alta eficacia (4 de educación primaria, 4 de educación secundaria, y 1 de ambas) y 8 de baja eficacia (4 de educación primaria y 4 de educación secundaria). Antes de comenzar los estudios de caso, se entrevistó a informantes clave (equipos directivos en los centros de alta eficacia e inspectores educativos en los centros de baja eficacia) para informar la organización del estudio. El procedimiento de estudios de caso consistió en la recogida de grandes cantidades de información cualitativa a través de diferentes instrumentos:

- Grupos de discusión con docentes y estudiantes
- Revisión de los documentos del centro
- Observaciones de aulas, reuniones de coordinación y otros espacios (recreos, comedor, etc.)
- Diario de campo de los investigadores

A través del análisis de toda esta información, el producto de esta segunda fase fue una comparativa de las prácticas educativas de los centros en función de 9 categorías, cuyo resumen se encuentra en la tabla 3.8.

Tabla 3.8. Resumen de las prácticas educativas detectadas en los centros de alta y baja eficacia, segunda fase del estudio UPV/ISEI

Categoría	Alta eficacia	Baja eficacia
Proyectos y formación	Participación elevada en actividades formativas impulsadas y facilitadas por la dirección	Actividades formativas no relacionadas con necesidades del centro o planes concretos, no repercuten en una mejora
Metodologías	No se observan metodologías concretas asociadas a la eficacia	En algunos centros se aplican metodologías para las que los docentes no están formados
Atención a la diversidad y seguimiento del alumnado	Seguimiento cercano del trabajo del alumnado mediante un plan de acción tutorial y contacto constante con las familias. Adecuada atención a la diversidad, detección precoz e intervención temprana, trabajo de dos docentes en el aula	Tratamiento aislado de alumnos con necesidades específicas. Agrupamiento de estudiantes por rendimiento
Evaluación	Evaluación del alumnado frecuente y formativa. Contenidos y criterios mínimos de evaluación consensuados y públicos. Evaluación periódica de programas, proyectos y profesorado	Criterios de evaluación en constante debate. No se menciona la evaluación del profesorado
Gestión del tiempo	Adecuado aprovechamiento del tiempo lectivo	Algunos muestran dificultades para el aprovechamiento del tiempo

Categoría	Alta eficacia	Baja eficacia
Liderazgo	Liderazgo firme, visión clara con metas compartidas	Dirección inestable, o muy estable sin reparto de responsabilidades
Gestión y organización del centro	Alto nivel de sistematización y planificación de la gestión Coordinación interna y externa bien organizadas Dedicación e implicación docente Atención y cuidado al profesorado (acogida y desarrollo de funciones)	Bajo nivel de participación de los docentes en la toma de decisiones Procesos y consecución de logros dificultados por falta de liderazgo firme Coordinación interna no sistematizada, limitada a información
Clima	Buen clima y convivencia, correcta resolución de conflictos	Buen clima en general, algunos centros con problemas de conflictividad.
Familias y comunidad	Alto nivel de implicación de las familias en las actividades del centro	Bajas expectativas sobre el rendimiento del alumnado Situaciones conflictivas con las familias

Fuente: Elaboración propia a partir de Lizasoain et al. (2015)

Estos dos estudios (Iberoamérica y País Vasco) sientan las bases teóricas y metodológicas del trabajo de investigación presentado en esta tesis. Su influencia se aprecia claramente en el proceso de investigación seguido, en el que se utilizan los factores de entrada/proceso para identificar aquellas escuelas cuyo rendimiento se muestre más alejado del rendimiento esperado en función de su composición, para luego estudiar qué factores de carácter procesual pueden estar relacionados con los niveles de eficacia de las escuelas. Por lo tanto, ambos trabajos se retomarán a la hora de presentar las conclusiones del presente estudio, con el objetivo de discutir las similitudes y diferencias encontradas.

3.3. LECCIONES APRENDIDAS

A lo largo de la historia de la investigación sobre eficacia escolar han surgido numerosas teorías, modelos y estudios empíricos con el objetivo de esclarecer cuáles son los factores a nivel de contexto, escuela, aula y estudiante que en último término benefician el desarrollo del alumno tanto en rendimiento académico como en la adquisición de otra clase de capacidades no necesariamente cognitivas.

En este capítulo se ha dedicado un espacio considerable a exponer aquellos factores que emergen de manera consistente como relevantes a lo largo de la tradición de la IEE (clima escolar, liderazgo educativo, expectativas sobre el rendimiento de los estudiantes, gestión del tiempo, prácticas de aula). Sin embargo, existen otras consideraciones relevantes que pueden concluirse de los diversos estudios analizados, quizá más relacionadas con el planteamiento de las investigaciones que con sus hallazgos. A modo de conclusión, en esta sección se exponen cinco de estas “lecciones aprendidas”.

La magnitud de los efectos del nivel de aula

La disciplina de la investigación en eficacia escolar se centró en un primer momento en evaluar los efectos de la escuela sobre el rendimiento de los estudiantes. Sin embargo, a medida que los estudios se fueron sofisticando, y se emplearon nuevos métodos de análisis de los datos, los autores comenzaron a darse cuenta de la gran influencia que tenía el nivel de aula sobre el rendimiento, mayor incluso que el de la escuela (ver tabla 3.5). Por este motivo, se hace necesaria la inclusión de variables fundamentales del nivel de aula, como las actuaciones pedagógicas del docente (Creemers & Kyriakides, 2008).

La inclusión de los factores de contexto

El hecho de incluir factores de entrada/contexto como covariables en los análisis de rendimiento de los estudiantes es una práctica relativamente habitual en la IEE, como ha podido observarse a lo largo de este capítulo. Sin embargo, esto no lo hace menos digno de mención, ya que la perspectiva acerca del rendimiento que proporciona el control de las variables de entrada/contexto es fundamental para comprender el aporte que las aulas, los centros o los países hacen al aprendizaje de los alumnos. Por tanto, la inclusión de estos factores en cualquier estudio cuyo objetivo sea el análisis de la eficacia de los centros o los sistemas educativos es esencial.

El uso de una metodología de análisis multinivel

Los datos procedentes de estudios escolares siempre presentan una estructura jerárquica multinivel, con los alumnos agrupados en aulas, estas en centros, estos en regiones/países, etc. Las técnicas clásicas de análisis, como la regresión, no son adecuadas para el estudio de estos datos, ya que obvian la variabilidad existente a distintos niveles y no permiten su inclusión en la investigación, produciendo de esta manera sesgos de estimación de los parámetros. Desde la década de los 80, muchos autores (Aitkin & Longford, 1986; Goldstein, 1987; Raudenbush & Bryk, 1986) abogan por el uso de esta técnica como estándar en la disciplina, y la predominancia de su uso se aprecia también en la actualidad (ver capítulo SLR, sección metodología), junto con otros nuevos métodos que también permiten la inclusión de la estructura anidada de los datos.

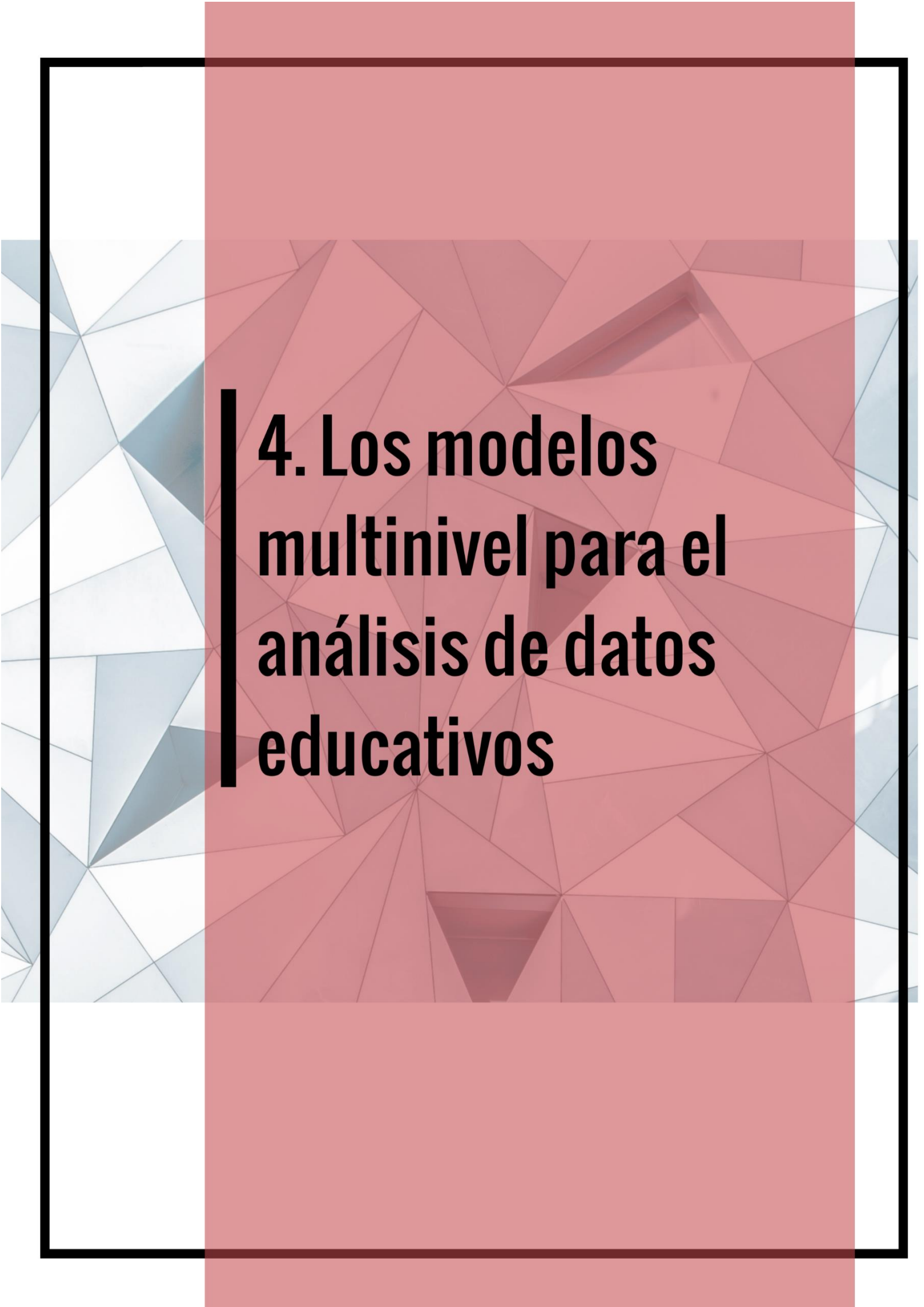
La garantía de la validez a través de medidas estandarizadas de rendimiento

El uso de medidas estandarizadas como base para el estudio del rendimiento de los estudiantes es una cuestión fundamental para la validez y fiabilidad de la investigación,

y para permitir las comparaciones entre los diversos centros estudiados (Murillo 2003). Por lo tanto, ya sea a través de pruebas de evaluación a gran escala ya existentes, como PISA (Gamazo et al., 2018) o las pruebas diagnósticas nacionales (Joaristi, Lizasoain & Azpillaga, 2014), o escalas validadas utilizadas específicamente para la investigación, resulta esencial registrar el nivel de rendimiento de una manera sistematizada.

La importancia de los resultados no cognitivos

La inclusión de factores de producto no cognitivos (motivación, desarrollo emocional, autoconcepto, etc.) es menos común que la utilización del rendimiento académico como medida del resultado del sistema educativo. Sin embargo, como apuntan algunos autores (Creemers & Kyriakides, 2008; Murillo, 2007), el desarrollo cognitivo académico no es (o no debería ser) el único objetivo de los sistemas educativos. Según la definición de la eficacia escolar presentada al principio de este capítulo, las escuelas eficaces (y, por ende, los sistemas educativos eficaces) deben comprometerse con el desarrollo integral de todos los estudiantes, cuestión que evidentemente trasciende los resultados puramente académicos. Por ello, para poder evaluar esta característica, los estudios de eficacia escolar deberían incluir variables de tipo no cognitivo.



4. Los modelos multinivel para el análisis de datos educativos

Los modelos multinivel para el análisis de datos educativos

La revisión de la literatura sobre investigación en eficacia escolar llevada a cabo en el capítulo anterior explora uno de los principales retos metodológicos a los que se enfrentaron los investigadores educativos trabajando con datos empíricos en el pasado. Este reto se refiere a la falta de independencia entre las observaciones realizadas a los estudiantes de un mismo centro educativo. En respuesta a esta limitación, en la década de los años 80 varios autores (Aitkin & Longford, 1986; Raudenbush & Bryk, 1986; Goldstein, 1987) propusieron el análisis de datos a través de técnicas multinivel para modelar los datos educativos de manera más ajustada a la realidad.

El análisis multinivel es una técnica de regresión estadística especialmente indicada para conjuntos de datos en los que hay observaciones anidadas en otras de orden superior, en este caso, los estudiantes (nivel 1) se encuentran agrupados en centros educativos (nivel 2). En los modelos multinivel, se modela una variable de resultado como una combinación lineal de variables predictoras, cada una de las cuales está asociada a un coeficiente que determina en qué medida su variabilidad afecta a la variable criterio (Hayes, 2006).

En un estudio en el que se aplican técnicas de análisis multinivel existe una variable de resultado o respuesta registrada en el nivel más bajo, y variables explicativas en todos los niveles (Hox, 2010). Este carácter jerárquico de las observaciones supone una serie de patrones complejos de variabilidad, cuya falta de consideración puede implicar errores tanto de estimación como de interpretación.

Los estudios que hacen uso del análisis multinivel pueden presentar dos, tres (Han, 2018; Dumay & Dupriez, 2014; Willms, 2010), o incluso cuatro niveles de análisis (Murillo, 2007). En investigación educativa, además del nivel del estudiante y el centro educativo suele incluirse también el nivel de aula y/o niveles superiores de organización territorial, como regiones o países. No obstante, el presente capítulo se centra en los supuestos y especificaciones de los modelos de dos niveles, ya que, debido a la naturaleza de los datos disponibles, es la metodología utilizada en la parte empírica de esta tesis.

4.1. SUPUESTOS

El uso de modelos multinivel para el análisis de datos conlleva una serie de supuestos sobre la distribución de los elementos que los componen, así como de las relaciones que se establecen entre ellos.

Dependencia de las observaciones

Uno de los principales supuestos de la estadística inferencial, y por tanto de las técnicas habituales de regresión lineal, es la independencia entre las observaciones que conforman la muestra de estudio. Sin embargo, en el contexto de los análisis multinivel el supuesto ha de ser el contrario por definición, ya que las observaciones de nivel 1 dependen de las de nivel 2, y por tanto para poder ejecutar un modelo multinivel será necesario que las puntuaciones de nivel 1 muestren dependencia en función de la unidad de nivel 2 a la que pertenezcan. Este supuesto se comprueba a través del cálculo del llamado Índice de Correlación Intraclase (ver 4.2.1).

En nuestro caso, el rendimiento de los estudiantes dependerá del centro educativo al que pertenecen por diversos motivos (pertenecer al mismo barrio, compartir profesorado, provenir de entornos sociofamiliares parecidos, etc.), y cuanto más parecido sea el rendimiento de los estudiantes de un centro en comparación con los de otros centros, más probabilidad hay de que las causas del rendimiento estén relacionadas con el centro educativo (Snijders & Bosker, 2012).

Linealidad de la relación entre la variable criterio y las variables explicativas

Al igual que en la regresión tradicional, las técnicas de regresión lineal asumen que las relaciones establecidas entre la variable criterio y las variables predictoras tienen una naturaleza lineal.

Distribución y varianza de los residuos

Los residuos establecidos por los modelos para ambos niveles se refieren a la diferencia entre la puntuación estimada para una unidad a partir de las variables que conforman el modelo y su puntuación real.

En el análisis multinivel se asume que estos residuos presentan una distribución normal, con una media de 0 y una varianza constante (Raudenbush, Bryk, Cheong,

Congdon & du Toit, 2011). Este supuesto se refleja para cada nivel en su correspondiente ecuación (4.3 y 4.4) en el siguiente apartado.

4.2. ESPECIFICACIÓN DEL MODELO

El uso de modelos multinivel para el análisis de datos conlleva la especificación de las relaciones estructurales que ocurren en cada uno de los niveles, así como su variabilidad residual. A continuación, se examinan las ecuaciones básicas que guían dicha especificación, tanto para el modelo nulo o incondicional, como para el modelo condicional final.

4.2.1. El modelo nulo

El modelo nulo es un modelo incondicional compuesto solamente por el intercepto, sin ninguna variable predictora añadida (Hayes, 2006; Hox, 2010). La ecuación 4.1 representa la composición de este modelo.

$$Y_{ij} = \gamma_{00} + u_{0j} + r_{ij} \quad (4.1)$$

Donde:

Y_{ij} es la puntuación de la unidad de nivel 2 en la variable criterio;

γ_{00} es la media general de la variable criterio;

u_{0j} es la diferencia entre la puntuación en la variable criterio de la unidad de nivel 2 y la media general; y

r_{ij} es el residuo del nivel 1.

La estimación de este modelo “vacío” sirve para valorar en qué medida las unidades de nivel 2 difieren entre sí en la variable criterio, y también para estimar la magnitud de la dependencia de las puntuaciones en la variable criterio entre las unidades de nivel 1 correspondientes a la misma unidad de nivel 2 (Hayes, 2006). Esta relación se expresa a través del Coeficiente de Correlación Intraclase (CCI) (ecuación 4.2).

$$CCI = \frac{\tau_{00}}{\tau_{00} + \sigma^2} \quad (4.2)$$

Donde:

τ_{00} es la variabilidad entre las unidades de nivel 2 y

σ^2 representa la variabilidad entre las unidades de nivel 1.

Para poder llevar a cabo un análisis multinivel con un conjunto de datos concreto, se recomienda que el CCI sea de al menos el 10% (Lee, 2000), es decir, que al menos un 10% de la variabilidad de la variable criterio sea atribuible a las unidades de nivel 2.

4.2.2.El modelo condicional

Una vez comprobada la idoneidad del uso de análisis multinivel para el conjunto de datos que se pretende estudiar, puede procederse a la construcción del modelo condicional, que es aquel que contiene las covariables cuya influencia en la variable criterio quiere comprobarse.

Para comprender mejor este proceso de construcción, así como los elementos presentes en cada uno de los niveles del modelo, se presentan a continuación las ecuaciones para cada nivel (Raudenbush et al., 2011). En ambas ecuaciones, el subíndice i representa a los estudiantes y el subíndice j a los centros educativos.

En la ecuación del modelo del nivel 1 (4.3), el resultado para el caso i dentro de la unidad de segundo nivel j es la siguiente:

$$\begin{aligned}
 Y_{ij} &= \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \dots + \beta_{Qj}X_{Qij} + r_{ij} & (4.3) \\
 &= \beta_{0j} + \sum_{q=1}^Q \beta_{qj}X_{qij} + r_{ij}
 \end{aligned}$$

Donde:

β_{qj} ($q = 0, 1, \dots, Q$) son los coeficientes de nivel 1;

X_{qij} es el predictor q del nivel 1 para el caso i en la unidad j ; y

r_{ij} es el efecto aleatorio del nivel 1, cuya varianza es σ^2 .

Se asume que el término aleatorio $r_{ij} \sim N(0, \sigma^2)$

Cada uno de los coeficientes de nivel 1 (β_{qj}) contenidos en la anterior ecuación se convierten en variables de resultado en el modelo de nivel 2 (ecuación 4.4):

$$\begin{aligned}
 \beta_{qj} &= \gamma_{q0} + \gamma_{q1}W_{1j} + \gamma_{q2}W_{2j} + \dots + \gamma_{qS_q}X_{S_qj} + u_{qj} & (4.4) \\
 &= \gamma_{0j} + \sum_{s=1}^{S_q} \gamma_{qs}W_{sj} + u_{qj}
 \end{aligned}$$

Donde:

γ_{qs} ($q = 0, 1, \dots, S_q$) son los coeficientes de nivel 2;

W_{sj} es un predictor del nivel 2; y

u_{qj} es el efecto aleatorio del nivel 2.

Se asume que para cada unidad j , el vector $(u_{0j}, u_{1j}, \dots, u_{Qj})'$ tiene una distribución normal multivariante, donde cada elemento de u_{qj} tiene una media de 0 y una varianza de τ_{qq} .

4.3. LOS EFECTOS ESCOLARES Y EL VALOR AÑADIDO

La aplicación de los modelos multinivel no sólo aporta información sobre el valor del coeficiente de las variables explicativas incluidas, sino que también puede utilizarse para valorar en qué medida el centro realiza una aportación significativa al rendimiento de los estudiantes. Esta contribución puede abordarse a través de dos conceptos clave: los efectos escolares y el valor añadido.

Los efectos escolares pueden definirse como “la capacidad de los centros educativos para influir en los resultados de sus alumnos” (Murillo, 2007, p.34), es decir, en qué medida el hecho de estar escolarizado en uno u otro centro puede afectar al rendimiento de un estudiante. En los modelos multinivel, este efecto se estudia a través de los residuos de segundo nivel, que constituyen indicadores de la diferencia entre la puntuación real de un centro y aquella esperada en función de las covariables que conforman el modelo.

Existen varios tipos de efectos escolares identificados en la literatura en función de las características que se tienen en cuenta para su cálculo. Los primeros dos tipos de efectos escolares descritos son los efectos A y B (Raudenbush & Willms, 1995). El efecto A indica la variación en el rendimiento de estudiantes con las mismas características contextuales que acuden a distintas escuelas. Dicho de otra manera, si tomamos un grupo de X estudiantes de idénticas características demográficas y socioeconómicas y los asignásemos aleatoriamente a X escuelas, el efecto A de cada una de las escuelas sería la diferencia entre el rendimiento de su estudiante asignado y la media de todos los estudiantes X (Raudenbush & Willms, 1995). Este efecto no podría considerarse información sobre el nivel de eficacia de una escuela, ya que al no tomar en consideración datos sobre los factores de contexto del centro, se encuentra influido por variables ajenas a él.

Por su parte, el efecto B sí controla las características contextuales del centro educativo, considerando el efecto en el rendimiento de los estudiantes de factores externos al control de la escuela como su entorno socioeconómico o la composición de su cuerpo estudiantil (Raudenbush & Willms, 1995).

Más adelante se incorporaron a la literatura otros dos tipos de efectos que controlan la incidencia de un número mayor de factores escolares en el análisis. Los efectos de tipo X (Keeves, Hungi & Afrassa, 2005) añaden a los factores contextuales del centro otras variables relacionadas con sus políticas que escapan al control de las escuelas, como su tamaño, el tamaño de las clases o su titularidad, o factores de estratificación como la región a la que pertenecen. Por último, los efectos de tipo Z (Darmawan & Keeves, 2006) incluyen además aquellas variables relacionadas con las políticas y prácticas escolares en las que el centro sí tiene influencia directa, como por ejemplo el agrupamiento de estudiantes en clases en función de su nivel de capacidad académica.

Todos estos efectos pueden calcularse en estudios de corte transversal, en los que existe una única medida del rendimiento de los estudiantes. Sin embargo, en la tradición de la investigación en eficacia escolar se planteó la necesidad de estudiar los efectos de la escuela no sólo en el nivel de rendimiento adquirido en un punto concreto en el tiempo, sino en el progreso académico conseguido por los estudiantes a lo largo del tiempo, es decir, estudios longitudinales que añadiesen el nivel previo de rendimiento de los estudiantes.

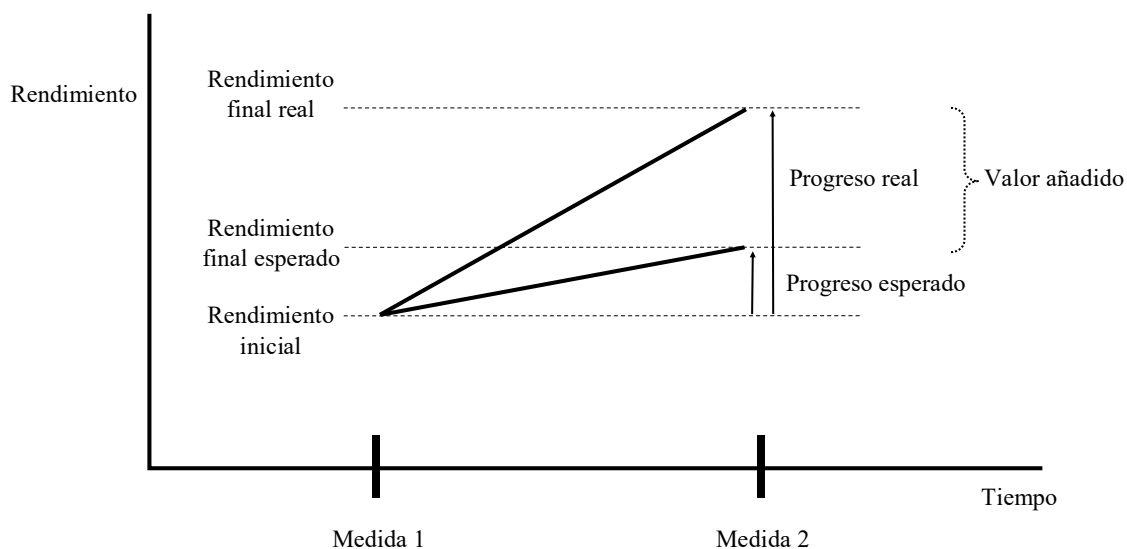


Figura 4.1. Representación gráfica del valor añadido. Adaptado de Martínez-Arias et al. (2009).

Este enfoque dio lugar al término *valor añadido* (figura 4.1), que puede caracterizarse como la contribución de un centro educativo “al progreso neto de los estudiantes hacia objetivos de aprendizaje establecidos, una vez eliminada la influencia de otros factores ajenos a la escuela que pueden contribuir a dicho progreso” (Martínez-Arias, Gaviria & Castro, 2009, p.17).

Existen diversos tipos de modelos que pueden utilizarse para el estudio de la magnitud de los efectos escolares o el valor añadido de un centro, diferenciados en función del tipo de covariables que utilizan (contextualizados y no contextualizados) y el número de medidas consecutivas del rendimiento que emplean (figura 4.2).

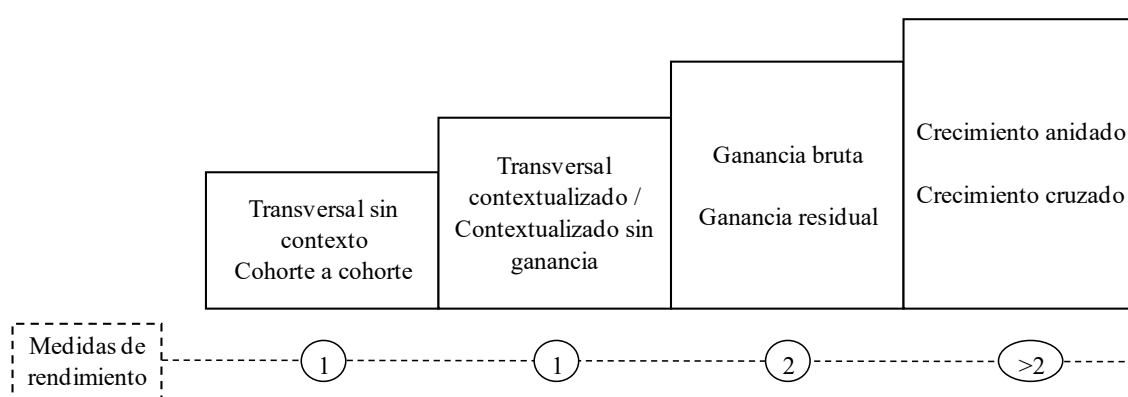


Figura 4.2. Modelos para el cálculo de los efectos escolares.

Entre los estudios transversales pueden encontrarse modelos que simplemente emplean los datos de rendimiento, sin ninguna covariable, los cuales sirven para comparar el rendimiento de una escuela con la media de un grupo de escuelas o un estándar determinado, y comprobar cuántos estudiantes lo superan.

Un caso especial de este tipo de modelo es el análisis de cohorte a cohorte. Aunque en este modelo se toman varias medidas de rendimiento a lo largo del tiempo, no se mide siempre a los mismos estudiantes, sino a diferentes cohortes en el mismo punto del sistema educativo (curso), por lo que se trata de modelos transversales consecutivos. Al igual que en los modelos transversales simples, el objetivo es observar cuántos alumnos superan el estándar de rendimiento prefijado y, en este caso, comparar el progreso de una escuela en el cumplimiento de dicho estándar.

Los modelos transversales también pueden tener en cuenta los efectos del contexto en el rendimiento, añadiendo covariables como el nivel socioeconómico o el estatus migratorio. De esta manera, los modelos transversales contextualizados, también

llamados modelos contextualizados sin ganancia (Martínez, Lizasoain, Castro & Joaristi, 2017), permiten el estudio de los efectos del contexto en el rendimiento, así como de la diferencia entre el rendimiento real de un centro y el esperado en función de sus características. Aunque esto no puede considerarse valor añadido como tal, ya que sólo se dispone de una puntuación de rendimiento, es un primer acercamiento a la contribución específica de los centros educativos en el rendimiento de sus estudiantes una vez se ha controlado el efecto del contexto.

Por otro lado, se encuentran los modelos que se centran en estudiar el cambio en las puntuaciones de una cohorte de estudiantes a través de diversas mediciones de su rendimiento a lo largo del tiempo, denominándose *ganancia* a aquellos modelos con dos mediciones, y *crecimiento* a los que tienen en cuenta al menos tres recogidas de datos.

Por su parte, los modelos de ganancia disponen de dos mediciones de rendimiento de los mismos estudiantes en distintos puntos temporales, por lo que pueden utilizarse para evaluar la evolución de sus puntuaciones. Estos modelos pueden ser de dos tipos en función de cómo se introduzcan ambas mediciones dentro del modelo. Se denomina *ganancia bruta* a aquellos modelos que introducen la diferencia entre el rendimiento actual y el previo como variable dependiente, y *ganancia residual* a los que incluyen la primera puntuación de rendimiento como covariable en el análisis de la segunda (Martínez, Lizasoain, Castro & Joaristi, 2017).

Por último, los modelos de crecimiento toman más de dos medidas del rendimiento de los mismos estudiantes a lo largo de tiempo, situando así el foco en la magnitud del cambio en el aprendizaje de los estudiantes que una escuela es capaz de conseguir en un periodo determinado de tiempo. Estos modelos pueden tomar dos formas diferenciadas, dependiendo de si conciben que los estudiantes se mantienen dentro de la misma aula/centro a lo largo de todo el proceso o tienen en cuenta los posibles cambios que pueda haber durante el periodo de estudio. Los modelos de crecimiento de efectos anidados responden a la primera opción, mientras que los modelos de crecimiento de efectos cruzados introducen en la estimación de los parámetros los cambios de aula o centro que puedan experimentar los estudiantes, de manera que sus datos estén en todo momento asociados a la unidad de análisis correcta.

4.4. CARACTERÍSTICAS METODOLÓGICAS

La aplicación de modelos multinivel conlleva la toma de una serie de decisiones metodológicas que dependerán de los datos de los que se dispone y de los objetivos de investigación. A continuación, se describen los principales tipos de estimación, tanto de coeficientes como de residuos, y las dos formas de modelar los interceptos y las pendientes de las covariables (efectos fijos y aleatorios).

4.4.1. Métodos de estimación

A la hora de llevar a cabo un análisis multinivel, pueden emplearse fundamentalmente cuatro técnicas de estimación de parámetros: máxima verosimilitud o máxima verosimilitud restringida para los coeficientes del modelo, empírica bayesiana para los residuos y estimación robusta para los errores.

Las técnicas de máxima verosimilitud y máxima verosimilitud restringida (ML y REML respectivamente, por sus siglas en inglés) estiman los parámetros a través de procesos iterativos que maximizan la idoneidad de su ajuste a las observaciones realizadas en la muestra (Boedeker, 2017). La principal diferencia entre estas dos técnicas es que la ML asume que los coeficientes de las variables son conocidos, y por tanto han sido medidos sin error, mientras que la REML asume que los coeficientes fueron estimados, y por tanto su estimación de la varianza es más robusta. Si bien es cierto que las diferencias entre ambos métodos en la estimación de los coeficientes de regresión son reducidas, sí se encuentran diferencias relevantes en cuanto a la estimación de los componentes de la varianza, ya que la ML presenta un sesgo a la baja del que la REML carece, derivado de que esta última técnica tiene en cuenta la pérdida de grados de libertad resultante de la estimación de los parámetros de regresión (Snijders & Bosker, 2012). Sin embargo, las diferencias entre estas dos técnicas se aprecian más claramente en muestras de pequeño tamaño (pocas unidades de segundo nivel, o pocas unidades de primer nivel en cada una de segundo nivel), para las que es más recomendable el uso de la REML (Boedeker, 2017).

Los estimadores empíricos de Bayes proporcionan información sobre los residuos del modelo, es decir, la diferencia entre la puntuación estimada por el modelo y la puntuación real de una unidad (Raudenbush et al., 2011). Estos estimadores combinan información de un grupo concreto con la del resto de grupos a través de técnicas bayesianas, aplicando a la media del grupo un factor de corrección (*shrinkage*) en función

de su fiabilidad o tamaño, con el objetivo de acercar casos extremos o menos fiables a la media de la muestra (Snijders & Bosker, 2012). Estos estimadores tienen dos ventajas principales a la hora de calcular el efecto de los centros en el rendimiento de los estudiantes, que son una alta correlación con la estimación del efecto verdadero, y la compensación de posibles estimaciones espurias debidas a muestras reducidas (Sanders & Wright, 2008).

Por último, las técnicas de máxima verosimilitud suelen dar como resultado la estimación de los errores estándar a través de la llamada matriz de información de Fisher (Snijders & Bosker, 2012). Sin embargo, estos errores pueden presentar sesgos de estimación procedentes de desviaciones de la normalidad en su distribución, o a la correlación entre grupos presente en las estructuras jerárquicas. Por ello, para análisis multinivel es más recomendable utilizar estimaciones robustas del error (también conocidas como estimadores *sándwich*) que tengan en cuenta la estructura anidada de los datos.

4.4.2.Efectos fijos y aleatorios

A la hora de plantear un estudio multinivel, otra de las decisiones que el investigador debe tomar es si los coeficientes de la regresión (intercepto y covariables) toman el mismo valor en todas las unidades de segundo nivel (efectos fijos) o si su valor puede variar entre unas unidades y otras (efectos aleatorios).

En general, en la investigación educativa el intercepto suele incluirse en los modelos multinivel como elemento aleatorio, ya que se espera que las unidades del nivel más alto de la jerarquía (clases, centros, etc.) varíen entre sí en función de su media en la variable dependiente (Hayes, 2006). Por su parte, los efectos de las covariables – o pendientes – sí pueden encontrarse modelados como fijos o aleatorios. Estas posibilidades dan lugar a dos tipos distintos de modelos, los de efectos mixtos (intercepto aleatorio y pendientes fijas) y los completamente aleatorios (intercepto y pendientes aleatorias).

En el modelo de pendientes fijas, las pendientes de las covariables se mantienen constantes entre las diferentes unidades de mayor nivel. Se utiliza en general cuando el objetivo de investigación es la inferencia sobre aquellos grupos que conforman la muestra.

En cambio, en el modelo de pendientes variables, cada covariable conserva su efecto fijo, pero se añade un término de variabilidad que permite que su valor cambie en función de la unidad de mayor nivel (Hayes, 2006). Este modelo es en principio el preferible cuando se pretende generalizar las conclusiones del análisis a todas las unidades que conforman la población de la que proviene la muestra (Rabe-Hesketh, & Skrondal, 2008), ya que se asume que las observaciones con las que se trabaja son una muestra pequeña de la población. También se recomienda este enfoque cuando el objetivo es estudiar la eficacia diferencial de los centros, es decir, el efecto diferencial de las covariables de interés en cada uno de los centros (Clarke, Crawford, Steele & Vignoles, 2010).

La elección del tipo de efectos a utilizar en una investigación debe basarse, por tanto, en la intención del estudio y en la naturaleza de los datos. Sin embargo, en algunas investigaciones con datos procedentes de evaluaciones a gran escala que han abordado la tarea de comprobar los modelos derivados del uso de pendientes fijas y aleatorias no se han encontrado grandes diferencias en cuanto a la estimación de parámetros. Así, diversos autores (Özdemir, 2016; Martínez-Garrido, 2015) plantean su toma de decisiones acerca de los efectos de las pendientes de las covariables en función del principio de parsimonia ya que, en igualdad de condiciones, el modelo de pendientes fijas presenta un menor número de parámetros y es por tanto más simple. Los métodos para comprobar la idoneidad de los modelos son diversos, pudiendo comparar las diferencias entre los coeficientes de ambos modelos (Özdemir, 2016), comparar las razones de verosimilitud de los modelos (Martínez-Garrido, 2015) o utilizar las estimaciones convencional y robusta del error como comprobación diagnóstica. En caso de que estas últimas presenten una gran diferencia, será una señal de que el modelo no se encuentra bien especificado, posiblemente debido a un establecimiento incorrecto de los efectos de las pendientes (Bell, Fairbrother & Jones, 2018).

4.5. VENTAJAS, LIMITACIONES Y COMPARATIVA CON OTRAS TÉCNICAS

Como ya se mencionó en la introducción, la principal ventaja de los modelos multinivel radica en la posibilidad de introducir datos de dos o más niveles distintos, posibilitando así el correcto análisis de realidades intrínsecamente jerárquicas como la educativa.

Sin embargo, el uso de esta técnica conlleva algunas limitaciones relacionadas con cuestiones como la necesidad de muestras grandes o la complejidad de su interpretación,

por lo que a continuación se expone una breve comparativa del uso de otras técnicas de regresión para el tratamiento de datos anidados, exponiendo las ventajas e inconvenientes de cada una de ellas.

La regresión lineal puede aplicarse a datos jerárquicos de dos maneras fundamentales: la desagregación y la agregación. En el primer caso, el foco del estudio se sitúa en el nivel más bajo, desagregando las variables de nivel 2 y situándolas a nivel individual. Este método supone ignorar la variación entre grupos que pueda existir e incumplir por tanto el supuesto de la dependencia de las observaciones (Woltman, Feldstain, MacKay & Rocchi, 2012). Adicionalmente, la desagregación puede llevar también a conclusiones erróneas a través de la falacia atomista, la cual ocurre cuando se extraen conclusiones sobre la variabilidad de los grupos basadas en datos individuales (Gaviria y Castro, 2014).

La agregación, por su parte, supone el proceso opuesto, es decir, agregar las variables individuales en función del nivel superior, de manera que el análisis se realice sobre los promedios de dichas variables. A través de este proceso se soluciona el problema de la dependencia de las observaciones, pero puede perderse hasta un 80-90% de variabilidad debida a las diferencias individuales, lo cual resulta en interpretaciones erróneas de las relaciones entre las variables (Raudenbush & Bryck, 1992). Al igual que el método anterior, esta técnica incurre en una falacia de interpretación, en este caso la llamada falacia ecológica, que se deriva del uso de datos grupales para extraer conclusiones sobre los individuos que conforman dichos grupos (Gaviria y Castro, 2014).

La regresión lineal permite otras opciones para el tratamiento de estos datos, como el establecimiento de una regresión independiente para cada uno de los grupos o la introducción de las unidades de segundo nivel como variables *dummy* en el análisis de variables individuales. Sin embargo, ambos métodos adolecen de la misma limitación principal; no permiten la exploración de los efectos de las variables de segundo nivel en la variable criterio ni de sus interacciones con las variables de primer nivel, y sólo resultan prácticos con un número reducido de unidades de nivel 2 (Diez-Roux, 2000).

La aplicación de modelos de regresión multinivel puede responder a todas las limitaciones presentadas por los otros métodos de regresión, ya que permite respetar la variabilidad correspondiente a cada uno de los niveles y analizar todas las unidades de

manera conjunta. Sin embargo, conviene considerar algunas de las limitaciones potenciales del empleo de esta técnica de análisis de datos.

Una de las principales desventajas de los modelos multinivel es el tamaño de la muestra necesaria. Para poder aplicar esta técnica con garantías (precisión y falta de sesgos en los coeficientes de regresión, los componentes de varianza y los errores típicos) es necesario disponer de una muestra suficientemente amplia de unidades al nivel más alto de análisis. Aunque el número exacto de unidades de mayor nivel necesarias depende de la complejidad del modelo (número de variables y relaciones planteadas entre ellas), una muestra aceptable podría estar compuesta como mínimo de entre 50 unidades para modelos sencillos hasta 100 unidades para modelos con mayor complejidad (Maas & Hox, 2005).

Por otro lado, el análisis simultáneo de diversos niveles supone una mayor complejidad a la hora de plantear las relaciones entre las variables, realizar los cálculos, e interpretar los resultados, y en consecuencia genera modelos menos parsimoniosos que otras técnicas tradicionales de regresión (Diez-Roux, 2000).

Por último, cabe destacar que los modelos multinivel no permiten el estudio de más de una variable criterio a la vez (Hox & Kreft, 1994). En el caso del estudio del rendimiento estudiantil, poder incorporar varias medidas de rendimiento en diversas materias en el mismo modelo sería útil e informativo, ya que no sólo permitiría el análisis de la relación entre las variables predictoras y cada una de las variables de rendimiento, sino que posibilitaría el estudio de la relación de estas últimas entre sí. Aunque los modelos multinivel tradicionales no permiten este tipo de análisis, existen algunos avances metodológicos en este sentido, como el uso de modelos de ecuaciones estructurales de múltiple indicador y múltiple causa (MIMIC), los cuales admiten el uso de variables criterio distintas (Tsai, Smith & Hauser, 2017).



5. Método

Método

El trabajo de investigación que a continuación se presenta se compone principalmente de dos estudios (figura 5.1), que se abordan de manera conjunta desde un enfoque de métodos mixtos secuenciales explicativos (Creswell, 2014).

En la primera fase, de naturaleza cuantitativa, se utiliza la información procedente de los instrumentos de PISA para seleccionar y caracterizar los centros educativos de alta y baja eficacia en España, así como para establecer qué variables de contexto están significativamente relacionadas con el rendimiento de los estudiantes en las tres competencias evaluadas. La segunda fase se sirve de una metodología cualitativa para realizar un estudio más profundo de alguno de los centros seleccionados por su nivel de eficacia (alta o baja), con el objetivo de evaluar los procesos organizativos y pedagógicos que puedan estar relacionados con la eficacia de los mismos.

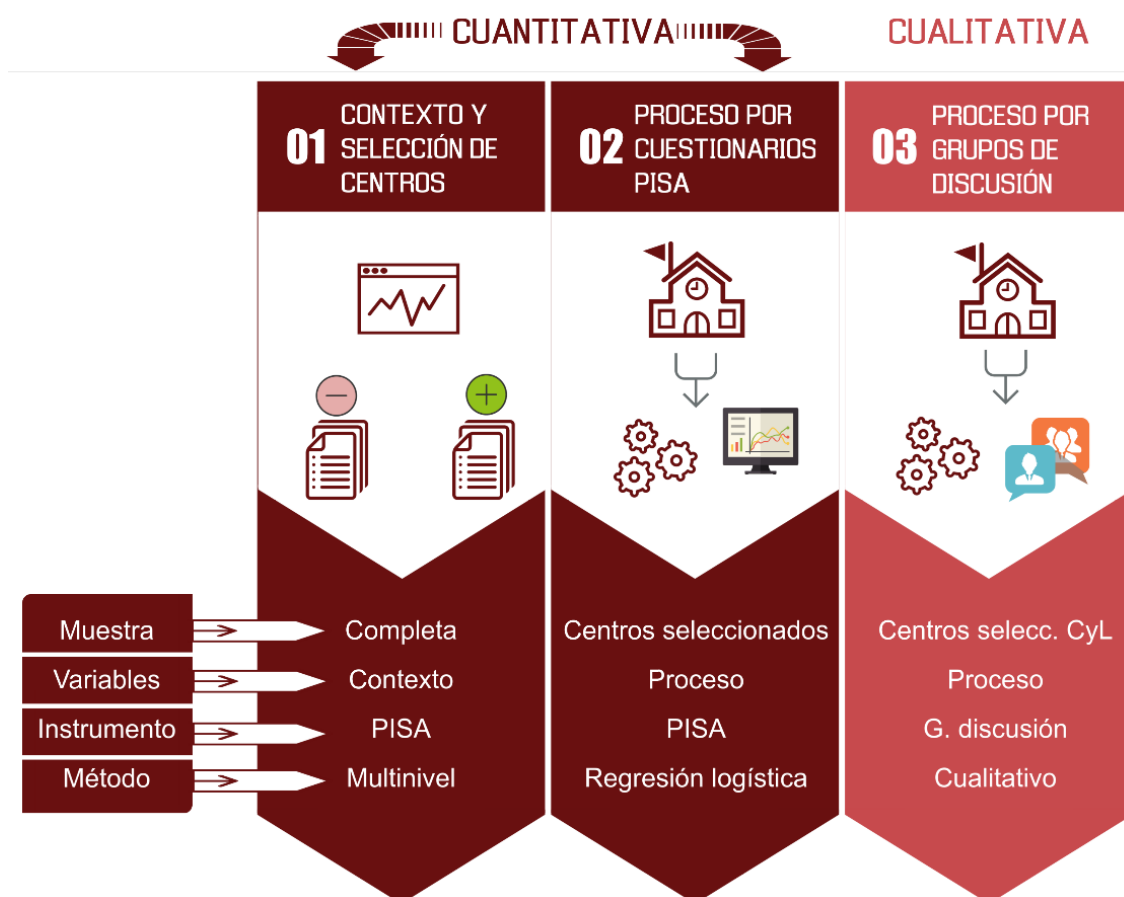


Figura 5.1. Resumen de las fases de la investigación.

A continuación, se exponen los objetivos, muestras, instrumentos, variables y procedimientos de cada una de estas fases.

5.1. OBJETIVOS

El objetivo general del presente trabajo de investigación es detectar aquellas prácticas escolares que se relacionan con la eficacia de los centros educativos de Educación Secundaria en Castilla y León, de cara a establecer una propuesta de buenas prácticas educativas que fomenten la eficacia.

Este objetivo general se concreta a su vez en tres objetivos específicos que guiarán la planificación de la investigación en sus diversas fases. Estos objetivos específicos son:

1. Estudiar los factores contextuales personales, escolares y familiares asociados al rendimiento de los alumnos participantes en PISA 2015 en España.
2. Detectar centros de alta y baja eficacia en PISA en España y Castilla y León.
3. Estudiar los factores de proceso asociados con la eficacia de los centros educativos de España y Castilla y León.

5.2. LOS MÉTODOS MIXTOS

El enfoque de métodos mixtos supone la puesta en práctica de procedimientos de recogida, análisis e interpretación de datos cuantitativos y cualitativos con el objetivo de estudiar el mismo fenómeno en el marco de un mismo estudio (Anguera, Portell, Chacón-Moscoso & Sanduvete-Chaves, 2018). Aunque el creciente interés de la comunidad científica por este enfoque a lo largo de las últimas décadas ha generado diversas opiniones acerca de los conceptos y definiciones básicos, Johnson, Onuwuegbuzie y Turner (2007) realizaron el esfuerzo de aunar las opiniones de los principales expertos en el campo para consensuar una definición:

Los métodos mixtos son un tipo de investigación en la cual un investigador o equipo de investigadores combinan elementos de los enfoques cuantitativo y cualitativo (p. ej. perspectivas, recogida de datos, análisis o técnicas de inferencia cuantitativas o cualitativas) con el objetivo general de alcanzar una comprensión y corroboración tanto amplias como profundas. (Johnson et al., 2007, p. 123)

El doble objetivo presentado por la última parte de la definición – amplitud y profundidad – es una de las principales ventajas de los métodos mixtos, ya que permiten aprovechar las ventajas de ambos enfoques metodológicos y beneficiarse del extenso

alcance de las investigaciones cuantitativas combinado con la comprensión multidimensional y profunda que proporcionan las técnicas cualitativas.

Los distintos tipos de métodos mixtos pueden representarse en un continuo (figura 5.2), cuyos extremos están representados por los métodos cuantitativos y cualitativos puros y los métodos mixtos ocupan el centro del espectro, generándose así varios tipos de enfoques mixtos.

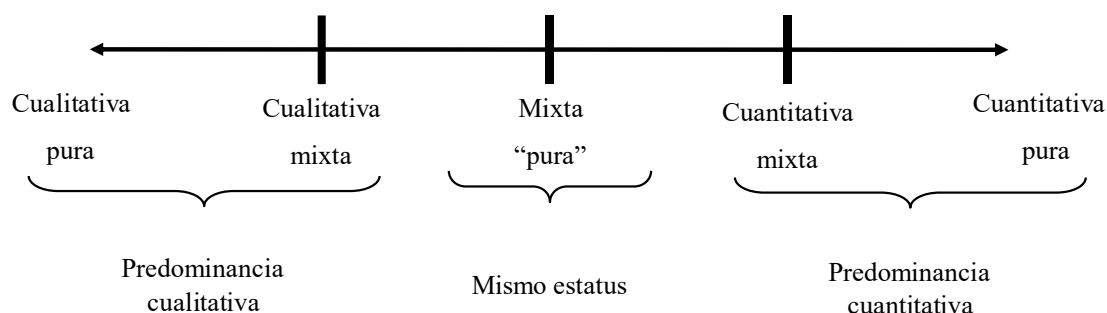


Figura 5.2. Enfoques de investigación (adaptado de Johnson et al., 2007).

Mediante la combinación de los distintos enfoques, y en función de la predominancia de uno u otro, además de los tres enfoques en su forma pura pueden observarse dos sub-enfoques que representan una complementariedad metodológica con un equilibrio desigual de influencias. Por su parte, la metodología cualitativa mixta muestra la predominancia del uso de datos, métodos y análisis cualitativos, que se complementan de alguna manera con métodos cuantitativos. Por el contrario, la metodología cuantitativa mixta responde a diseños en los que la perspectiva cuantitativa guía la mayor parte del proceso investigador, pero se reconoce el valor que puede añadir el uso de técnicas cualitativas complementarias (Johnson et al., 2007).

Por otro lado, los estudios que hacen uso de métodos mixtos pueden caracterizarse también en función de la secuenciación de las fases cuantitativa y cualitativa en la recogida y el análisis de datos. Así, los estudios mixtos pueden ser principalmente de carácter secuencial, en los que el investigador se centra en expandir los hallazgos de un método con otro diferente, o concurrente, en los que se recogen y analizan los datos cuantitativos y cualitativos de manera simultánea (Creswell, 2014). En el caso de los estudios mixtos secuenciales, habitualmente es la fase cuantitativa la que precede a la cualitativa (secuencial explicativo), aunque en función de los objetivos de los investigadores, también pueden encontrarse estudios que utilicen un orden inverso de las fases (secuencial exploratorio) (Ivankova, Creswell & Stick, 2006).

Una vez expuesta la naturaleza y distintas categorizaciones de los métodos mixtos, cabe caracterizar el presente estudio como una investigación secuencial explicativa con predominancia cuantitativa, en la que la recogida y análisis de los datos cuantitativos preceden a la fase cualitativa. La determinación de este orden de fases se sustenta en el razonamiento de que el análisis cuantitativo proporcionará a los investigadores una visión amplia y general del problema de investigación, y a partir de ellos la fase cualitativa supondrá una exploración más profunda del tema a través del estudio de las vivencias de los participantes.

5.3. FASE 1: ESTUDIO CUANTITATIVO

El estudio cuantitativo está a su vez compuesto por diversas fases, cada una de ellas dirigida a un objetivo concreto, y con su metodología y variables propias.

5.3.1. Objetivos

El estudio cuantitativo pretende abarcar todos los objetivos planteados para esta investigación mediante dos procedimientos diferenciados. El objetivo 1, referido al estudio del efecto de los factores contextuales en el rendimiento de los estudiantes, se aborda en la primera parte del estudio empírico, y sus resultados sirven como base para alcanzar el objetivo 2, la selección de centros de alta y baja eficacia. Por último, en la segunda parte de la fase cuantitativa también se realiza un primer acercamiento al objetivo 3 a través del estudio de los factores de proceso medidos por PISA asociados a la eficacia de los centros españoles.

5.3.2. Diseño

El estudio empírico sigue un diseño de carácter no experimental o ex-post-facto, ya que se trata de un análisis secundario de datos recogidos de manera ajena a la presente investigación. Asimismo, también constituye un estudio transversal, ya que la medición se ha realizado en un punto único en el tiempo.

5.3.3. Población y Muestra

La prueba PISA va dirigida a todos aquellos estudiantes entre 15,3 y 16,2 años de edad en el momento de la evaluación. En la aplicación de 2015, la población española del estudio se compuso por todos aquellos estudiantes que cumplían ese criterio de edad, es decir, que hubiesen nacido en 1999, y estuviesen escolarizados en un centro ordinario de educación secundaria localizado en territorio español.

Como ya se expuso en el capítulo 1 (sección 1.3.1), el proceso de muestreo es bietápico; primero se seleccionan los centros educativos y después se seleccionan estudiantes susceptibles de ser evaluados dentro de dichos centros. Con el objetivo de contar con los efectos de dicho diseño muestral, el presente estudio incorpora los pesos muestrales en todos los procedimientos estadísticos

En 2015, la muestra española estuvo compuesta por 32.330 estudiantes y 976 centros educativos (OCDE, 2017). Sin embargo, para la realización de este estudio se eliminaron aquellos centros con menos de 20 estudiantes evaluados con el objetivo de asegurar un correcto análisis de las variables agrupadas a nivel de centro (Hayes, 2006; Joaristi, Lizasoain & Azpillaga, 2014; Martínez-Abad, Lizasoain, Castro & Joaristi, 2017; Meunier, 2011). La muestra final quedó compuesta por 31.273 estudiantes procedentes de 897 centros educativos.

En esta edición de la prueba, todas las Comunidades Autónomas (CCAA) decidieron ampliar su muestra para poder ser comparadas a nivel internacional, por lo que se cuenta con una muestra superior a la de otros países participantes (Ministerio de Educación, Cultura y Deporte, 2016). La distribución de estudiantes según su región de procedencia es relativamente uniforme; cada Comunidad Autónoma aporta entre un 4,5% y un 6% del total de la muestra final (tabla 5.1). La única excepción es el País Vasco, que representa el 10,7% del total de participantes¹¹.

¹¹ El aumento de la muestra en esta región se debe a su interés por introducir la modalidad lingüística del centro como variable de estudio, para lo cual se necesita una población lo suficientemente representativa de cada una de las tres modalidades.

Tabla 5.1. Número de estudiantes y centros participantes y porcentaje que representan de la muestra total, por Comunidad Autónoma

CCAA	Estudiantes		Centros	
	n	%	n	%
Andalucía	1779	5,7	52	5,8
Aragón	1756	5,6	49	5,5
Asturias, Principado de	1723	5,5	49	5,5
Balears, Illes	1732	5,5	50	5,6
Canarias	1755	5,6	49	5,5
Cantabria	1876	6,0	53	5,9
Castilla y León	1806	5,8	53	5,9
Castilla-La Mancha	1810	5,8	50	5,6
Cataluña	1744	5,6	50	5,6
Comunitat Valenciana	1570	5,0	49	5,5
Extremadura	1773	5,7	50	5,6
Galicia	1769	5,7	52	5,8
Madrid, Comunidad de	1791	5,7	50	5,6
Murcia, Región de	1780	5,7	52	5,8
Navarra, Comunidad Foral de	1852	5,9	50	5,6
País Vasco	3360	10,7	99	11,0
Rioja, La	1397	4,5	40	4,5

Fuente: Elaboración propia a partir de la base de datos de PISA 2015 (OECD, 2017).

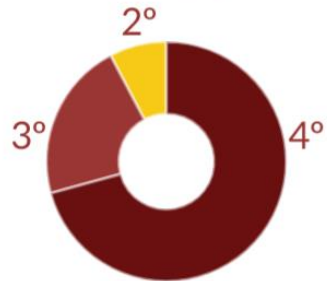
La base de datos de PISA proporciona numerosas variables de caracterización de la muestra, como cuestiones demográficas (género, estatus migratorio, etc.) o aspectos relacionados con la escolarización de los estudiantes (repetición de curso, número de cambios de centro realizados, etc.). Como puede observarse en la figura 5.3, la muestra presentó una distribución equitativa en cuanto al género (50.6% de mujeres), una mayoría de estudiantes de 4º de ESO (70%), no repetidores (71%) y que no han experimentado ningún cambio de centro educativo (65%), así como una mayoría de estudiantes en cuyos hogares se habla la lengua en la que realizaron las pruebas competenciales (83%) y una predominancia de estudiantes nativos (83%) seguidos por los inmigrantes de 1ª generación (9%).

MUESTRA DE ESTUDIANTES

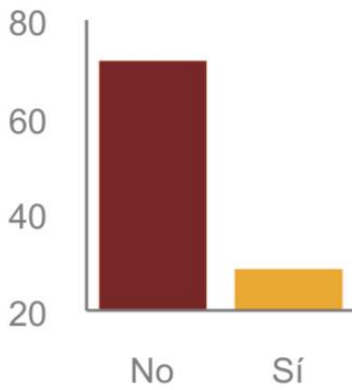
50% GÉNERO



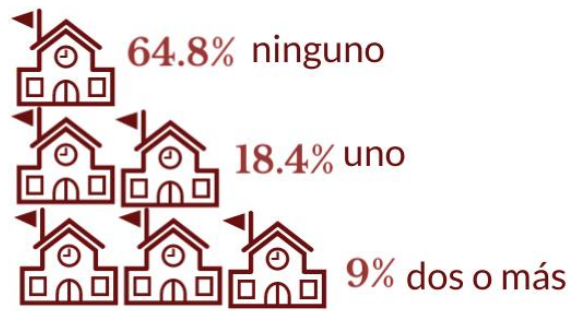
CURSO



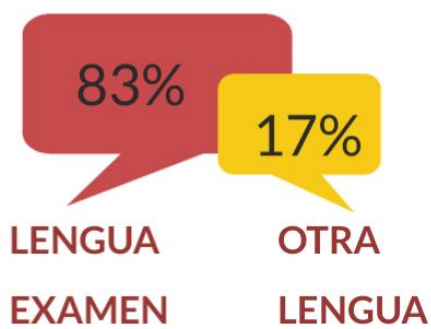
REPETICIÓN



CAMBIOS DE CENTRO



LENGUA EN EL HOGAR



ESTATUS MIGRATORIO

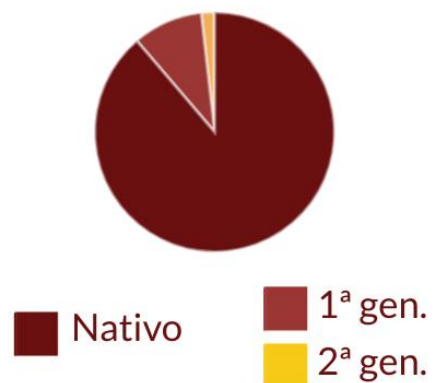


Figura 5.3. Características de la muestra de estudiantes.

Otro de los datos relevantes sobre la muestra de estudiantes es la distribución de la variable que determina su índice socioeconómico y cultural (ESCS) (figura 5.4). En la muestra española, esta variable presenta una media de -0.45 y una desviación típica de 1.15.

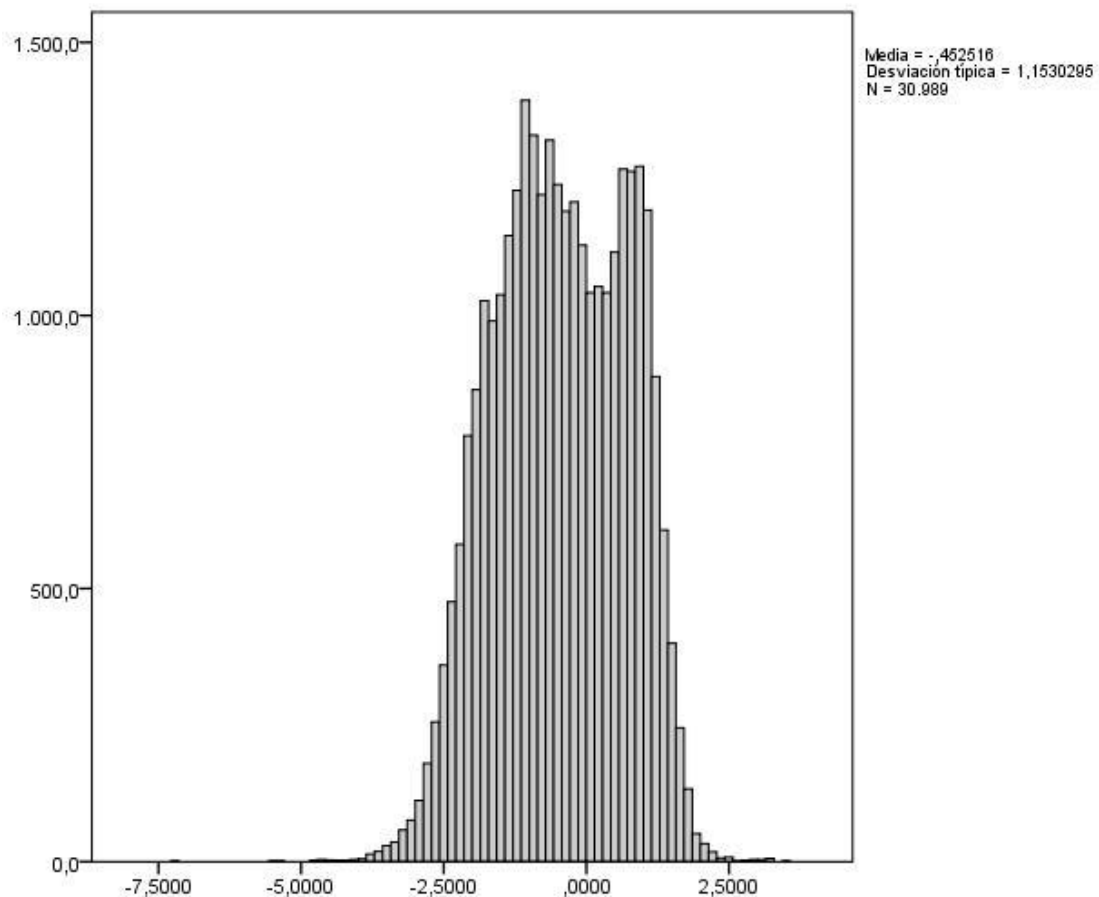


Figura 5.4. Distribución de la variable ESCS para la muestra española.

Por su parte, los centros que componen la muestra (figura 5.5) son en su mayoría públicos (66%), tienen un tamaño medio de 736 estudiantes, y sus clases tienen 27 estudiantes y una ratio profesor-estudiante promedio de 12.5.

MUESTRA DE CENTROS

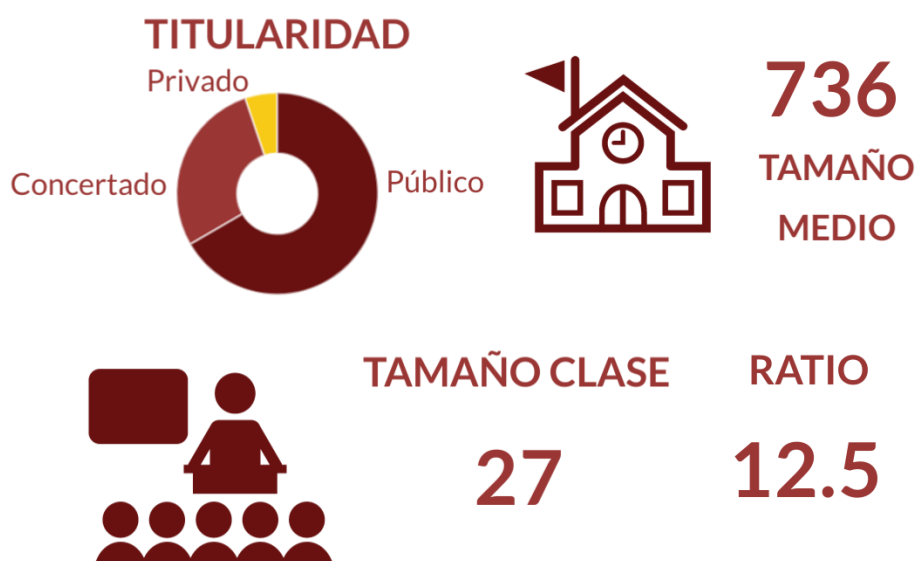


Figura 5.5. Características de la muestra de centros.

5.3.4. Instrumentos

Los instrumentos utilizados para esta fase son los cuestionarios propios de la evaluación PISA 2015, ya descritos en el capítulo 1 (sección 1.2). Concretamente, se utilizaron tanto los cuestionarios de evaluación del nivel competencial de los estudiantes en lectura, matemáticas y ciencias, como los cuestionarios de contexto aplicados a estudiantes y centros educativos.

5.3.5. Variables

Como se ha introducido en la figura 5.1, en esta etapa del estudio se hace uso de distintas clases de variables, que debido a la naturaleza no experimental del diseño son denominadas *criterio* y *predictoras*.

Para el primer análisis, las variables criterio utilizadas son los niveles competenciales de los estudiantes en lectura, matemáticas y ciencias (la definición de estas variables puede encontrarse en la sección 1.2.1). La puntuación de cada estudiante participante en cada una de estas variables está comprendida por 10 valores plausibles, los cuales fueron utilizados de manera simultánea para realizar los cálculos pertinentes.

Por otro lado, las variables predictoras son los factores contextuales – o de entrada – tanto de estudiantes como de centros educativos (tabla 5.2). La información para la mayoría de estas variables procede solamente de un ítem, cuyo tipo de respuesta se indica

en la tabla 5.2. Sin embargo, cabe definir tres variables concretas (ESCS, escasez de recursos y escasez de profesorado), ya que se trata de índices compuestos por información procedente de diversos ítems.

La variable ESCS está compuesta por datos sobre el nivel educativo y la ocupación de los padres, así como los recursos disponibles en el hogar, incluyendo equipos electrónicos o digitales, recursos educativos, libros, o posesiones generales relacionadas con el poder adquisitivo de la familia.

Por su parte, la escasez de recursos del centro informa sobre los materiales de enseñanza y la infraestructura del centro (insuficiencia e inadecuación), y el índice de escasez de personal se establece a partir de información sobre la insuficiencia e inadecuación del personal docente y el personal de apoyo.

Tabla 5.2. Variables de contexto

	Variable	Tipo	Rango¹
Nivel 1 - Alumnos	Género	Nominal	0: Masculino 1: Femenino
	Mes de nacimiento	Discreta	1 – 12
	Curso	Ordinal	1º ESO – 1º Bachillerato
	Índice socioeconómico y cultural (ESCS)	Continua	-7.58 – 3.52
	Estatus migratorio	Nominal	0: Nativo, 1: Inmigrante de 2ª generación 2: Inmigrante de 1ª generación
	Condición de repetidor	Nominal	0: No 1: Sí
	Número de cambios de centro educativo	Nominal	0: Ningún cambio 1: Un cambio 2: Dos o más cambios
	Idioma hablado en el hogar	Nominal	0: Idioma de la prueba 1: Otro idioma
	Tamaño del centro	Discreta	25 – 4034
	Tamaño de las clases	Discreta	13 – 53
Nivel 2 - Centros	Escasez de recursos	Continua	-1.32 – 3.63
	Escasez de profesorado	Continua	-1.68 – 3.72
	Titularidad del centro	Nominal	1: Privado 2: Concertado 3: Público
	Ratio profesor-alumnos	Continua	1 – 50.48
	ESCS medio	Continua	-1.92 – 1.14
	Tasa de repetidores	Continua	0 – .79
	Tasa de alumnos inmigrantes	Continua	0 – .85
	Porcentaje de chicas	Continua	0 – 1

¹Las variables cualitativas con más de dos categorías se convirtieron en variables *dummy* (tantas como el número de categorías de la variable original menos una, siendo la categoría de referencia siempre la más frecuente).

En este punto cabe señalar una cuestión sobre las variables de identificación del alumnado descubierta durante la fase cualitativa de estudio, que es la participación de estudiantes con algún tipo de necesidad educativa especial o necesidad específica de apoyo educativo en la prueba. A este respecto, el Informe Técnico de PISA 2015 (OECD, 2017b) expone que uno de los motivos por los que pueden excluirse alumnos de la muestra para participar en la prueba es su condición de alumno NEE¹², definida por las siguientes categorías: discapacidad intelectual que conlleve un retraso cognitivo suficiente como para no poder participar en la prueba, discapacidad funcional (física) que impida la participación en la prueba, estudiantes cuya lengua nativa no es la vehicular y que han recibido menos de un año de instrucción en la lengua vehicular, o estudiantes no evaluables por otros motivos (categoría definida a nivel nacional). Por otro lado, este informe indica también que

En interés de asegurar que PISA fuera lo más inclusiva posible, la participación de los estudiantes y las razones para su exclusión fueron codificadas por separado en el *Student Tracking Form*¹³. Este hecho permitió que estudiantes NEE pudiesen ser incluidos en la muestra cuando sus necesidades no fuesen lo suficientemente serias como para constituir una barrera a su participación. (OCDE, 2017b, p. 86)

Por lo tanto, PISA permite hasta cierto punto la participación de estudiantes con NEE cuando estas no son lo suficientemente graves como para garantizar su exclusión. Sin embargo, los informes de la OCDE no dejan claro si los datos de estos estudiantes son incluidos en la base de datos pública junto con el resto de los datos. Esta falta de transparencia en la documentación hace necesario acudir a otras fuentes para poder comprender el alcance de este asunto.

¹² La definición de alumno con necesidades educativas especiales recogida en este punto se basa en el término *Special Education Needs* expresado por la documentación de PISA, el cual no se corresponde completamente con la definición de Alumnos con Necesidades Específicas de Apoyo Educativo, que es el término actualmente recogido en la legislación española.

¹³ Formulario a rellenar por el coordinador de PISA del centro educativo, en el que se recogen los datos de los estudiantes que participan (o son excluidos) en la evaluación.

LeRoy, Samuel, Deluca y Evans (2018) reportan en su artículo la existencia de una base de datos privada, disponible para los investigadores bajo demanda, en la que se encuentran los datos de los estudiantes NEE que han participado en la prueba. A este respecto, Barabara LeRoy, profesora e investigadora en el *Michigan Developmental Disabilities Institute* (Wayne University) y coautora del mencionado artículo, confirma la existencia de esta base de datos separada para los estudiantes NEE, aunque afirma que a veces hay estudiantes que son percibidos como NEE por el centro, pero no encajan en la definición que la OCDE hace de esta condición, por lo que deben participar si son seleccionados y se les incluye en la base de datos general (B. LeRoy, comunicación personal, 13 de junio de 2019). Esta autora sugiere que nos pongamos en contacto con la organización española encargada de la aplicación de PISA para aclarar el tratamiento de los datos de estos estudiantes en nuestro país, y, en su caso, solicitar la base de datos separada.

En España, la evaluación PISA es organizada por el Instituto Nacional de Evaluación Educativa (INEE), por lo que se dirigió la consulta al personal técnico encargado de la prueba dentro de esta organización. Desde el INEE se informó de que, en España, PISA no se aplica a estudiantes con NEE diagnosticadas. Aunque reconocen la existencia de la modalidad *Une Heure* para este tipo de alumnado, manifiestan que a su juicio no está bien adaptada a sus características y por ello se ha decidido no participar a nivel nacional. Sin embargo, informan de una nueva modalidad para poder aplicar la prueba a alumnado con NEE que se encuentra en fase de pre-piloto (L. Cercadillo, comunicación personal, 17 de junio de 2019).

Por lo tanto, puede concluirse que, en la base de datos general española puede haber alumnos con algún tipo de necesidad específica de apoyo que pueda afectar a su rendimiento en la prueba, y que estos alumnos no se encuentran identificados a través de ningún tipo de variable, lo cual puede suponer una dificultad a la hora de analizar los datos.

Para la siguiente fase, la variable criterio se deriva del análisis anterior, y consiste en una variable dicotómica que indica si un centro pertenece al grupo de alta o de baja eficacia. En este caso, las variables predictoras son los factores de proceso ofrecidos en

la base de datos de PISA en forma de índices (es decir, aquellas que agrupan información de varios ítems en un solo factor) tanto del nivel de estudiantes como de centros¹⁴.

Las variables de nivel 1 analizadas fueron motivación, disposición al trabajo colaborativo, creencias epistemológicas, disfrute e interés por la ciencia, competencia, interés y autonomía en el uso de las nuevas tecnologías, clima de disciplina en el aula de ciencias, apoyo del profesor, conciencia y optimismo medioambiental, estatus ocupacional esperado, pertenencia al centro, apoyo emocional y académico parental, retroalimentación percibida e imparcialidad del profesorado.

Las variables de nivel 2 empleadas fueron: liderazgo, desarrollo del currículum, desarrollo profesional, responsabilidad del centro sobre los recursos y el currículum, participación del profesorado, autonomía escolar y actividades extraescolares de carácter creativo.

5.3.6. Procedimiento y análisis de datos

El estudio de los factores contextuales asociados al rendimiento se realizó a través de la aplicación de modelos jerárquicos lineales (Snijders & Bosker, 2012). Se definió un modelo diferente para cada una de las competencias evaluadas, cuya versión final consistió sólo en aquellas variables predictoras que mostraron una relación significativa con la variable ($\alpha = .05$). La construcción de los modelos multinivel permitió el cálculo de la diferencia entre la puntuación real de los centros y aquella estimada en función de las variables contenidas en el modelo. Esta diferencia también se denomina residuo, y se calcula a través de estimadores de Bayes empírico (Raudenbush et al., 2011). Tras un estudio de la colinealidad de las variables predictoras, y partiendo de las evidencias mostradas por Özdemir (2016) y Martínez-Garrido (2015) sobre la falta de diferencias sustanciales entre modelos multinivel de pendientes fijas y aleatorias a partir de datos de las pruebas PISA, se decidió evaluar las diferencias entre ambos modelos con los datos de PISA 2015 para España, decidiendo al final introducir las variables contextuales en el primer nivel como covariables de efectos fijos, sin incluir los efectos aleatorios. Todos los modelos incluyeron en su cálculo los 10 valores plausibles del nivel competencial, los

¹⁴ La formulación de los ítems que componen estos índices puede consultarse en el Anexo II.

pesos muestrales de los estudiantes y los centros, y la estimación robusta de los errores (ver capítulo 4 para más información sobre estos elementos).

Para determinar los centros de alto y bajo residuo, se elaboró un protocolo con los criterios de selección partiendo, por un lado, del trabajo de Joaristi et al. (2014) en el que fue fijado el percentil 80 como región límite para los 6 modelos estimados. Por otro lado, se tiene en cuenta el trabajo de Martínez-Abad et al. (2017), en el que se optó por establecer como límite inferior (alta eficacia) el percentil 66 y como límite superior (baja eficacia) el 33, requiriendo el cumplimiento en 5 de los 8 modelos. Así, y dados los datos disponibles, el presente trabajo optó por un criterio ecléctico entre ambos procedimientos, tomando los percentiles 33 y 66 como regiones límite, pero estableciendo la necesidad de su cumplimiento en los 3 modelos obtenidos finalmente.

A continuación, se analizó la correlación entre las principales variables de proceso (nivel 1 y 2) y la variable dicotómica generada (centro de alto o bajo residuo, variable criterio) a través del coeficiente de correlación biserial puntual. En caso de alcanzarse correlaciones significativas ($\alpha=.05$), y tras descartarse efectos importantes de colinealidad entre las variables predictoras del modelo (de proceso), se procedió con la aplicación de técnicas regresión logística para el estudio multivariante de la relación de las variables de proceso (predictoras) con la variable anterior. Estas técnicas son las más recomendables para variables criterio de naturaleza dicotómica.

La construcción de los modelos jerárquicos lineales fue realizada con el programa HLM7, y la regresión logística se llevó a cabo con el paquete estadístico SPSS v.20.

5.4. FASE 2: ESTUDIO CUALITATIVO

La segunda fase del estudio empírico consistió en el planteamiento de una investigación de corte cualitativo cuyo objetivo fue profundizar en las cuestiones de proceso que caracterizan a los centros de alta y baja eficacia a través de la realización de una serie de grupos de discusión.

Antes de describir el diseño y procedimiento de esta fase, cabe destacar algunas características de la metodología cualitativa relevantes a la hora de plantear e interpretar una investigación de estas características. Así como la investigación cuantitativa se apoya en la objetividad de los datos recogidos y los análisis realizados, en la investigación cualitativa se plantean una serie de visiones subjetivas sobre la realidad estudiada

(incluyendo la del propio investigador) y, por tanto, el objetivo es interpretar estas visiones para poder alcanzar un significado *intersubjetivo*, un conocimiento común sobre el tema (Ballesteros, 2014). Por otro lado, otro de los pilares fundamentales de este tipo de investigación es el *holismo*, desde el cual se concibe el objeto de estudio como un fenómeno complejo que ha de ser estudiado de manera profunda y en su totalidad. Dado que la conducta de los seres humanos (y de las organizaciones) no está compartimentada, sino que sus elementos se encuentran interrelacionados, es necesario interpretarla de manera global (Ballesteros, 2014).

Así en esta fase se plantea el estudio de los procesos que se llevan a cabo en los centros educativos a través de diversos informantes, para poder obtener diferentes testimonios que permitan una interpretación intersubjetiva y holística del fenómeno de estudio.

5.4.1. Objetivos

Esta fase se centró fundamentalmente en el estudio del objetivo 3, es decir, el estudio de los factores relacionados con la organización y las prácticas educativas realizadas en los centros de alto y bajo residuo identificados en Castilla y León.

5.4.2. Enfoque

Dentro de las diversas visiones bajo las que puede plantearse una investigación cualitativa (etnográfico, narrativo, fenomenológico, teoría fundamentada y estudio de casos), el presente trabajo se sitúa en el enfoque de estudio de casos, que puede definirse como el “estudio de la particularidad y complejidad de un caso concreto, comprendiendo su actividad dentro de circunstancias relevantes” (Stakes, 1995, p. xi).

Existen distintos tipos de estudio de casos en función del objeto de estudio seleccionado y la intención de los investigadores, pudiendo encontrarse estudios instrumentales, colectivos e intrínsecos. La diferencia entre los dos primeros tipos es el número de casos seleccionados para ilustrar el problema de investigación, uno o varios respectivamente, mientras que el tercer tipo se refiere a estudios en los que el foco de interés es el caso en sí mismo debido a que presenta alguna particularidad (Creswell, 2007). En función de estas definiciones, la investigación presentada en esta tesis responde al segundo tipo, el estudio de casos colectivo.

5.4.3.Muestra

La muestra del estudio cualitativo estuvo compuesta por dos centros educativos participantes en la evaluación PISA 2015 en Castilla y León, seleccionados por su relevancia para la investigación (un centro de alto residuo y otro de bajo residuo). Con el objetivo de preservar su anonimato, en este trabajo se identificarán como centro A (alta eficacia) y centro B (baja eficacia). Ambos centros se encuentran localizados en una capital de provincia de la citada Comunidad Autónoma y presentan las siguientes características (tabla 5.3):

Tabla 5.3. Características de los centros de la fase cualitativa.

Variable	Centro A	Centro B
Titularidad	Público	Concertado
Tamaño	1303	949
Ratio profesor/alumno	18	15
ESCS medio	-0.34	0.36
% alumnado inmigrante	16%	3%
% alumnado repetidor	31%	23%
Residuo lectura	21.06	-21.73
Residuo matemáticas	14.94	-18.77
Residuo ciencias	21.5	-26.69

5.4.4.Informantes

El estudio de casos se realizó a través de información procedente de dos tipos de informantes: docentes y equipos directivos de los dos centros seleccionados (tabla 5.4).

En relación con el profesorado, se establecieron en cada centro dos grupos compuestos por entre seis y ocho docentes¹⁵, uno de los grupos dirigido a docentes con menos de 10 años de experiencia en el centro y el otro para docentes con más de 10. Esta distinción se realizó con el objetivo de obtener la perspectiva de profesionales con distintos grados de antigüedad en el centro. Con motivo de facilitar la exposición de los resultados, a estos grupos se les han asignado las etiquetas de docentes *noveles* y *veteranos* respectivamente, aunque cabe señalar que, si bien los docentes con menos experiencia del centro B podrían ser descritos como noveles en el sentido estricto de la palabra, los del centro A en general tienen una experiencia previa más o menos amplia en

¹⁵ Por motivos de organización e incompatibilidad de horarios, alguno de los grupos estuvo compuesto por menos integrantes, siendo cuatro el número mínimo de informantes en un grupo.

otros centros educativos. Adicionalmente, también se propuso la presencia de docentes de distintas áreas de enseñanza (lenguas, ciencias, etc.).

Por otro lado, también se contactó con los equipos directivos de ambos centros, en los que se recogió la opinión de agentes como directores, jefes de estudios o secretarios del centro. En el centro B, la sesión estuvo compuesta por dos informantes, por lo que se caracteriza como una entrevista grupal en lugar de un grupo de discusión.

Tabla 5.4. Informantes de la fase cualitativa

Participantes	Centro A	Centro B
Docentes	Noveles 4 docentes – 1,5-7 años de antigüedad	5 docentes – 1-3 años de antigüedad
	Veteranos 4 docentes – 11-17 años de antigüedad	6 docentes – 23-35 años de antigüedad
Equipo directivo	5 participantes (Director, Jefa de Estudios General, Jefas de Estudios de ESO y Bachillerato, y Secretario)	2 participantes (Jefe de Estudios y Director de Secundaria)

5.4.5. Instrumentos

Para la recogida de datos de esta fase se utilizaron grupos de discusión, que constituyen una técnica no directiva centrada en el planteamiento de determinados temas de interés para su debate por parte de los integrantes del grupo participante (Gil Flores, 1993).

Para su preparación, se utilizó como base el sistema de categorías establecido por Lizasoain et al. (2015), compuesto por los siguientes nueve temas: proyectos y formación, metodologías, atención a la diversidad y seguimiento del alumnado, evaluación, gestión del tiempo, liderazgo, gestión y organización del centro, clima, y familias y comunidad.

La planificación de los grupos de discusión, y el subsiguiente tratamiento de la información obtenida, se realizó teniendo en cuenta las fases delimitadas por Suárez Ortega (2005) para este tipo de técnicas:

- Exploratoria: acercamiento al tema de estudio, realizado a través de una profunda revisión de la literatura sobre el estudio de los factores escolares de proceso en el ámbito de la eficacia escolar.
- Preparatoria: establecimiento de los informantes y las circunstancias concretas de la aplicación de la técnica (tiempo, espacio, etc.).

- Aplicación: realización de los grupos de discusión de acuerdo con la planificación previa.
- Análisis e interpretación de la información: selección y aplicación de las técnicas de análisis de los datos obtenidos.
- Elaboración del informe: síntesis de las conclusiones extraídas a partir del análisis.

En cuanto al papel del moderador, pueden distinguirse fundamentalmente dos enfoques: el psicoanalítico, en el que el moderador se limita a plantear el tema y escuchar, y el conductista-cognitivo, en el que el papel del moderador consiste en dinamizar la sesión a través de preguntas que promuevan el debate (Flick, 2007). En este caso se optó por el segundo enfoque, con el objetivo de fomentar la reflexión entre los participantes y poder recoger una mayor cantidad de información relevante.

5.4.6.Procedimiento

El primer paso para poder llevar a cabo la fase cualitativa de este estudio fue la obtención de los nombres de los centros educativos correspondientes a los códigos designados por PISA. A partir de la selección de centros de interés en función de su eficacia llevada a cabo en la fase anterior, el equipo de investigación se puso en contacto con la Consejería de Educación de la Junta de Castilla y León con el objetivo de obtener los datos de identificación de los centros seleccionados en esta Comunidad Autónoma. Tras mantener una reunión informativa sobre el proyecto con los representantes de esta Comunidad Autónoma, y presentarles posteriormente una planificación detallada de las evaluaciones que pretendían llevarse a cabo, la Consejería de Educación facilitó el acceso a los nombres de los centros educativos de alta y baja eficacia participantes en PISA 2015 en Castilla y León.

Una vez obtenida la información, se seleccionaron dos centros concretos en función de criterios de conveniencia y accesibilidad. Se llevó a cabo un primer contacto telefónico con los directores de ambos centros para explicarles los objetivos del proyecto, y se concertó una cita con ambos para desarrollar la propuesta con mayor profundidad. En estas reuniones, mantenidas a lo largo del mes de diciembre de 2018, se explicó a los directores el desarrollo del proyecto de investigación, así como el valor y alcance de su participación en su fase cualitativa. Adicionalmente, se les hizo entrega de un acuerdo de confidencialidad en el que se exponían las cláusulas relativas al cumplimiento del Real Decreto 1720/2007, de 21 de diciembre, por el que se aprueba el Reglamento de

desarrollo de la Ley Orgánica 15/1999, de 13 de diciembre, de protección de datos de carácter personal, y del Reglamento General de Protección de Datos. Dicho acuerdo (Anexo III) define lo que se entiende por información confidencial en el marco del proyecto (fundamentalmente cualquier tipo de información que permita a personas ajenas al proyecto identificar los centros) y cuáles son los protocolos de gestión de dicha información.

Entre los meses de abril y mayo de 2019 se concertaron y llevaron a cabo los grupos de discusión, que fueron grabados con el consentimiento de los participantes y transcritos para su posterior análisis. Previamente a la recogida de información, se informó a los participantes sobre la estructura y los temas a tratar en la sesión a través de documentos enviados por e-mail (Anexo III).

5.4.7. Análisis de datos

La principal técnica de análisis de datos que guio esta parte de la investigación fue el análisis cualitativo de contenido de las transcripciones de todos los grupos de discusión realizados.

El análisis cualitativo implica “ordenar y organizar la información, elaborar patrones, categorías y unidades de análisis” (Tójar, 2006, p.284). Para poder llevar a cabo este proceso, hay que seguir una serie de pasos que operativicen el análisis de datos cualitativos (figura 5.6), como el establecimiento de las categorías, la separación de los textos en unidades, la clasificación de estas unidades en función de las categorías, y su síntesis para la presentación de los resultados.

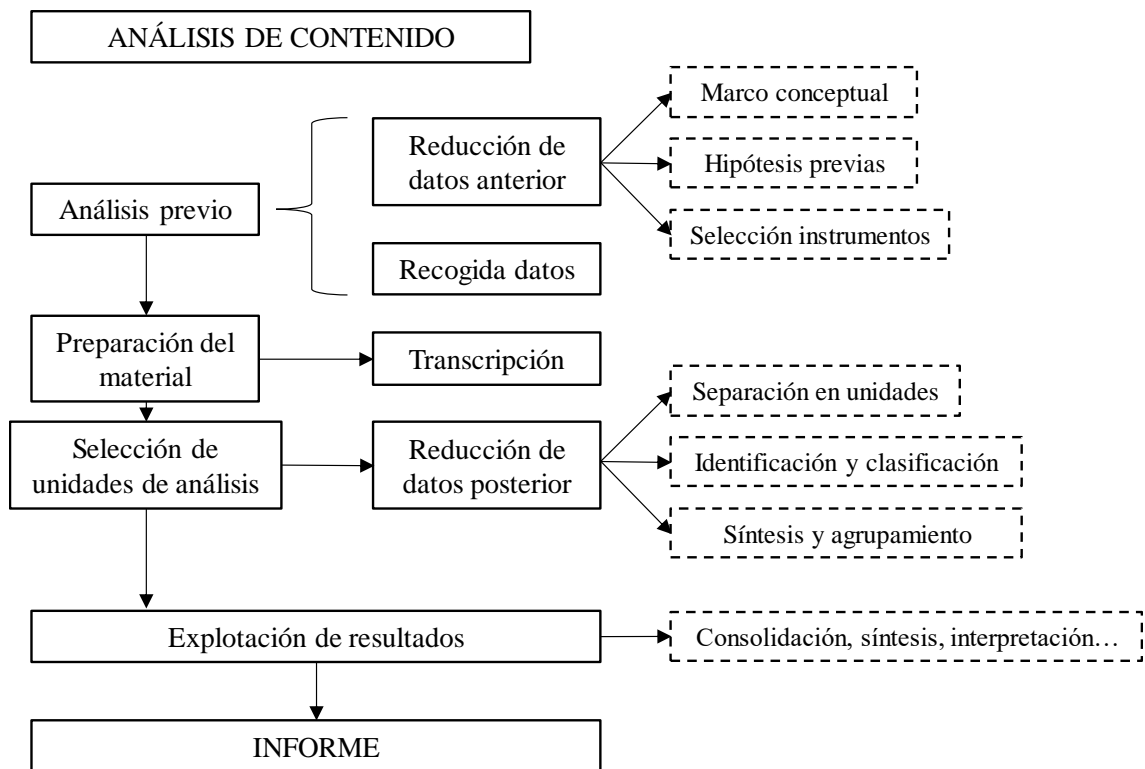


Figura 5.6. Procedimiento de análisis cualitativo de datos.

Los procedimientos de reducción cualitativa de datos suponen una buena parte del proceso, ya que se llevan a cabo tanto antes como durante el análisis. En la parte de análisis previo, es necesario establecer un marco conceptual con el cual poder guiar e interpretar el resto del proceso. De manera general, el presente trabajo se enmarca en la línea de la investigación en eficacia escolar (ver capítulo 3), y más concretamente se utiliza el sistema de categorías establecido por Lizasoain y Angulo (2014) para su estudio cualitativo de los factores escolares asociados a la eficacia. Asimismo, en esta fase deben seleccionarse los instrumentos de recogida de datos que, como ya se mencionó, en este caso son los grupos de discusión.

Tras la recogida de la información, se procedió a la transcripción de las grabaciones realizadas en las sesiones con el apoyo del programa de transcripción automática *Transcribe Wreally*. Cada grupo de discusión se transcribió en un documento separado, y posteriormente se introdujeron en el programa NVivo v.XX para su posterior análisis.

La siguiente fase es la selección de unidades de análisis a través de la reducción de datos, cuyos principales procedimientos son la separación en unidades, la identificación y clasificación de unidades, y la síntesis y agrupamiento de la información (Tójar, 2006).

La separación de los textos recogidos en unidades más pequeñas tiene por objetivo generar extractos de la información más manejables para su análisis. Las unidades pueden separarse en función de distintos criterios (temáticos, cronológicos, gramaticales, etc.), eligiéndose en este caso un criterio temático en función de las categorías que ya se han establecido a priori.

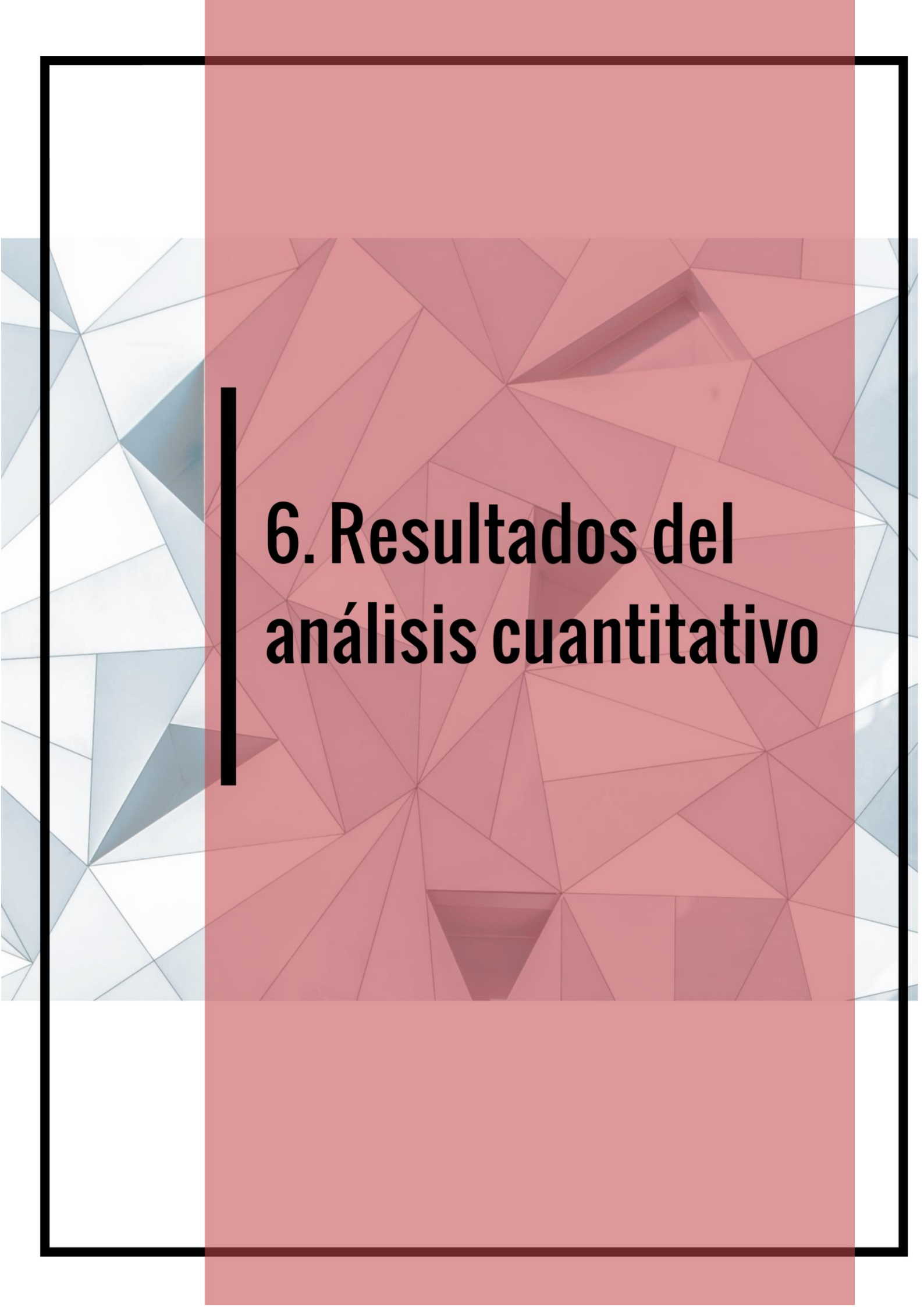
Una vez los documentos han sido separados en unidades de registro en función del tema al que se refieren, estas unidades han de ser reagrupadas en sus categorías correspondientes. Las unidades que pertenecen a la misma categoría han de presentar una cierta homogeneidad interna entre ellas, así como una heterogeneidad externa frente a otras categorías, aunque las categorías no son excluyentes y puede darse el caso de que compartan algunos fragmentos entre ellas (Tójar, 2006).

Este proceso de categorización puede realizarse fundamentalmente de tres maneras, abierta, cerrada o mixta. El procedimiento de categorización abierta es inductivo, es decir, no hay unas categorías preestablecidas, y estas van emergiendo a medida que se analizan los documentos. En el caso de la categorización cerrada, existe un sistema de categorías establecido a priori que guía el proceso, y no pueden añadirse o quitarse categorías durante el análisis. Por su parte, en los procedimientos mixtos, que son los elegidos para llevar a cabo el presente análisis, existe un sistema de categorías preestablecido, pero se permite que emerjan nuevas categorías no recogidas en dicho sistema durante el análisis, o que se combinen varias categorías en una sola en función lo que hayan expresado los participantes (Miles & Huberman, 1994).

Cuando el sistema de categorías ha sido establecido (o de manera paralela, en caso de que se utilice un procedimiento abierto o mixto), se procede a la asignación de las distintas unidades de registro a las categorías que se han determinado. En este proceso, que se denomina síntesis y agrupamiento, se llevan a cabo procedimientos de agregación y ordenación de unidades de registro, y también se comienzan a vislumbrar las relaciones entre las unidades que pertenecen a la misma categoría, o los vínculos establecidos entre las distintas categorías (Goetz y LeCompte, 1988).

Finalmente, el último paso consiste en realizar la explotación de los resultados y el informe a través de la descripción de la información contenida en cada una de las categorías, la interpretación de los significados expresados por los sujetos en cada uno de los temas, y la comparación de la información en función de los diferentes atributos que

les hemos asignado a cada uno de los recursos (centro – A o B – y puesto laboral – equipo directivo, docente novel o docente veterano), con el objetivo de comprobar si los distintos colectivos de un mismo centro expresan opiniones parecidas, o qué características comparten o diferencian a los dos centros investigados. En último término, tal como se exponía al principio de esta sección, el objetivo de esta fase es intentar construir una intersubjetividad que refleje la realidad del centro a través de las ideas expresadas por cada uno de los sujetos participantes, así como obtener una visión holística de los procesos pedagógicos y organizativos que caracterizan cada uno de los centros investigados.



6. Resultados del análisis cuantitativo

Resultados del análisis cuantitativo

En este capítulo se presentan los principales resultados de la parte cuantitativa del estudio, que consiste en la aplicación de modelos multinivel para el estudio de los efectos del contexto sobre el rendimiento de los estudiantes y para la selección de centros de alta y baja eficacia, y el estudio cuantitativo de la relación entre las variables de proceso de los centros y su eficacia.

En primer lugar, se presentan los datos de la construcción de los modelos multinivel para cada una de las competencias, detallando el proceso seguido para determinar los modelos finales, así como el cumplimiento de los supuestos (ver sección 4.1).

Una vez establecidos los modelos, se realiza la selección de centros de alta y baja eficacia a través del estudio de los residuos de segundo nivel de cada una de las competencias, y se hace una descripción de estos centros en función de sus características contextuales.

Por último, se procede al análisis de las variables de proceso proporcionadas por PISA en función de la selección de centros realizada, con el objetivo de investigar la posible relación entre las prácticas organizativas y pedagógicas descritas por los directores de los centros en los cuestionarios de contexto aplicados por PISA.

6.1. ANÁLISIS DE LOS EFECTOS DE LAS VARIABLES CONTEXTUALES EN EL NIVEL COMPETENCIAL DE LOS ESTUDIANTES ESPAÑOLES

Como ya se ha expuesto en diversas partes del documento, el estudio de los efectos del contexto sobre la adquisición de competencias de los estudiantes de secundaria se realiza en función de las tres competencias principales evaluadas por PISA: matemáticas, lectura y ciencias. Cabe recordar que estos análisis se han realizado con los 10 valores plausibles de cada una de las competencias como variable criterio, e introduciendo las ponderaciones que reflejan los procesos y el diseño de selección de la muestra. El análisis de cada una de las competencias se ajusta al siguiente esquema:

1. Modelo nulo y cálculo del Coeficiente de Correlación Intraclase (comprobación del supuesto de dependencia)
2. Prueba de colinealidad para la selección de variables a incluir en el modelo condicional.
3. Modelo condicional: Selección de pendientes fijas o aleatorias, presentación del modelo final.

4. Comprobación de los supuestos relacionados con los residuos del modelo (distribución normal, media 0 y homocedasticidad).

6.1.1. Consideraciones previas

En el análisis basado en modelos multinivel resulta necesario realizar un estudio previo de la colinealidad entre las variables (tablas 6.1 y 6.2), para evitar la inclusión de dos variables altamente relacionadas en el mismo modelo. Ya que las covariables de carácter contextual a incluir en el modelo condicional son las mismas para las tres competencias evaluadas, se expone a continuación este análisis que servirá de referencia para los tres modelos. Para cada par de variables se calcula el índice de correlación correspondiente en función del tipo de variable¹⁶.

Tabla 6.1. Correlaciones entre las covariables de nivel 1.

	1	2	3	4	5	6	7	8
1. Género	-							
2. Repetición	-.91	-						
3. Lengua	-.009	.053	-					
4. Inmigrante	.008	.187	.257	-				
5. Cambio centro	.026	.213	.1	.239	-			
6. Curso	.1	.693	.073	.221		-		
7. Mes nacimiento	.006	.055	-.007	-.015	.004	-.065	-	
8. ESCS	-.005	-.349	.053	.190	.217	.355	-.007	-

Como puede observarse, las únicas variables que presentan una correlación media-alta son curso y repetición, aunque esta no se considera suficientemente elevada como para justificar la retirada de una de estas covariables.

Tabla 6.2. Correlaciones entre las covariables de nivel 2.

	1	2	3	4	5	6	7	8	9	10
1. Tamaño centro	-									
2. Tamaño clase	.180	-								
3. Ratio	.464	-.149	-							
4. Escasez recursos	-.092	.004	-.147	-						
5. Escasez personal	-.062	-0.26	-.128	.447	-					
6. Titularidad	.761	.218	.805	.529	.549	-				
7. Proporción inmigrantes	-.028	-.023	-.119	.079	.060	.189	-			
8. Proporción repetidores	-.293	-.147	-.297	.204	.202	.413	.413	-		
9. Proporción alumnas	.037	-.003	.064	-.039	-.059	-.027	-.031	-.115	-	
10. ESCS centro	.449	.12	.416	-.293	-.248	-.545	-.314	-.718	.041	-

Nota: Todas las correlaciones por encima de .070 son significativas

¹⁶ Coeficiente Fi para los pares de variables dicotómicas, correlación biserial-puntual para pares de variables dicotómicas y cuantitativas, coeficiente de contingencia para pares de variables dicotómica-politómica, politómica-politómica, dicotómica-ordinal y politómica-ordinal, coeficiente de Spearman para pares de variables ordinales y ordinal-cuantitativa, y coeficiente de Pearson para pares de variables cuantitativas.

En general, no se presentan correlaciones elevadas entre ninguna de las variables, excepto en el caso de la titularidad del centro, que presenta una correlación alta o media-alta tanto con el tamaño del centro como con la ratio profesores-alumnos, así como el ESCS medio del centro en relación con la tasa de repetidores. La decisión sobre qué covariable utilizar habrá de tomarse en función de los resultados de cada modelo, en el caso de que en alguno de los modelos se presente alguno de estos pares de covariables de manera conjunta.

A continuación, se expone la construcción de los modelos multinivel para la competencia en matemáticas, lectura y ciencias.

6.1.2. Competencia matemática

Como primer paso en el establecimiento del modelo, es necesario calcular el modelo nulo o incondicional para comprobar el supuesto de dependencia de las observaciones, es decir, hasta qué punto las puntuaciones de rendimiento en matemáticas muestran variabilidad en el nivel del centro educativo. El modelo nulo para matemáticas se expresa a través de la ecuación 6.1.

$$\text{Matemáticas}_{ij} = \gamma_{00} + u_{0j} + r_{ij} \quad (6.1)$$

Donde:

Matemáticas_{ij} es la puntuación media del centro en matemáticas;

γ_{00} es la media de todos los centros en matemáticas;

u_{0j} es la diferencia entre la puntuación del centro en matemáticas y la media general de todos los centros; y

r_{ij} es el residuo del nivel 1.

La aplicación del modelo nulo permite el cálculo de los componentes de varianza de ambos niveles de análisis (tabla 6.3) para poder estudiar el Coeficiente de Correlación Intraclase.

Tabla 6.3. Componentes de varianza del modelo nulo para la competencia matemática.

Efecto aleatorio	Desviación Típica	Componente de varianza	g.l.	χ^2	<i>p</i> -valor
Nivel 2	30.2	911.99	895	5662.48	<0.001
Nivel 1	78.13	6104.24			

Con los datos de la tercera columna, que informan de la variabilidad presentada por los datos en el nivel de estudiantes y de centros, puede procederse a calcular el CCI (ecuación 6.2).

$$CCI_{\text{Matemáticas}} = \frac{\tau_{00}}{\tau_{00} + \sigma^2} = \frac{911.99}{(911.99 + 6104.24)} = 0.13 \quad (6.2)$$

A partir de este coeficiente, se puede afirmar que la variabilidad atribuible al nivel de centro educativo para el nivel competencial de los estudiantes españoles es de aproximadamente un 13%, por lo que se cumple el supuesto de dependencia de las observaciones y se considera adecuado realizar un análisis multinivel.

El siguiente paso en el proceso es ajustar el modelo multinivel en sí mismo mediante la inclusión de las covariables de nivel 1 y 2 de interés. En este caso, se sigue un procedimiento de eliminación paso a paso, en el que se van eliminando del modelo aquellas variables que no presentan una relación significativa con la puntuación en competencia matemática.

Para determinar qué tipo de efectos de pendientes (fijas o aleatorias) es más adecuado utilizar, se calculan ambos modelos para poder compararlos. En primer lugar, se presenta el modelo de intercepto aleatorio y pendientes fijas (ecuación 6.3)

$$\begin{aligned} \text{Matemáticas}_{ij} &= \gamma_{00} + \gamma_{01} * \text{ESCS nivel } 2_j + \gamma_{10} * \text{Género}_{ij} + \gamma_{20} \\ &* \text{Inmigrante } 1^{\text{a}} \text{ generación}_{ij} + \gamma_{30} * \text{Curso}_{ij} + \gamma_{40} \\ &* \text{Repetición}_{ij} + \gamma_{50} * \text{ESCS}_{ij} + \gamma_{60} * \text{Cambio centro } 1_{ij} \\ &+ \gamma_{70} * \text{Cambio centro } 2_{ij} + u_{0j} + r_{ij} \end{aligned} \quad (6.3)$$

Donde:

Matemáticas_{ij} es la puntuación media del centro en matemáticas;

γ_{00} es la media de todos los centros en matemáticas;

γ_{01} es la covariable de nivel 2;

$\gamma_{10} - \gamma_{70}$ son las covariables de nivel 1;

u_{0j} es la diferencia entre la puntuación del centro en matemáticas y la media general de todos los centros; y

r_{ij} es el residuo del nivel 1.

Tras la retirada una a una de aquellas variables que no presentan una relación significativa con la variable criterio, el modelo queda conformado por las variables indicadas en la tabla 6.4.

Tabla 6.4. Modelo de intercepto aleatorio y pendientes fijas para matemáticas.

Efecto fijo	Estimación robusta del error				Estimación ordinaria del error			
	Coef.	Error típico	t-ratio	p-valor	Coef.	Error típico	t-ratio	p-valor
Intercepto, γ_{00}	543.07	1.73	313.44	<0.001	543.07	1.34	404.55	<0.001
ESCS centro, γ_{01}	13.00	1.91	6.82	<0.001	13.00	1.46	8.9	<0.001
Género, γ_{10}	-22.58	1.98	-11.47	<0.001	-22.59	1.69	-13.33	<0.001
Inmigrante gen, γ_{20} 1 ^a	-11.81	4.85	-2.43	0.026	-11.81	4.34	-2.72	0.020
Curso, γ_{30}	34.86	3.85	9.05	<0.001	34.86	3.26	10.69	<0.001
Repetición, γ_{40}	-35.80	4.94	-7.24	<0.001	-35.80	4.05	-8.85	<0.001
ESCS, γ_{50}	9.79	1.12	8.73	<0.001	9.79	1.01	9.66	<0.001
Cambio de centro (1), γ_{60}	-10.65	1.86	-5.72	<0.001	-10.65	1.41	-7.53	<0.001
Cambio de centro (2+), γ_{70}	-15.60	3.30	-4.72	<0.001	-15.60	2.72	-5.73	<0.001

Por su parte, el modelo de intercepto aleatorio y pendientes aleatorias se presenta en la ecuación 6.4.

$$\begin{aligned}
 \text{Matemáticas}_{ij} &= \gamma_{00} + \gamma_{01} * \text{ESCS nivel } 2_j + \gamma_{10} * \text{Género}_{ij} + \gamma_{20} \\
 &* \text{Inmigrante } 1^{\text{a}} \text{ generación}_{ij} + \gamma_{30} * \text{Curso}_{ij} + \gamma_{40} \\
 &* \text{Repetición}_{ij} + \gamma_{50} * \text{ESCS}_{ij} + \gamma_{60} * \text{Cambio centro } 1_{ij} \\
 &+ \gamma_{70} * \text{Cambio centro } 2_{ij} + u_{0j} + u_{1j} * \text{Género}_{ij} + u_{2j} \\
 &* \text{Inmigrante } 1^{\text{a}} \text{ generación}_{ij} + u_{3j} * \text{Curso}_{ij} + u_{4j} \\
 &* \text{Repetición}_{ij} + u_{5j} * \text{ESCS}_{ij} + u_{6j} * \text{Cambio centro}_{ij} \\
 &+ u_{7j} * \text{Cambio centro } 2_{ij} + r_{ij}
 \end{aligned} \tag{6.4}$$

Donde¹⁷:

$u_{1j}-u_{7j}$ es la diferencia entre el coeficiente de la pendiente variable del centro en la covariable de nivel 1 para cada centro y su efecto fijo (componente aleatorio de las pendientes).

¹⁷ Para la especificación del resto de elementos, ver ecuación 6.3.

Los valores del modelo de pendientes aleatorias pueden observarse en la tabla 6.5.

Tabla 6.5. Modelo de intercepto y pendientes aleatorias para matemáticas.

Efecto fijo	Estimación robusta del error				Estimación ordinaria del error			
	Coef.	Error típico	<i>t</i> -ratio	<i>p</i> -valor	Coef.	Error típico	<i>t</i> -ratio	<i>p</i> -valor
Intercepto, γ_{00}	543.075	1.70	318.85	<0.001	543.07	1.36	398.59	<0.001
ESCS centro, γ_{01}	12.95	1.87	6.93	<0.001	12.95	1.45	8.91	<0.001
Género, γ_{10}	-22.4	1.97	-11.38	<0.001	-22.4	1.73	-12.95	<0.001
Inmigrante 1 ^a gen, γ_{20}	-10.26	4.88	-2.10	0.048	-10.26	4.51	-2.28	0.042
Curso, γ_{30}	35.75	3.98	8.97	<0.001	35.75	3.54	10.11	<0.001
Repetición, γ_{40}	-34.64	5.12	-6.76	<0.001	-34.64	4.41	-7.85	<0.001
ESCS, γ_{50}	9.8	1.12	8.76	<0.001	9.8	1.03	9.53	<0.001
Cambio de centro (1), γ_{60}	-10.82	1.86	-5.81	<0.001	-10.82	1.47	-7.35	<0.001
Cambio de centro (2+), γ_{70}	-15.73	3.28	-4.89	<0.001	-15.73	2.75	-5.71	<0.001

Una primera exploración de los resultados de ambos modelos no arroja diferencias aparentes, pero con el objetivo de tomar una decisión informada en cuanto al tipo de modelo a utilizar para el análisis, se realizó también una comparación de los residuos de los centros estimados por ambos modelos. Se realizó una prueba *t* para grupos relacionados, cuyos resultados indicaron una correlación de .944 y una ausencia de diferencias significativas entre ambos grupos de residuos (tabla 6.6).

Tabla 6.6. Diferencias entre los residuos de los centros en los modelos de pendientes fijas y aleatorias para matemáticas.

Par	Res. aleatorio – Res. fijo	Diferencias				<i>t</i>	gl	Sig.	
		Media	Desv. típica	Error típico medio	Intervalo de confianza 95% de la diferencia				
					Bajo				Alto
		-,109	4,074	,136	-,376	,158	-,801	895 ,423	

Teniendo en cuenta la ausencia de diferencias aparentes entre las estimaciones ordinarias y robustas de los errores en el modelo de pendientes fijas, lo cual es indicativo de una correcta especificación del modelo (Bell et al., 2018), así como la ausencia de diferencias significativas entre los residuos de los centros producidos por los modelos de

pendientes fijas y aleatorias, el modelo utilizado será el de pendientes fijas (tabla 6.7), atendiendo al principio de parsimonia.

Tabla 6.7. Modelo final para matemáticas, estimación robusta de los errores.

Efecto fijo	Coefficiente	Error típico	t-ratio	g.l.	p-valor
Intercepto, γ_{00}	543.069423	1.732621	313.438	89	<0.001
ESCS centro, γ_{01}	13.000162	1.906068	6.820	99	<0.001
Género, γ_{10}	-22.578047	1.967796	-11.474	26	<0.001
Inmigrante 1ª gen, γ_{20}	-11.806617	4.851714	-2.433	18	0.026
Curso, γ_{30}	34.859660	3.852886	9.048	27	<0.001
Repetición, γ_{40}	-35.803964	4.943071	-7.243	34	<0.001
ESCS, γ_{50}	9.787405	1.121277	8.729	19	<0.001
Cambio de centro (1), γ_{60}	-10.649863	1.860594	-5.724	115	<0.001
Cambio de centro (2+), γ_{70}	-15.602348	3.302820	-4.724	34	<0.001

Atendiendo a las puntuaciones t de las covariables, se puede observar cómo la variable que más incidencia tiene en la competencia matemática de los estudiantes es su género, diferencia que en este caso favorece a los estudiantes varones, que puntúan de media 22.6 puntos por encima de sus compañeras. Las siguientes variables en importancia son el curso y el ESCS individual de los estudiantes, ambos con efecto positivo en el rendimiento, y la repetición de curso, en este caso con efecto negativo. Cabe destacar que, si bien la condición de ser inmigrante de 1ª generación resulta significativa en la competencia matemática de los estudiantes, los inmigrantes de 2ª generación no presentan diferencias estadísticamente significativas con el rendimiento de sus compañeros nativos.

La estimación de los componentes de varianza del modelo final (tabla 6.8) permitirá el cálculo del CCI una vez se han controlado todas las variables significativas, por lo que servirá para establecer la magnitud de la explicación de la varianza por parte del modelo.

Tabla 6.8. Componentes de varianza del modelo condicional para matemáticas.

Efecto aleatorio	Desviación Típica	Componente de varianza	g.l.	χ^2	p-valor
Nivel 2	14.13213	199.71716	894	2252.79160	<0.001
Nivel 1	64.82719	4202.56440			

En función de estos datos, el CCI final tiene un valor de 0.045, es decir, que queda un 4.5% de varianza no explicada al nivel de centro. Teniendo en cuenta que el CCI inicial era de 13%, se puede afirmar que el modelo establecido explica un 8.5% de la varianza entre centros.

El último paso para establecer el modelo es la comprobación del cumplimiento de los supuestos. El supuesto de dependencia de las observaciones ya se comprobó al inicio del proceso a través del índice CCI, por lo que restaría comprobar los supuestos relacionados con los residuos del modelo.

En relación a los residuos de nivel 1, presentan una media cercana a 0 (tabla 6.9), tomando un valor de 2 en una escala de aproximadamente 600 puntos de amplitud, y una desviación típica de 64.21.

Tabla 6.9. Estadísticos descriptivos de los residuos de nivel 1 para matemáticas.

	N	Mínimo	Máximo	Media	Desviación típica	Asimetría		Curtosis	
						Valor	Error típico	Valor	Error típico
Residuo N1	28299	-291,67	268,76	2,11	64,21	-,07	,01	,01	,03

La distribución del residuo es normal (figura 6.1), siendo el nivel de significación de la prueba de Kolmogorov-Smirnov de 0.63.

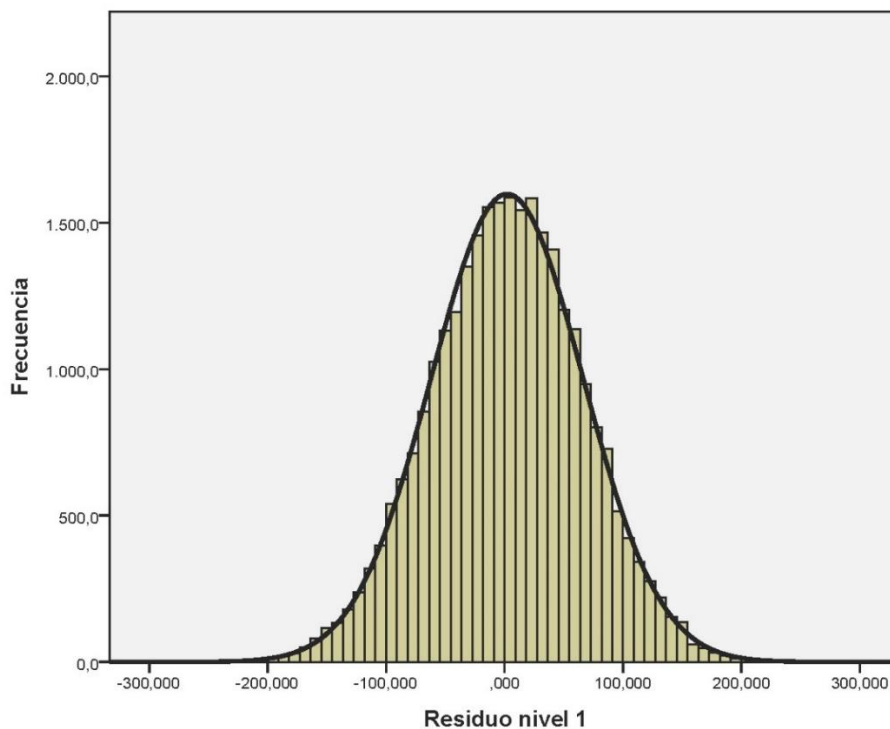


Figura 6.1. Distribución de los residuos de nivel 1 de matemáticas.

El supuesto de la homocedasticidad de los residuos del primer nivel, comprobado a través de su dispersión en función de los niveles del valor estimado por el modelo para el primer nivel, también se cumple (figura 6.2), presentando un R^2 de 5.530E-4.

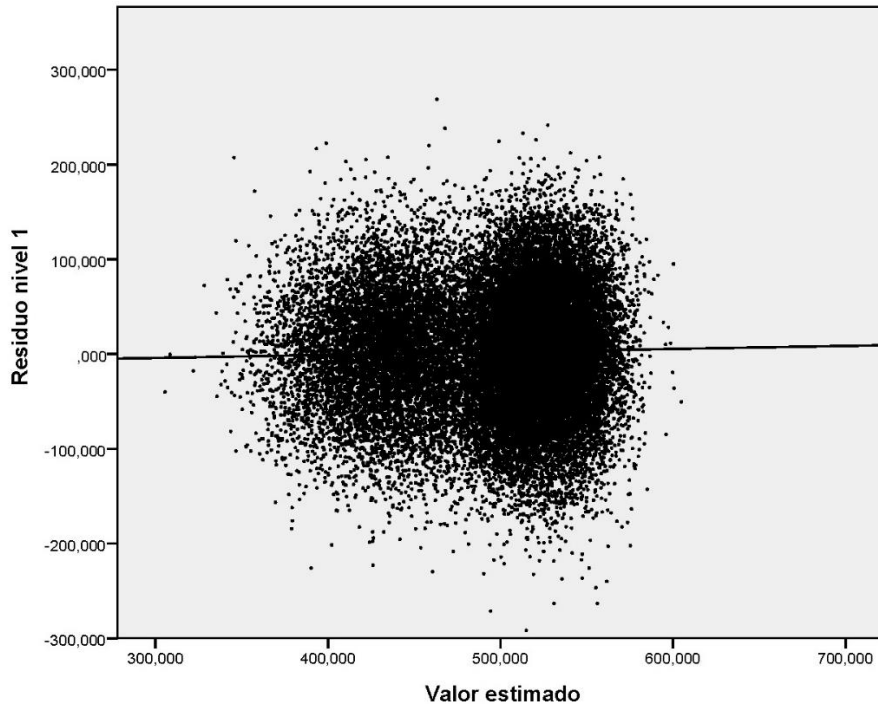


Figura 6.2. Comprobación del supuesto de homocedasticidad para los residuos de nivel 1 en matemáticas.

Los residuos de segundo nivel también presentan una media cercana a 0, y una desviación típica de 10.64 (tabla 6.10).

Tabla 6.10. Estadísticos descriptivos de los residuos de nivel 2 para matemáticas.

	N	Mínimo	Máximo	Media	Desviación típica	Asimetría		Curtosis	
						Valor	Error típico	Valor	Error típico
Residuo N2	896	-49,08	45,12	,64	10,64	-,44	,08	2,51	,16

La distribución de estos residuos se desvía ligeramente de la normalidad (nivel de significación del índice Kolmogorov-Smirnov de .001), como se puede apreciar en la figura 6.3, debido a una concentración alta de valores en torno a la media (distribución leptocúrtica) y una asimetría negativa.

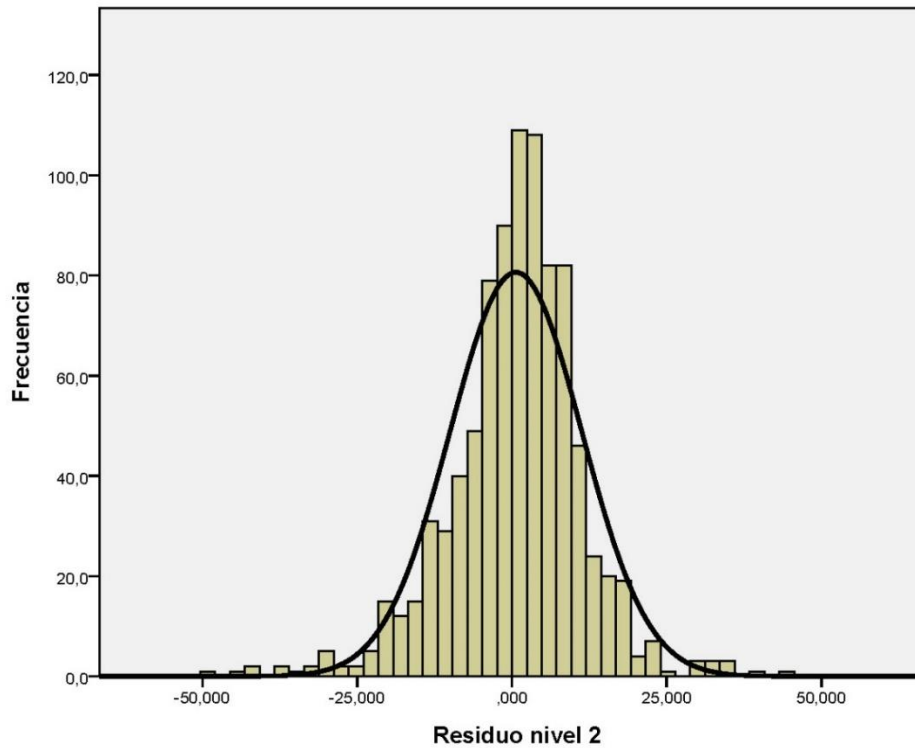


Figura 6.3. Distribución de los residuos de nivel 2 de matemáticas.

El gráfico Q-Q de la distribución (figura 6.4) muestra que la mayoría de los valores se ajustan a la curva normal, habiendo algunos *outliers* en los extremos de la distribución que pueden estar produciendo la desviación observada.

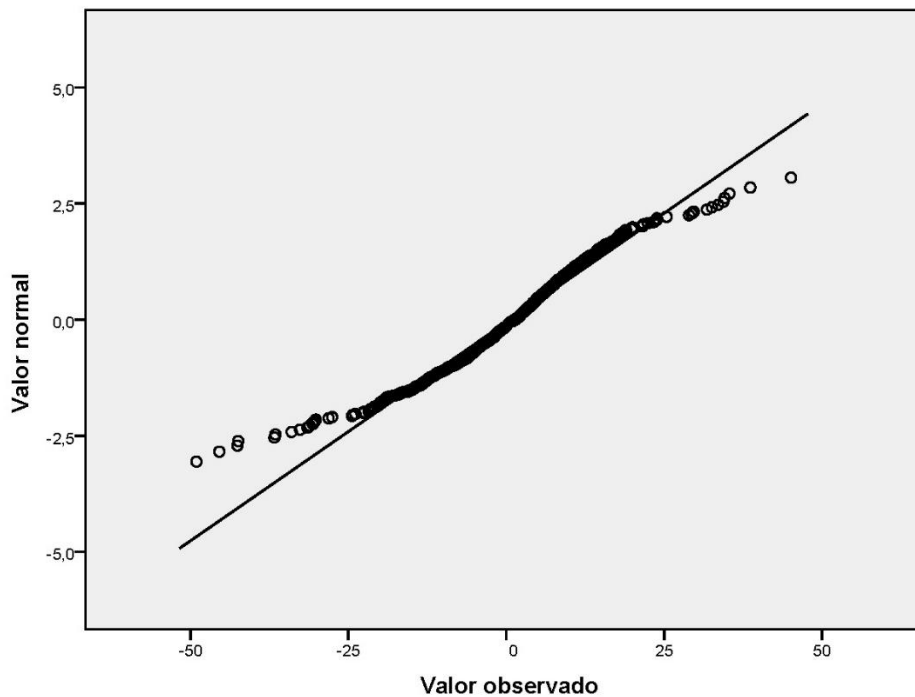


Figura 6.4. Gráfico Q-Q para la distribución de los residuos de nivel 2 de matemáticas.

El supuesto de homocedasticidad, comprobado a través de la figura 6.5, se cumple, siendo el R2 igual a 0.002.

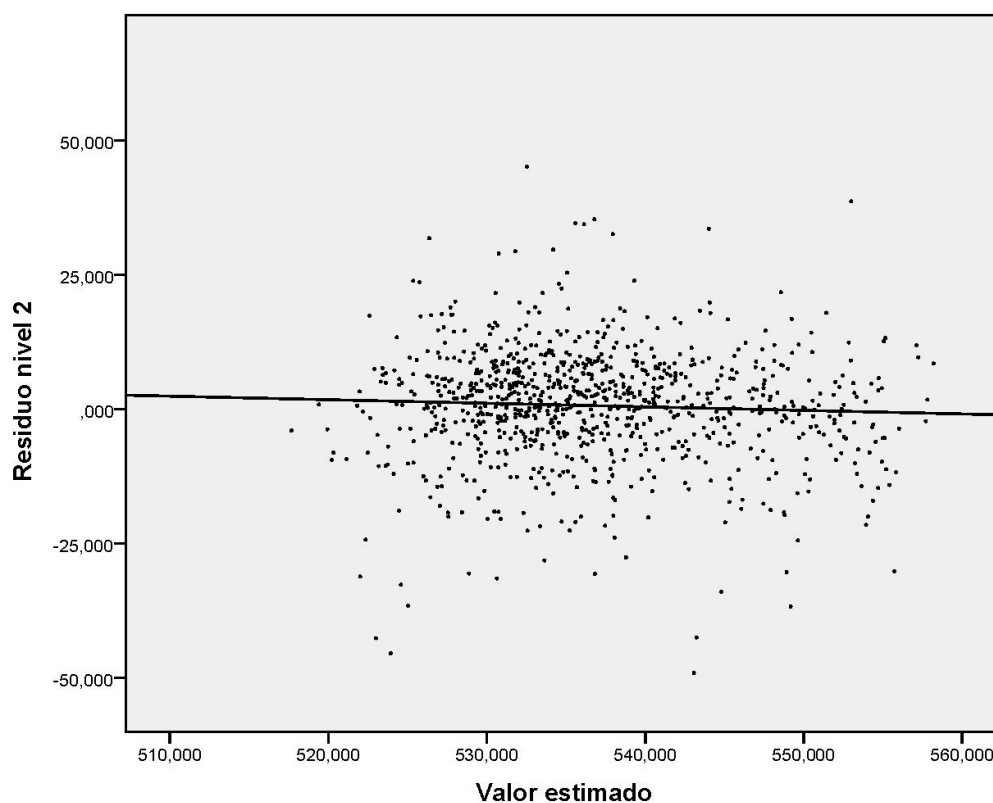


Figura 6.5. Comprobación del supuesto de homocedasticidad para los residuos de nivel 2 de matemáticas.

Ya que todos los supuestos relacionados con los residuos (distribución normal, media 0, homocedasticidad) parecen cumplirse de manera aceptable, el modelo presentado en la tabla 6.7 queda determinado como el modelo final para la competencia matemática.

6.1.3. Competencia lectora

En primer lugar, se presenta el modelo nulo o incondicional para las puntuaciones en la competencia lectora, expresado a través de la ecuación 6.5.

$$Lectura_{ij} = \gamma_{00} + u_{0j} + r_{ij} \quad (6.5)$$

Donde:

$Lectura_{ij}$ es la puntuación media del centro en lectura;

γ_{00} es la media de todos los centros en lectura;

u_{0j} es la diferencia entre la puntuación del centro en lectura y la media general de todos los centros; y

r_{ij} es el residuo del nivel 1.

A partir de este modelo se calculan los componentes de varianza de ambos niveles de análisis (tabla 6.11) para poder estudiar el Coeficiente de Correlación Intraclase.

Tabla 6.11. Componentes de varianza del modelo nulo para lectura.

Efecto aleatorio	Desviación Típica	Componente de varianza	g.l.	χ^2	p-valor
Nivel 2	31.7	1004.96	895	5551.58	<0.001
Nivel 1	82.57	6818.6			

Así, a partir de los datos de la variabilidad presentada por los datos en el nivel de estudiantes y de centros, puede procederse a calcular el CCI (ecuación 6.6).

$$CCI_{Lectura} = \frac{\tau_{00}}{\tau_{00} + \sigma^2} = \frac{1004.96}{(1004.96 + 6818.6)} = 0.1285 \quad (6.6)$$

En este caso, la variabilidad atribuible al nivel centro para la puntuación en comprensión lectora de los estudiantes españoles es de 12.85%.

A continuación, se presentan los modelos de pendientes fijas y aleatorias con el cálculo de ambos tipos de error (ordinario y robusto) para compararlos y poder decidir cuál se ajusta mejor al análisis. En primer lugar, tras retirar una a una las variables que no presentan una relación significativa con la variable criterio, el modelo de intercepto aleatorio y pendientes fijas (ecuación 6.7) se conforma de la siguiente manera:

$$\begin{aligned} Lectura_{ij} = & \gamma_{00} + \gamma_{01} * ESCS \text{ nivel } 2_j + \gamma_{02} * \text{Porcentaje de repetidores}_j \\ & + \gamma_{03} * \text{Porcentaje de alumnas}_j + \gamma_{10} * \text{Género}_{ij} + \gamma_{20} \\ & * \text{Curso}_{ij} + \gamma_{30} * \text{Repetición}_{ij} + \gamma_{40} * ESCS_{ij} + \gamma_{50} \\ & * \text{Idioma hogar}_{ij} + \gamma_{60} * \text{Cambio centro } 1_{ij} + \gamma_{70} \\ & * \text{Cambio centro } 2_{ij} + u_{0j} + r_{ij} \end{aligned} \quad (6.7)$$

Donde:

$lectura_{ij}$ es la puntuación media del centro en lectura;

γ_{00} es la media de todos los centros en lectura;

γ_{01} - γ_{03} son las covariables de nivel 2;

γ_{10} - γ_{70} son las covariables de nivel 1;

u_{0j} es la diferencia entre la puntuación del centro en lectura y la media general de todos los centros; y

r_{ij} es el residuo del nivel 1.

Como puede observarse en la tabla 6.12, los errores, valores t y niveles de significación entre la estimación ordinaria y robusta de los errores no difieren mucho entre sí.

Tabla 6.12. Modelo de intercepto aleatorio y pendientes fijas para lectura.

Efecto fijo	Estimación robusta del error				Estimación ordinaria del error			
	Coef.	Error típico	t-ratio	p-valor	Coef.	Error típico	t-ratio	p-valor
Intercepto, γ_{00}	529.43	7.53	70.34	<0.001	529.43	5.02	105.46	<0.001
ESCS centro, γ_{01}	20.34	2.69	7.55	<0.001	20.34	1.96	10.39	<0.001
Porcentaje repetidores, γ_{02}	36.83	10.91	3.38	<0.001	36.83	7.54	4.88	<0.001
Porcentaje alumnas, γ_{03}	23.93	11.06	2.16	0.031	23.93	7.18	3.34	<0.001
Género, γ_{10}	6.98	1.79	3.90	<0.001	6.98	1.43	4.90	<0.001
Curso, γ_{20}	42.62	3.41	12.51	<0.001	42.62	2.57	16.60	<0.001
Repetición, γ_{30}	-30.72	5.05	-6.09	<0.001	-30.72	4.00	-7.69	<0.001
ESCS, γ_{40}	8.22	0.95	8.68	<0.001	8.22	0.77	10.68	<0.001
Idioma hogar, γ_{50}	-7.10	3.05	-2.33	0.026	-7.10	2.49	-2.85	0.012
Cambio de centro (1), γ_{60}	-8.89	2.17	-4.11	<0.001	-8.89	1.66	-5.37	<0.001
Cambio de centro (2+), γ_{70}	-20.88	3.19	-6.53	<0.001	-20.88	2.45	-8.50	<0.001

Por su parte, el modelo de intercepto aleatorio y pendientes aleatorias para lectura se presenta en la ecuación 6.8.

(6.8)

$$\begin{aligned}
 Lectura_{ij} = & \gamma_{00} + \gamma_{01} * ESCS\ nivel\ 2_j + \gamma_{02} * Porcentaje\ de\ repetidores_j + \gamma_{03} \\
 & * Porcentaje\ de\ alumnas_j + \gamma_{10} * Género_{ij} + \gamma_{20} * Curso_{ij} + \gamma_{30} \\
 & * Repetición_{ij} + \gamma_{40} * ESCS_{ij} + \gamma_{50} * Cambio\ centro\ 1_{ij} + \gamma_{60} \\
 & * Cambio\ centro\ 2_{ij} + u_{0j} + u_{1j} * Género_{ij} + u_{2j} * Curso_{ij} + u_{3j} \\
 & * Repetición_{ij} + u_{4j} * ESCS_{ij} + u_{5j} * Cambio\ centro_{ij} + u_{6j} \\
 & * Cambio\ centro\ 2_{ij} + r_{ij}
 \end{aligned}$$

Donde¹⁸:

$u_{1j}-u_{6j}$ es la diferencia entre el coeficiente de la pendiente variable del centro en la covariable de nivel 1 para cada centro y su efecto fijo (componente aleatorio de las pendientes).

Cabe destacar que, en el caso de la lectura, al modelar las pendientes como aleatorias el modelo (6.13) ha perdido una covariable con respecto al de efectos fijos, que es la variable relacionada con la diferencia entre el idioma hablado en el hogar y el idioma en el que se ha hecho la prueba.

Tabla 6.13. Modelo de intercepto y pendientes aleatorias para lectura.

Efecto fijo	Estimación robusta del error				Estimación ordinaria del error			
	Coef.	Error típico	t-ratio	p-valor	Coef.	Error típico	t-ratio	p-valor
Intercepto, γ_{00}	520.45	5.87	88.72	<0.001	520.45	4.06	128.04	<0.001
ESCS centro, γ_{01}	20.05	2.50	8.02	<0.001	20.05	1.74	11.56	<0.001
Porcentaje repetidores, γ_{02}	36.57	10.36	3.53	<0.001	36.57	6.99	5.24	<0.001
Porcentaje alumnas, γ_{03}	25.17	10.34	2.44	0.015	25.17	7.38	3.41	<0.001
Género, γ_{10}	7.17	1.82	3.95	<0.001	7.17	1.51	4.76	<0.001
Curso, γ_{20}	43.76	3.53	12.40	<0.001	43.76	2.91	15.06	<0.001
Repetición, γ_{30}	-29.31	5.24	-5.59	<0.001	-29.31	4.43	-6.62	<0.001
ESCS, γ_{40}	8.40	0.95	8.88	<0.001	8.40	0.79	10.57	<0.001
Cambio de centro (1), γ_{60}	-9.26	2.18	-4.25	<0.001	-9.26	1.75	-5.30	<0.001
Cambio de centro (2+), γ_{70}	-22.06	3.13	-7.05	<0.001	-22.06	2.51	-8.80	<0.001

Para completar el análisis de la idoneidad de los modelos en función del efecto de sus pendientes, se realiza una prueba t para muestras relacionadas (tabla 6.14) con el objetivo de comprobar si hay diferencias estadísticamente significativas entre los residuos del nivel de centros generados por el modelo de pendientes fijas y el de pendientes aleatorias. La correlación entre estas dos variables es de .963.

¹⁸ Para la especificación del resto de elementos, ver ecuación 6.7.

Tabla 6.14. Diferencias entre los residuos de los centros en los modelos de pendientes fijas y aleatorias para lectura.

Par	Res. aleatorio – Res. fijo	Diferencias					t	gl	Sig.
		Media	Desv. típica	Error típico medio	Intervalo de confianza 95% de la diferencia				
					Bajo	Alto			
		-,03	3,58	,12	-,27	,20	-,27	895	,79

Al igual que en el caso de la competencia matemática, la ausencia de grandes diferencias entre los modelos de pendientes fijas y aleatorias, entre sus respectivos residuos de segundo nivel, y entre la estimación ordinaria y robusta del modelo de pendientes fijas, se establece el modelo de pendientes fijas como modelo final (tabla 6.15) para el rendimiento en comprensión lectora de los alumnos españoles.

Tabla 6.15. Modelo final para lectura, con estimación robusta de los errores.

Efecto fijo	Coef.	Error típico	t-ratio	p-valor
Intercepto, γ_{00}	529.43	7.53	70.34	<0.001
ESCS centro, γ_{01}	20.34	2.69	7.55	<0.001
Porcentaje repetidores, γ_{02}	36.83	10.91	3.38	<0.001
Porcentaje alumnas, γ_{03}	23.93	11.06	2.16	0.031
Género, γ_{10}	6.98	1.79	3.90	<0.001
Curso, γ_{20}	42.62	3.41	12.51	<0.001
Repetición, γ_{30}	-30.72	5.05	-6.09	<0.001
ESCS, γ_{40}	8.22	0.95	8.68	<0.001
Idioma hogar, γ_{50}	-7.10	3.05	-2.33	0.026
Cambio de centro (1), γ_{60}	-8.89	2.17	-4.11	<0.001
Cambio de centro (2+), γ_{70}	-20.88	3.19	-6.53	<0.001

En el modelo final para el rendimiento en comprensión lectora, la variable que más influencia demuestra es el curso, seguido por el ESCS individual y el ESCS medio del centro, todos con un impacto positivo en el rendimiento. La repetición y los cambios de centro (especialmente si han sido dos o más) son las variables con impactos negativos más altos en la comprensión lectora. Cabe destacar el rol del género en este modelo ya que, aunque su impacto es menor que en el modelo de matemáticas, muestra un signo opuesto, es decir, a favor de las alumnas, quienes puntuarían unos 7 puntos de media más que sus compañeros.

La estimación de los componentes de varianza del modelo final (tabla 6.16) es necesaria para poder calcular en qué medida el modelo ha controlado la variabilidad de las observaciones a nivel de centro educativo.

Tabla 6.16. Componentes de varianza del modelo condicional para lectura.

Efecto aleatorio	Desviación Típica	Componente de varianza	g.l.	χ^2	p-valor
Nivel 2	15.99558	255.85856	892	2432.18427	<0.001
Nivel 1	69.06516	4769.99664			

En función de estos datos, el CCI final tiene un valor de 0.051, es decir, que queda un 5.1% de varianza no explicada al nivel de centro. Teniendo en cuenta que el CCI inicial era de 12.85%, se puede afirmar que el modelo establecido explica un 7.75% de la varianza entre centros.

El último paso para establecer el modelo es la comprobación del cumplimiento de los supuestos. El supuesto de dependencia de las observaciones ya se comprobó al inicio del proceso a través del índice CCI, por lo que restaría comprobar los supuestos relacionados con los residuos del modelo.

Los residuos de nivel 1 presentan una media cercana a 0 (tabla 6.17), tomando un valor de -1.5 en una escala de aproximadamente 600 puntos de amplitud, y una desviación típica de 68.7.

Tabla 6.17. Estadísticos descriptivos de los residuos de nivel 1 para lectura.

	N	Mínimo	Máximo	Media	Desviación típica	Asimetría		Curtosis	
						Valor	Error típico	Valor	Error típico
Residuo N1	28685	-301,16	278,29	-1,51	68,70	-,14	,01	,05	,03

Como puede observarse en la figura 6.6, la distribución de los residuos de nivel 1 se asemeja a la distribución normal, aunque presente una muy ligera desviación fruto de una distribución asimétrica negativa (nivel de significación del índice Kolmogorov-Smirnov de .000).

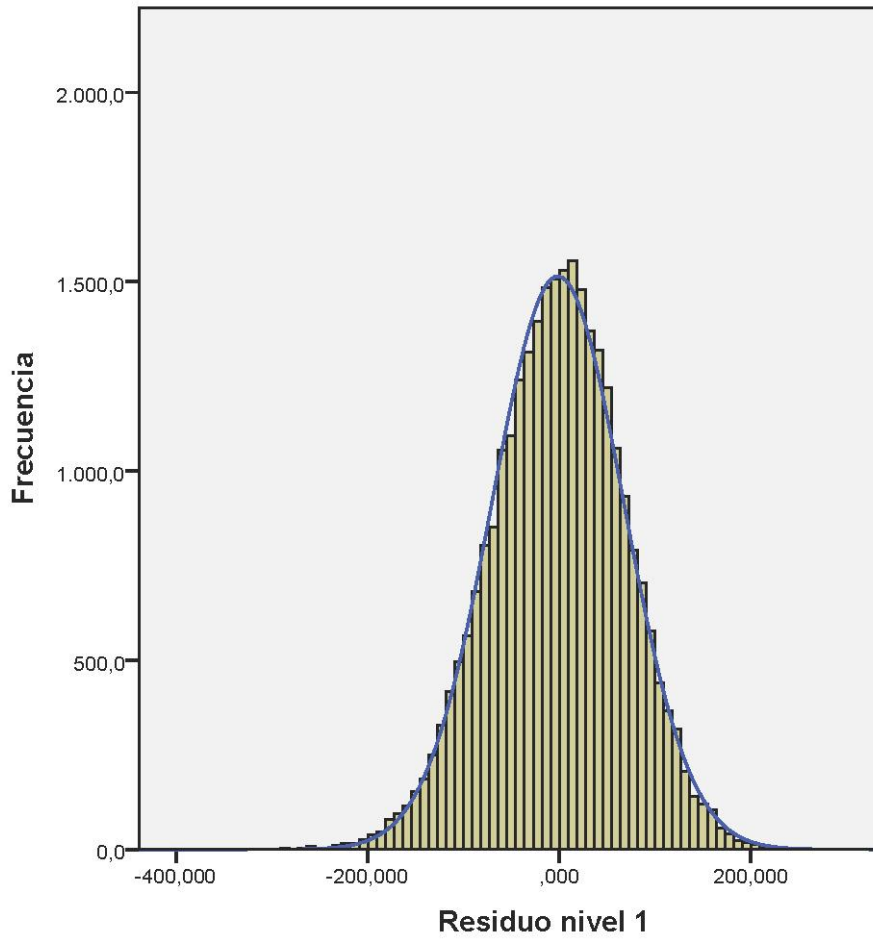


Figura 6.6. Distribución de los residuos de nivel 1 de lectura.

El gráfico Q-Q muestra un ajuste general a la distribución normal, con la excepción de algunos valores extremos que se alejan de esta (figura 6.7).

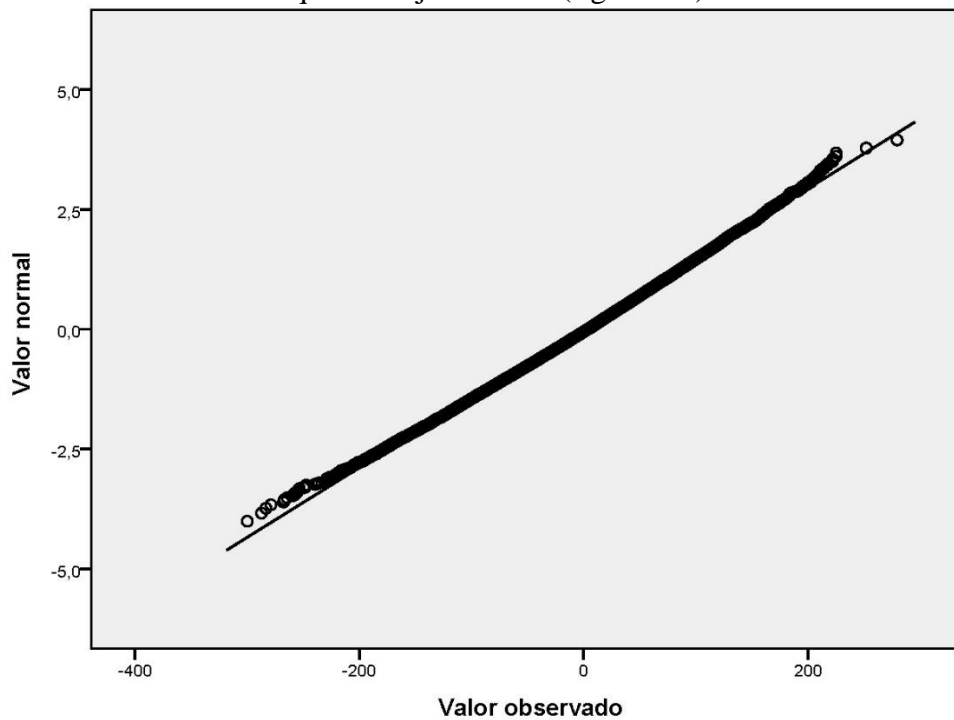


Figura 6.7. Gráfico Q-Q para los residuos de nivel 1 de lectura.

En cuanto al supuesto de homocedasticidad para los residuos de nivel 1, la figura 6.8 muestra que el supuesto se cumple en este caso, presentando un R^2 de $2.820E-4$.

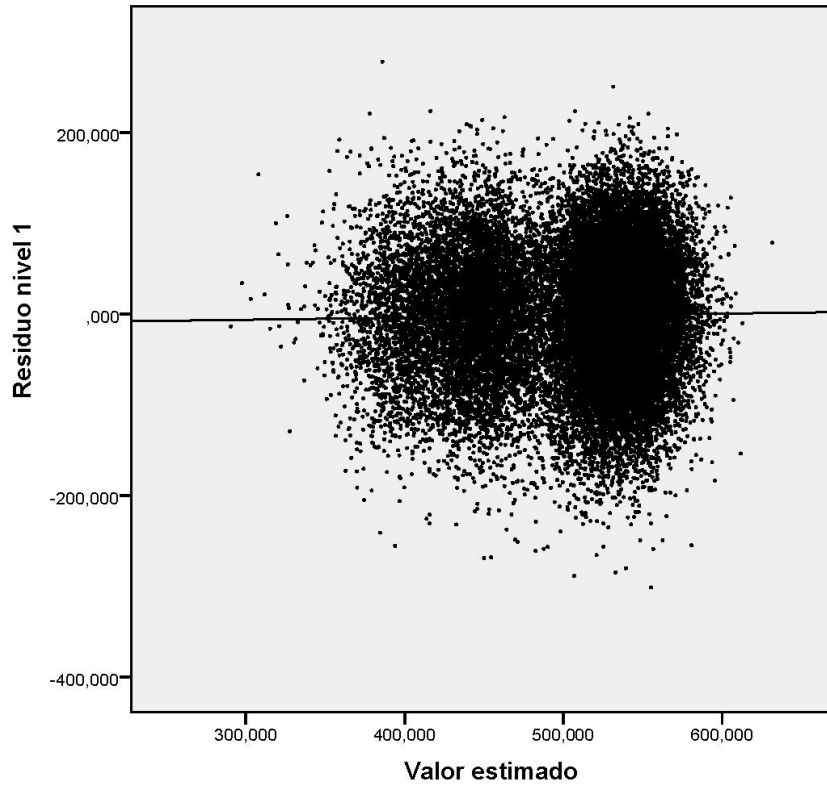


Figura 6.8. Verificación del supuesto de homocedasticidad para los residuos de nivel 1 de lectura.

Con respecto a los residuos de nivel 2, presentan una media cercana a 0 (tabla 6.17), y una desviación típica de 12.63.

Tabla 6.18. Estadísticos descriptivos de los residuos de nivel 2 para lectura.

	N	Mínimo	Máximo	Media	Desviación típica	Asimetría		Curtosis	
						Valor	Error típico	Valor	Error típico
Residuo N2	896	-66,85	50,43	-,35	12,63	,02	,08	2,07	,16

La distribución de estos residuos (figura 6.9) se ajusta razonablemente a la distribución normal, presentando una ligera desviación debida a una asimetría positiva (nivel de significación del índice de Kolmogorov-Smirnov de .003).

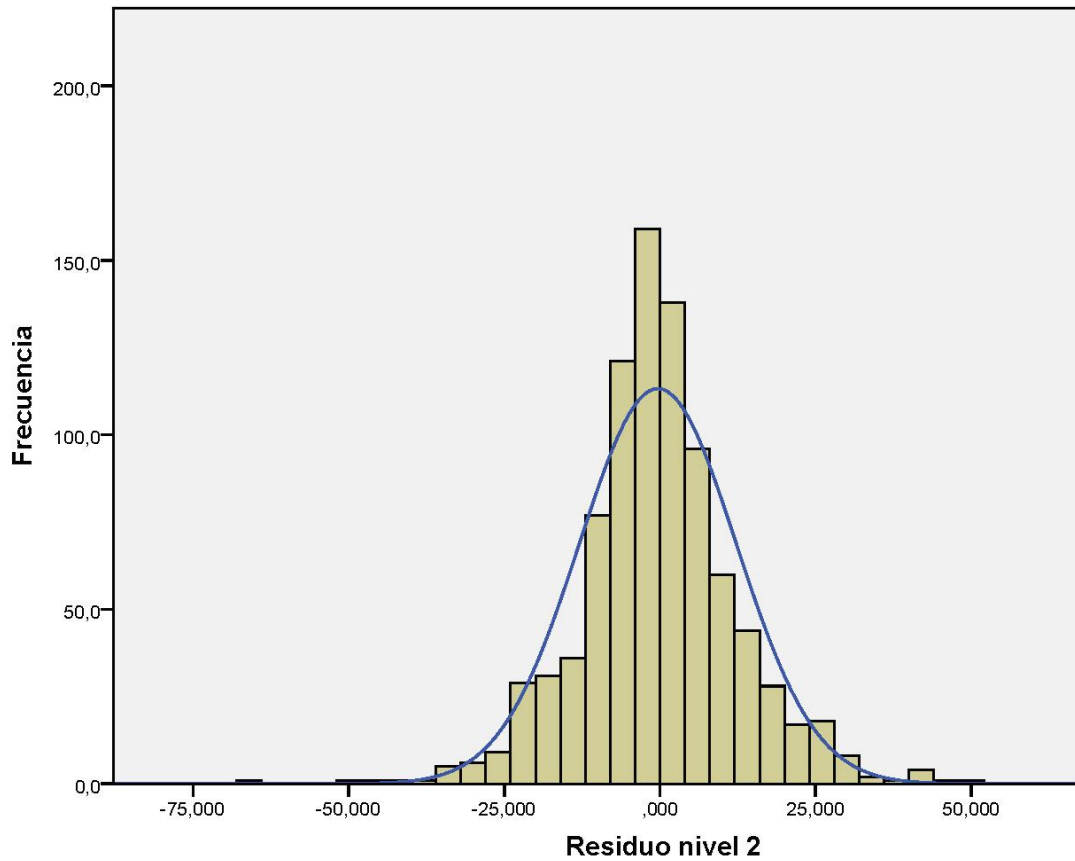


Figura 6.9. Distribución de los residuos de nivel 2 de lectura.

La exploración de esta desviación a través del gráfico Q-Q para la distribución (figura 6.10) revela que, al margen de algunos *outliers* en los extremos de la distribución, la gran mayoría de valores centrales se muestran muy cercanos a la distribución normal.

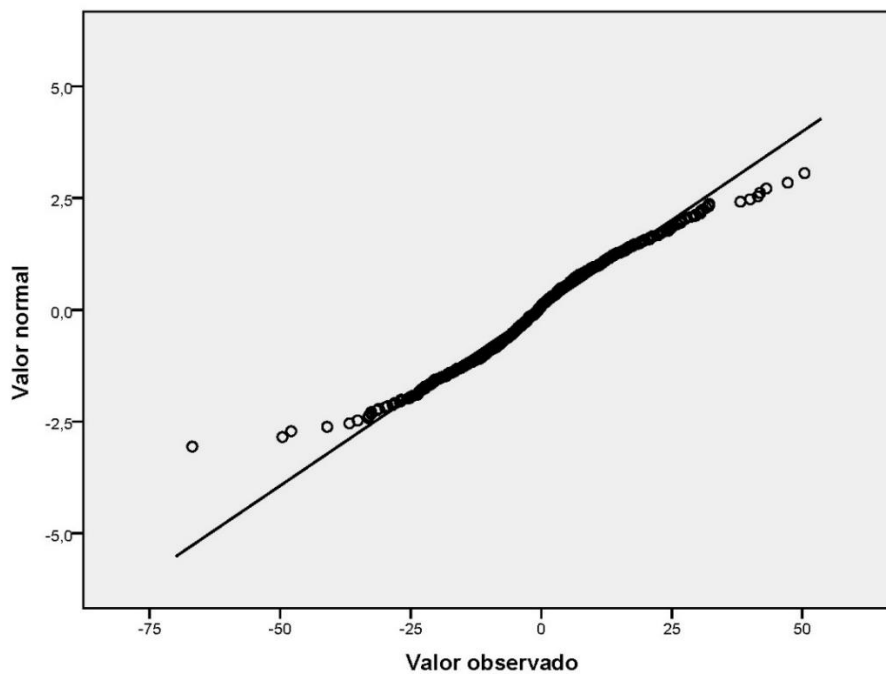


Figura 6.10. Gráfico Q-Q de los residuos de nivel 2 de lectura.

La figura 6.11, que informa sobre la homocedasticidad de los residuos de nivel 2 para la competencia lectora, indica que se cumple dicho supuesto, presentando un R^2 de 0.001.

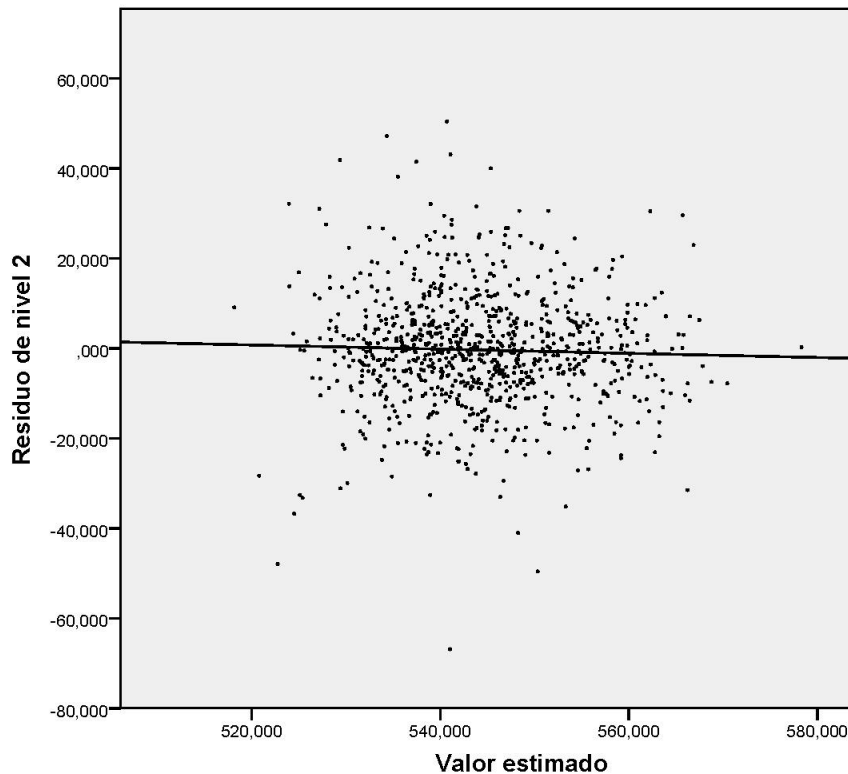


Figura 6.11. Verificación del supuesto de homocedasticidad para los residuos de nivel 2 de matemáticas.

Ya que todos los supuestos relacionados con los residuos (distribución normal, media 0, homocedasticidad) se cumplen razonablemente, el modelo presentado en la tabla 6.15 queda determinado como el modelo final para la competencia en comprensión lectora.

6.1.4. Competencia científica

En primer lugar, se presenta el modelo nulo o incondicional para las puntuaciones en la competencia científica, expresado a través de la ecuación 6.9.

$$Ciencia_{ij} = \gamma_{00} + u_{0j} + r_{ij} \quad (6.9)$$

Donde:

$Ciencia_{ij}$ es la puntuación media del centro en ciencias;

γ_{00} es la media de todos los centros en ciencia;

u_{0j} es la diferencia entre la puntuación del centro en ciencia y la media general de todos los centros; y

r_{ij} es el residuo del nivel 1.

A partir de este modelo se calculan los componentes de varianza de ambos niveles de análisis (tabla 6.19) para poder estudiar el Coeficiente de Correlación Intraclase.

Tabla 6.19. Componentes de varianza del modelo nulo para ciencia.

Efecto aleatorio	Desviación Típica	Componente de varianza	g.l.	χ^2	p-valor
Nivel 2	31.98	1022.6	895	5723.87	<0.001
Nivel 1	81.87	6703.38			

Así, a partir de los datos de la variabilidad presentada por los datos en el nivel de estudiantes y de centros, puede procederse a calcular el CCI (ecuación 6.9).

$$CCI_{Ciencia} = \frac{\tau_{00}}{\tau_{00} + \sigma^2} = \frac{1022.6}{(1022.6 + 6703.38)} = 0.1324 \quad (6.9)$$

En este caso, la variabilidad atribuible al nivel centro para la puntuación en comprensión lectora de los estudiantes españoles es de 13.24%.

Para determinar qué tipo de efectos de pendientes (fijas o aleatorias) es más adecuado utilizar, se calculan ambos modelos para poder compararlos. En primer lugar, se presenta el modelo de intercepto aleatorio y pendientes fijas (ecuación 6.10)

$$\begin{aligned}
 Ciencia_{ij} = & \gamma_{00} + \gamma_{01} * Tama\~{n}o\ centro_j + \gamma_{02} * Escasez\ personal_j + \gamma_{03} \\
 & * ESCS\ nivel\ 2_j + \gamma_{04} * Porcentaje\ de\ repetidores_j + \gamma_{05} \\
 & * Porcentaje\ de\ alumnas_j + \gamma_{10} * Mes\ nacimiento_{ij} + \gamma_{20} \\
 & * G\~{e}nero_{ij} + \gamma_{30} * Inmigrante\ 1^a\ generaci\~{o}n_{ij} + \gamma_{40} * Curso_{ij} + \gamma_{50} \\
 & * Repetic_i\~{o}n_{ij} + \gamma_{60} * ESCS_{ij} + \gamma_{70} * Cambio\ centro\ 1_{ij} + \gamma_{80} \\
 & * Cambio\ centro\ 2_{ij} + u_{0j} + r_{ij}
 \end{aligned}$$

Donde:

$Ciencia_{ij}$ es la puntuación media del centro en ciencia;

γ_{00} es la media de todos los centros en ciencia;

γ_{01} - γ_{05} son las covariables de nivel 2;

γ_{10} - γ_{80} son las covariables de nivel 1;

u_{0j} es la diferencia entre la puntuación del centro en ciencia y la media general de todos los centros; y

r_{ij} es el residuo del nivel 1.

Tras la retirada una a una de aquellas variables que no presentan una relación significativa con la variable criterio, el modelo queda conformado por las variables indicadas en la tabla 6.20.

Tabla 6.20. Modelo de intercepto aleatorio y pendientes fijas para ciencia.

Efecto fijo	Estimación robusta del error				Estimación ordinaria del error			
	Coef.	Error típico	t-ratio	p-valor	Coef.	Error típico	t-ratio	p-valor
Intercepto, γ_{00}	537.68	7.29	73.71	<0.001	537.68	4.36	123.19	<0.001
Tamaño centro, γ_{01}	-0.01	0.00	-2.24	0.025	-0.01	0.00	-2.82	0.005
Escasez personal, γ_{02}	2.49	1.16	2.14	0.032	2.49	0.79	3.16	0.002
ESCS centro, γ_{03}	20.75	2.61	7.96	<0.001	20.75	1.84	11.29	<0.001
Porcentaje repetidores, γ_{04}	31.95	9.53	3.35	<0.001	31.95	6.65	4.80	<0.001
Porcentaje alumnas, γ_{05}	31.84	11.98	2.66	0.008	31.84	7.59	4.20	<0.001
Mes nacimiento, γ_{10}	-0.59	0.23	-2.56	0.011	-0.59	0.17	-3.54	0.001
Género, γ_{20}	-19.76	1.69	-11.70	<0.001	-19.76	1.30	-15.21	<0.001
Inmigrante 1ª generación, γ_{30}	-9.63	2.89	-3.34	<0.001	-9.63	1.79	-5.37	<0.001
Curso, γ_{40}	38.29	2.89	13.25	<0.001	38.29	2.13	18.00	<0.001
Repetición, γ_{50}	-34.29	3.87	-8.86	<0.001	-34.29	2.74	-12.50	<0.001
ESCS, γ_{60}	9.78	0.70	13.99	<0.001	9.78	0.48	20.18	<0.001
Cambio de centro (1), γ_{70}	-10.36	1.94	-5.34	<0.001	-10.36	1.40	-7.41	<0.001
Cambio de centro (2+), γ_{80}	-16.96	2.64	-6.43	<0.001	-16.96	1.78	-9.54	<0.001

Por su parte, el modelo de intercepto aleatorio y pendientes aleatorias para lectura se presenta en la ecuación 6.11.

$$(6.11)$$

$$\begin{aligned}
Ciencia_{ij} = & \gamma_{00} + \gamma_{01} * \text{Tamaño centro}_j + \gamma_{02} * \text{Escasez personal}_j + \gamma_{03} \\
& * \text{ESCS nivel } 2_j + \gamma_{04} * \text{Porcentaje de repetidores}_j + \gamma_{05} \\
& * \text{Porcentaje de alumnas}_j + \gamma_{10} * \text{Mes nacimiento}_{ij} + \gamma_{20} \\
& * \text{Género}_{ij} + \gamma_{30} * \text{Inmigrante } 1^{\text{a}} \text{ generación}_{ij} + \gamma_{40} * \text{Curso}_{ij} + \gamma_{50} \\
& * \text{Repetición}_{ij} + \gamma_{60} * \text{ESCS}_{ij} + \gamma_{70} * \text{Cambio centro } 1_{ij} + \gamma_{80} \\
& * \text{Cambio centro } 2_{ij} + u_{0j} + u_{1j} * \text{Mes nacimiento}_{ij} + u_{2j} \\
& * \text{Género}_{ij} + u_{3j} * \text{inmigrante } 1^{\text{a}} \text{ generación}_{ij} + u_{4j} * \text{Curso}_{ij} \\
& + u_{5j} * \text{Repetición}_{ij} + u_{6j} * \text{ESCS}_{ij} + u_{7j} * \text{Cambio centro } 1_{ij} + u_{8j} \\
& * \text{Cambio centro } 2_{ij} + r_{ij}
\end{aligned}$$

Donde¹⁹:

$u_{1j}-u_{8j}$ es la diferencia entre el coeficiente de la pendiente variable del centro en la covariable de nivel 1 para cada centro y su efecto fijo (componente aleatorio de las pendientes).

Tabla 6.21. Modelo de intercepto y pendientes aleatorias para ciencias.

Efecto fijo	Estimación robusta del error				Estimación ordinaria del error			
	Coef.	Error típico	t-ratio	p-valor	Coef.	Error típico	t-ratio	p-valor
Intercepto, γ_{00}	538.64	7.01	76.82	<0.001	538.64	4.57	117.96	<0.001
Tamaño centro, γ_{01}	-0.01	0.00	-2.26	0.024	-0.01	0.00	-2.75	0.006
Escasez personal, γ_{02}	2.59	1.08	2.39	0.017	2.59	0.79	3.27	0.001
ESCS centro, γ_{03}	20.06	2.61	7.68	<0.001	20.06	1.88	10.68	<0.001
Porcentaje repetidores, γ_{04}	29.43	9.05	3.25	0.001	29.43	6.53	4.50	<0.001
Porcentaje alumnas, γ_{05}	30.79	11.44	2.69	0.007	30.79	7.89	3.90	<0.001
Mes nacimiento, γ_{10}	-0.62	0.23	-2.73	0.007	-0.62	0.17	-3.61	<0.001
Género, γ_{20}	-19.45	1.70	-11.46	<0.001	-19.45	1.37	-14.24	<0.001
Inmigrante 1 ^a generación, γ_{30}	-8.08	2.78	-2.91	0.004	-8.08	2.00	-4.04	<0.001
Curso, γ_{40}	38.85	2.90	13.38	<0.001	38.85	2.33	16.66	<0.001
Repetición, γ_{50}	-33.47	3.90	-8.59	<0.001	-33.47	2.98	-11.24	<0.001
ESCS, γ_{60}	9.87	0.69	14.31	<0.001	9.87	0.52	18.90	<0.001
Cambio de centro (1), γ_{70}	-10.34	1.91	-5.41	<0.001	-10.34	1.44	-7.19	<0.001
Cambio de centro (2+), γ_{80}	-17.28	2.59	-6.68	<0.001	-17.28	1.93	-8.93	<0.001

¹⁹ Para la especificación del resto de elementos, ver ecuación 6.10.

La tabla 6.21 presenta los coeficientes y estadísticos para las variables resultantes como significativas en el modelo de intercepto y pendientes aleatorias.

Para completar el análisis de la idoneidad de los modelos en función del efecto de sus pendientes, se realiza una prueba t para muestras relacionadas (tabla 6.22) entre los residuos del nivel de centros generados por el modelo de pendientes fijas y el de pendientes aleatorias. La correlación entre estas dos variables es de .965.

Tabla 6.22. Diferencias entre los residuos de los centros en los modelos de pendientes fijas y aleatorias para ciencias.

Par	Res. aleatorio – Res. fijo	Diferencias					t	gl	Sig.
		Media	Desv. típica	Error típico medio	Intervalo de confianza 95% de la diferencia				
					Bajo	Alto			
		-,11	4,31	,14	-,39	,18	-,74	-,11	4,31

Siguiendo el ejemplo de las dos competencias anteriormente analizadas, teniendo en cuenta la ausencia de grandes diferencias entre los modelos de pendientes fijas y aleatorias, entre sus respectivos residuos de segundo nivel, y entre la estimación ordinaria y robusta del modelo de pendientes fijas, se establece el modelo de pendientes fijas como modelo final (tabla 6.23) para el rendimiento en comprensión lectora.

Tabla 6.23. Modelo final para ciencia, con estimación robusta de los errores.

Efecto fijo	Coef.	Error típico	t-ratio	p-valor
Intercepto, γ_{00}	537.68	7.29	73.71	<0.001
Tamaño centro, γ_{01}	-0.01	0.00	-2.24	0.025
Escasez personal, γ_{02}	2.49	1.16	2.14	0.032
ESCS centro, γ_{03}	20.75	2.61	7.96	<0.001
Porcentaje repetidores, γ_{04}	31.95	9.53	3.35	<0.001
Porcentaje alumnas, γ_{05}	31.84	11.98	2.66	0.008
Mes nacimiento, γ_{10}	-0.59	0.23	-2.56	0.011
Género, γ_{20}	-19.76	1.69	-11.70	<0.001
Inmigrante 1ª generación, γ_{30}	-9.63	2.89	-3.34	<0.001
Curso, γ_{40}	38.29	2.89	13.25	<0.001
Repetición, γ_{50}	-34.29	3.87	-8.86	<0.001
ESCS, γ_{60}	9.78	0.70	13.99	<0.001
Cambio de centro (1), γ_{70}	-10.36	1.94	-5.34	<0.001
Cambio de centro (2+), γ_{80}	-16.96	2.64	-6.43	<0.001

Los resultados del modelo final para la competencia científica se asemejan a los del modelo explicativo de matemáticas en cuanto a las variables que tienen más impacto y el signo de dicha influencia. Las variables con un nivel más alto de incidencia en el

rendimiento en ciencias son el ESCS individual y el curso en el que se encuentra el estudiante. El género también juega un rol importante dentro del modelo, indicando que los estudiantes puntúan de media 20 puntos por encima de sus compañeras en esta competencia. Otras variables con relevancia dentro del modelo son la repetición de curso y los cambios de centro (especialmente dos o más cambios), con impacto negativo en el rendimiento, y el ESCS medio del centro, con influencia positiva. Al igual que en el modelo de matemáticas, el estatus migratorio de los estudiantes presenta un efecto negativo en el desempeño en la competencia científica, pero sólo para los inmigrantes de 1ª generación, que puntúan aproximadamente 10 puntos por debajo de sus compañeros.

La estimación de los componentes de varianza del modelo final (tabla 6.24) permitirá el cálculo del CCI una vez se han controlado todas las variables significativas, por lo que servirá para establecer la magnitud de la explicación de la varianza por parte del modelo.

Tabla 6.24. Componentes de varianza del modelo condicional para ciencia.

Efecto aleatorio	Desviación Típica	Componente de varianza	g.l.	χ^2	<i>p</i> -valor
Nivel 2	16.64029	276.89910	890	2575.66144	<0.001
Nivel 1	68.43336	4683.12443			

En función de estos datos, el CCI final tiene un valor de 0.056, es decir, que queda un 5.6% de varianza no explicada al nivel de centro. Teniendo en cuenta que el CCI inicial era de 13.24%, se puede afirmar que el modelo establecido explica un 7.64% de la varianza entre centros.

El último paso para establecer el modelo es la comprobación del cumplimiento de los supuestos. El supuesto de dependencia de las observaciones ya se comprobó al inicio del proceso a través del índice CCI, por lo que restaría comprobar los supuestos relacionados con los residuos del modelo.

En relación con los residuos de nivel 1, presentan una media de 0 (tabla 6.25), y una desviación típica de 68.07.

Tabla 6.25. Estadísticos descriptivos de los residuos de nivel 1 para ciencia.

	N	Mínimo	Máximo	Media	Desviación típica	Asimetría		Curtosis	
						Valor	Error típico	Valor	Error típico
Residuo N1	28299	-260,13	233,63	,00	68,07	-,09	,01	-,07	,03

Como puede observarse en la figura 6.12, la distribución de los residuos de nivel 1 presenta una distribución normal, siendo el nivel de significación de la prueba Kolmogorov-Smirnov de .183.

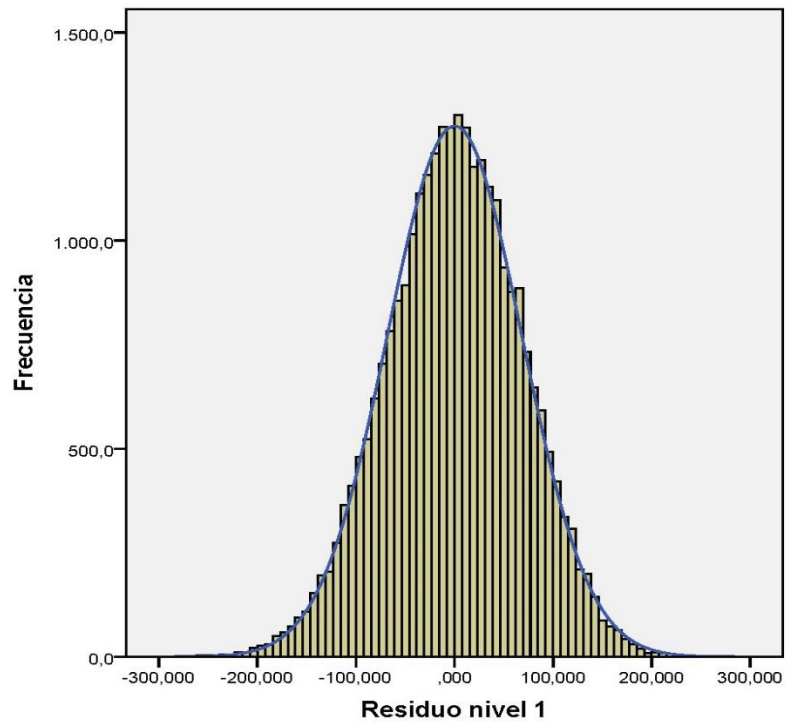


Figura 6.12. Distribución de los residuos de nivel 1 para ciencia.

El supuesto de homocedasticidad, comprobado a partir de la dispersión de los residuos con respecto a las puntuaciones estimadas por los modelos (figura 6.13), se cumple también para el caso de los residuos de nivel 1, presentando un R^2 de 2.122E-4.

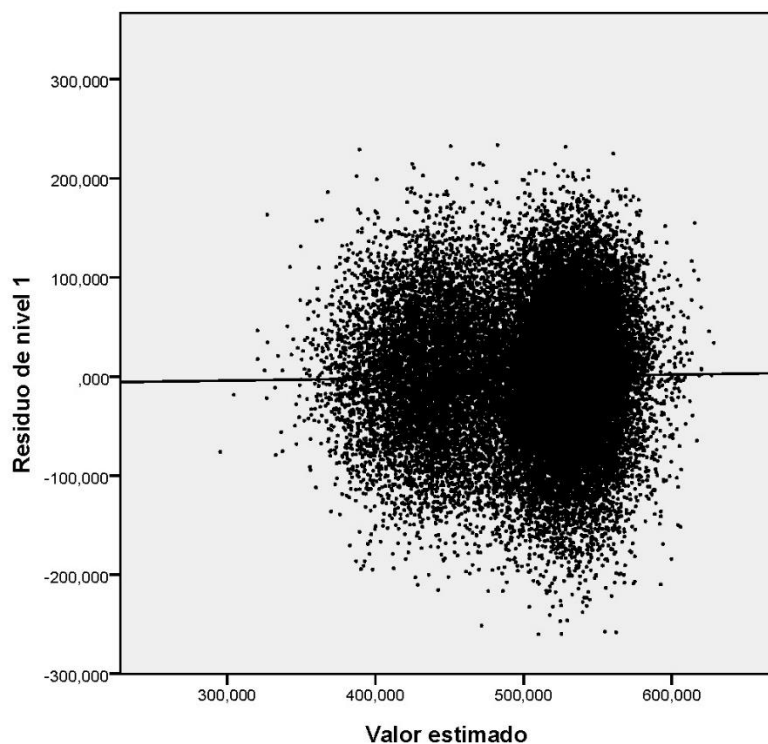


Figura 6.13. Verificación del supuesto de homocedasticidad para los residuos de nivel 1 de ciencia.

Por su parte, los residuos de nivel 2 de la competencia científica presentan una media muy cercana a 0 (tabla 6.26) y una desviación típica de 13.6.

Tabla 6.26. Estadísticos descriptivos de los residuos de nivel 2 para ciencia.

	N	Mínimo	Máximo	Media	Desviación típica	Asimetría		Curtosis	
						Valor	Error típico	Valor	Error típico
Residuo N1	896	-57,52	61,33	,16	13,60	-,13	,08	2,26	,16

La distribución de los residuos de segundo nivel (figura 6.14) se asemeja a la normal (nivel de significación del índice Kolmogorov-Smirnov de .002), aunque presenta una desviación en cuanto a su curtosis (distribución leptocúrtica) y una ligera asimetría negativa.

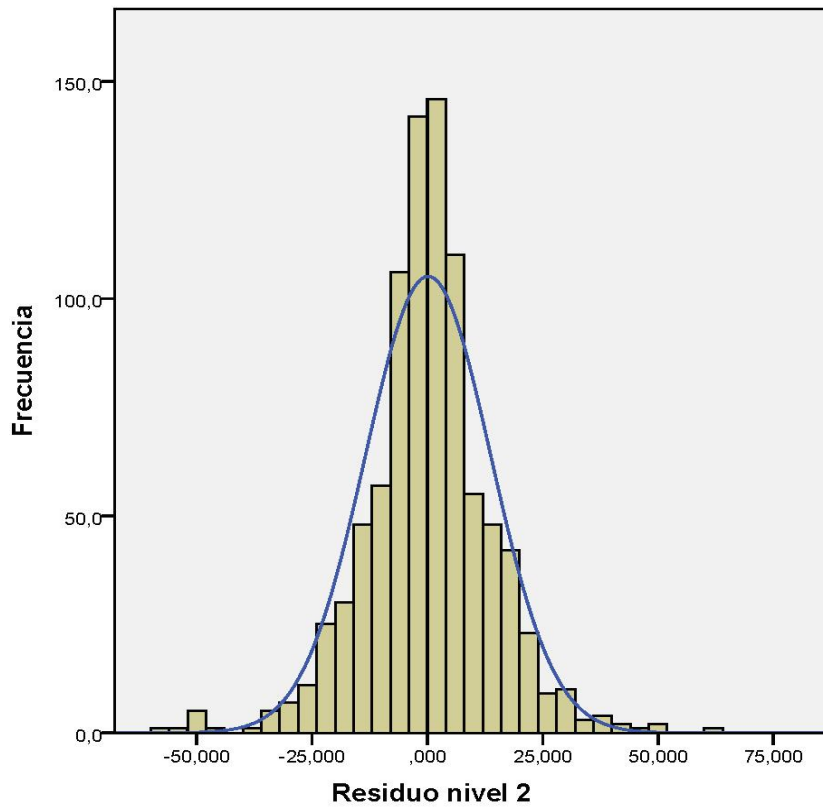


Figura 6.14. Distribución de los residuos de nivel 2 para ciencia.

Debido a la ligera desviación de la normalidad que presenta esta variable, se decide explorar también un gráfico de tipo Q-Q (figura 6.15), encontrando que, en general, los valores observados se ajustan a la normalidad, aunque existen algunos *outliers* en los extremos que se alejan de ella.

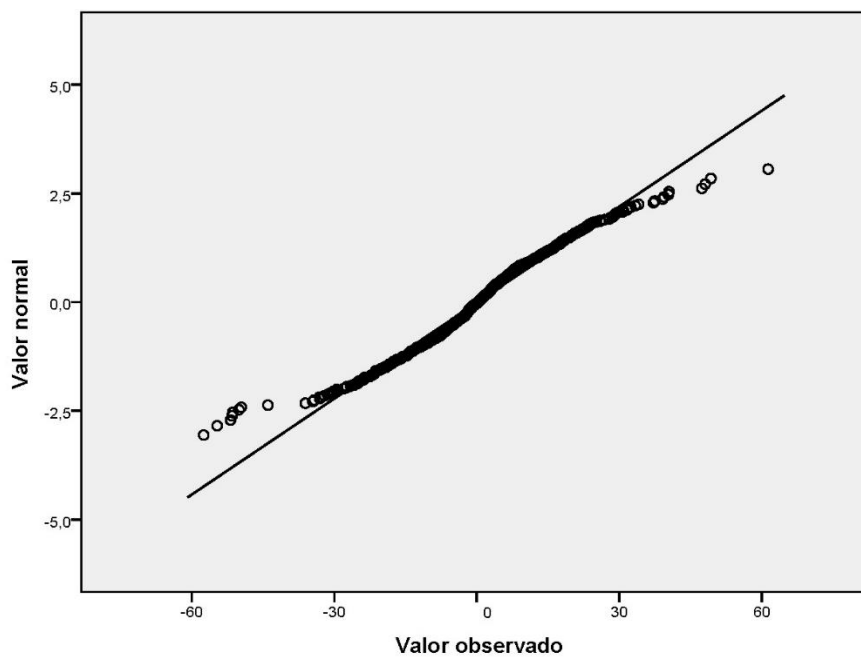


Figura 6.15. Gráfico Q-Q de los residuos de nivel 2 de ciencia.

Por último, los residuos de segundo nivel de la competencia científica cumplen con el supuesto de homocedasticidad con un R^2 de 0.005, como puede observarse en la figura 6.16.

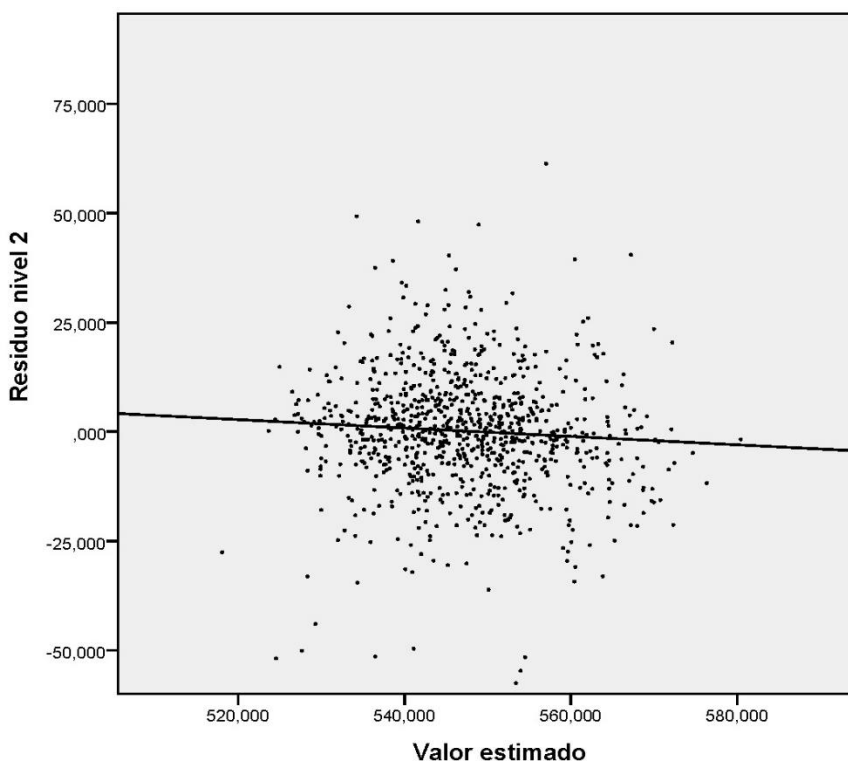


Figura 6.16. Verificación del supuesto de homocedasticidad para los residuos de nivel 2 de ciencia.

Teniendo en cuenta que todos los supuestos se cumplen de manera aceptable, el modelo explicativo para la competencia científica de los estudiantes españoles presentado en la tabla 6.23 se establece como definitivo.

6.2. SELECCIÓN DE CENTROS DE ALTA Y BAJA EFICACIA

A partir de la estimación de las puntuaciones medias esperadas para los centros y sus correspondientes residuos de nivel 2, se realizó una selección de las escuelas con residuos más altos y más bajos en las tres competencias evaluadas.

Para ello, se llevó a cabo el mismo protocolo en las tres competencias, consistente en los siguientes pasos:

1. Clasificación de los centros de mayor a menor en función del valor de su residuo.
2. Asignación a cada centro de un número correspondiente a su orden en la lista.

3. Cálculo de la posición relativa (percentil) de cada centro (posición en la lista dividida por el total de unidades).
4. Asignación de un punto positivo (1) a aquellos centros que se encuentran en el 33% inferior de la distribución, es decir, aquellos con residuos más altos.
5. Asignación de un punto negativo (-1) a aquellos centros que se encuentran en el 33% superior, es decir, aquellos con residuos más bajos.
6. Asignación de un 0 a aquellos centros que no hayan sido seleccionados en los dos pasos anteriores.

Tras realizar estos pasos para cada una de las competencias, se realiza la suma de los puntos asignados a cada centro en las tres clasificaciones, quedando un rango de posibles puntuaciones de -3 (centros seleccionados como de baja eficacia en las tres competencias) a 3 (centros seleccionados como de alta eficacia en las tres competencias). A partir de este cálculo se realiza la selección final de centros de interés para su posterior estudio, conformada por aquellos centros que tengan puntuaciones de -3 (196 centros) o 3 (189 centros).

Como paso previo al estudio de los factores de proceso en relación con la eficacia de los centros, se realiza un estudio descriptivo de ambos grupos (alta y baja eficacia) (tabla 6.27) para comprobar su equivalencia en las variables contextuales de caracterización y en sus puntuaciones esperadas para las tres competencias, ya que debería haber una distribución equivalente de estas variables en ambos grupos.

Tabla 6.27. Estadísticos descriptivos de las variables de los centros en función de su eficacia.

Variable	Baja eficacia				Alta eficacia			
	Min.	Máx.	Media	D.t.	Min.	Máx.	Media	D.t.
Tamaño centro	25.00	2635.00	747.54	445.44	100.00	2117.00	676.91	365.77
Tamaño clase	13.00	53.00	27.23	9.43	13.00	53.00	26.89	8.33
Escasez recursos	-1.25	3.61	0.18	1.07	-1.25	3.63	0.19	1.17
Escasez profesores	-1.68	3.72	0.24	0.93	-1.68	2.42	0.21	0.96
Ratio	1.85	39.24	12.81	3.99	2.38	41.51	12.18	3.97
ESCS centro	-1.92	0.97	-0.42	0.66	-1.55	1.14	-0.53	0.56
Prop. repetidores	0	.67	.27	0.16	0	.79	.30	0.14
Prop. inmigrantes	0	.69	.09	0.13	0	.58	.13	0.11
Prop. alumnas	0	.80	.49	0.10	0	.71	.50	0.09
P. esp. mate	517.66	556.01	537.55	8.79	522.59	558.19	536.13	7.39
P. esp. lectura	520.79	570.42	544.64	10.23	524.98	566.85	543.51	8.34
P. esp. ciencia	518.09	576.33	547.82	10.64	524.99	572.19	546.73	8.73

La realización de una prueba t para grupos independientes ($\alpha=.005$) revela que no existen diferencias estadísticamente significativas en ninguna de estas variables según la eficacia de los centros, a excepción de la proporción de alumnos inmigrantes, que es significativamente mayor en el grupo de centros de alta eficacia ($p=.002$).

Con respecto a la titularidad de los centros, su distribución (tabla 6.28) se ajusta a la muestra total tanto en la selección de centros de alta eficacia ($\chi^2=3.11$, $p=.21$) como en la de centros de baja eficacia ($\chi^2=2.58$, $p=.27$).

Tabla 6.28. Distribución de los centros en función de su titularidad.

		Titularidad			Total
		Pública	Concertada	Privada	
Muestra total	Recuento	564	239	44	847
	%	62.9	26.7	4.9	100
Alta eficacia	Recuento	135	44	7	186
	%	72.6	23.7	3.8	100
Baja eficacia	Recuento	110	58	12	180
	%	61.1	32.2	6.7	100

Por último, la distribución de centros de alta y baja eficacia por Comunidad Autónoma (figura 6.17) no es equivalente a la muestra inicial en la que todas las CCAA aportan el mismo número de centros (excepto el País Vasco que aporta el doble), percibiéndose CCAA en las que se seleccionan muy pocos centros (La Rioja, Asturias, Cantabria, todas con 10 centros o menos), o muchos centros, como es el caso de Castilla y León, Andalucía y Cataluña, con 30 centros o más.

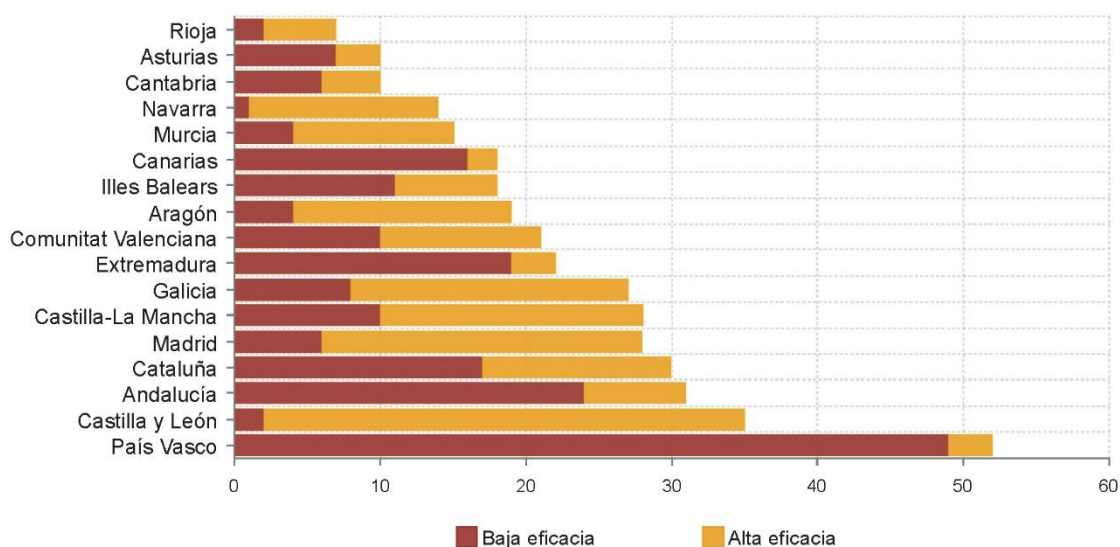


Figura 6.17. Distribución de centros de alta y baja eficacia por Comunidad Autónoma.

6.3. ANÁLISIS DEL EFECTO DE LAS VARIABLES DE PROCESO EN LA EFICACIA DE LOS CENTROS

Para responder al tercer objetivo de la tesis, que consiste en estudiar qué variables de proceso pueden estar asociadas con el nivel de eficacia de los centros educativos, se plantean dos análisis complementarios. El primero, que se presenta a continuación, se basa en un análisis de regresión logística, que es la técnica más indicada para el estudio de variables criterio dicotómicas (Kleinbaum, Dietz, Gail, Klein, & Klein, 2002).

En primer lugar, se seleccionaron aquellas variables de tipo procesual relacionadas con el centro educativo sobre las que los directores habían informado a través del cuestionario de contexto de PISA (ver sección 5.3.5 y Anexo II para más información sobre estas variables).

El análisis correlacional realizado como paso previo al empleo de la técnica de regresión señaló la ausencia de relaciones significativas entre las variables introducidas en el modelo y la eficacia de los centros, por lo que no se procedió con el análisis de regresión.

En segundo lugar, se procedió a repetir el análisis con las variables procesuales del nivel de estudiante agregadas al nivel de centro (ver sección 5.3.5 y Anexo II), para comprobar la existencia de alguna posible relación. El estudio correlacional mostró relaciones significativas entre un gran número de variables y la variable criterio de selección de centro, por lo que se continuó con el análisis basado en la regresión para comprobar el efecto de estas variables con la eficacia de los centros (tabla 6.29).

En la regresión logística, la magnitud de la relación entre una variable predictora y la variable criterio (*Odds Ratio*) se traduce en que el incremento de una unidad en la predictora produce una mayor probabilidad de que el centro pertenezca al grupo de alta eficacia. Antes de analizar los resultados, cabe señalar que todas las variables presentadas en la tabla están estandarizadas, con una media aproximada de 0 y desviación típica de 1, por lo que son directamente comparables. La única excepción es la variable “estatus ocupacional esperado”, que tiene una media de 60 y una desviación típica de 16. Debido a su escala, se entiende que el impacto de esta variable es muy bajo.

Tabla 6.29. Regresión logística para variables de proceso de nivel 1 agregadas a nivel de centro

Variable	B	Sig.	Odds Ratio	Odds Ratio convertida
Autonomía en el uso de TIC	2.77	0.001	15.92	
Adaptación de la instrucción	2.47	0.004	11.77	
Creencias epistemológicas	2.02	0.007	7.51	
Autoeficacia en ciencias	1.78	0.001	5.94	
Disfrute de la cooperación	1.63	0.029	5.09	
Conciencia medioambiental	1.44	0.018	4.23	
Clima de disciplina en ciencias	1.44	0.001	4.20	
Estatus ocupacional esperado	-0.09	0.008	0.91	1.10
Injusticia percibida del profesorado	-0.88	<.001	0.41	2.42
Uso de TIC en el centro	-1.33	0.001	0.26	3.79
Pertenencia al centro	-1.63	0.003	0.20	5.10
Retroalimentación percibida	-1.65	0.006	0.19	5.21
Apoyo del profesor en ciencias	-1.87	0.014	0.15	6.49
TIC como tema de interacción social	-2.57	0.005	0.08	12.99
Apoyo emocional familiar	-3.61	<.001	0.03	37.04

Nota: La columna “Odds Ratio convertida” presenta la probabilidad de aquellas variables que presentan una OR menor que 1 convertida (1/OR) para facilitar su comparabilidad.

Las variables con valores positivos han sido de diversa naturaleza. En cuanto a las prácticas de aula, tanto la adaptación de la instrucción por parte del docente a las necesidades de los estudiantes como el clima de disciplina aumentan la posibilidad de que un centro pertenezca a la categoría de alta eficacia. Por su parte, las actitudes o habilidades de los estudiantes relacionadas con la autonomía en el uso de las nuevas tecnologías, sus creencias epistemológicas en relación con la ciencia, su eficacia autopercibida, hasta qué punto disfrutaban del trabajo en equipo, o cuál es su nivel de conciencia con respecto al medioambiente también han resultado tener una relación positiva con la eficacia de los centros.

Por otro lado, las variables con una relación negativa más relevante con la eficacia se han correspondido principalmente con dos temas: las relaciones de los estudiantes con su familia y con los docentes, y las TIC. La variable que indica una mayor probabilidad de que un centro esté en la categoría de baja eficacia, con una *Odds Ratio* notablemente superior a todas las demás, es el apoyo emocional familiar. Esto quiere decir que, cuanto más apoyo familiar perciben los alumnos de un centro en promedio, más probabilidades tiene dicho centro de presentar una baja eficacia. Otras variables que también muestran esta relación son el apoyo prestado por el docente durante las clases y la retroalimentación académica percibida por los alumnos. En relación con las TIC, tanto su uso en el centro

como herramienta didáctica, como el hecho de que las tecnologías constituyan un tema en el que se basan las interacciones sociales de los alumnos también establecen una relación negativa con la variable de la eficacia.

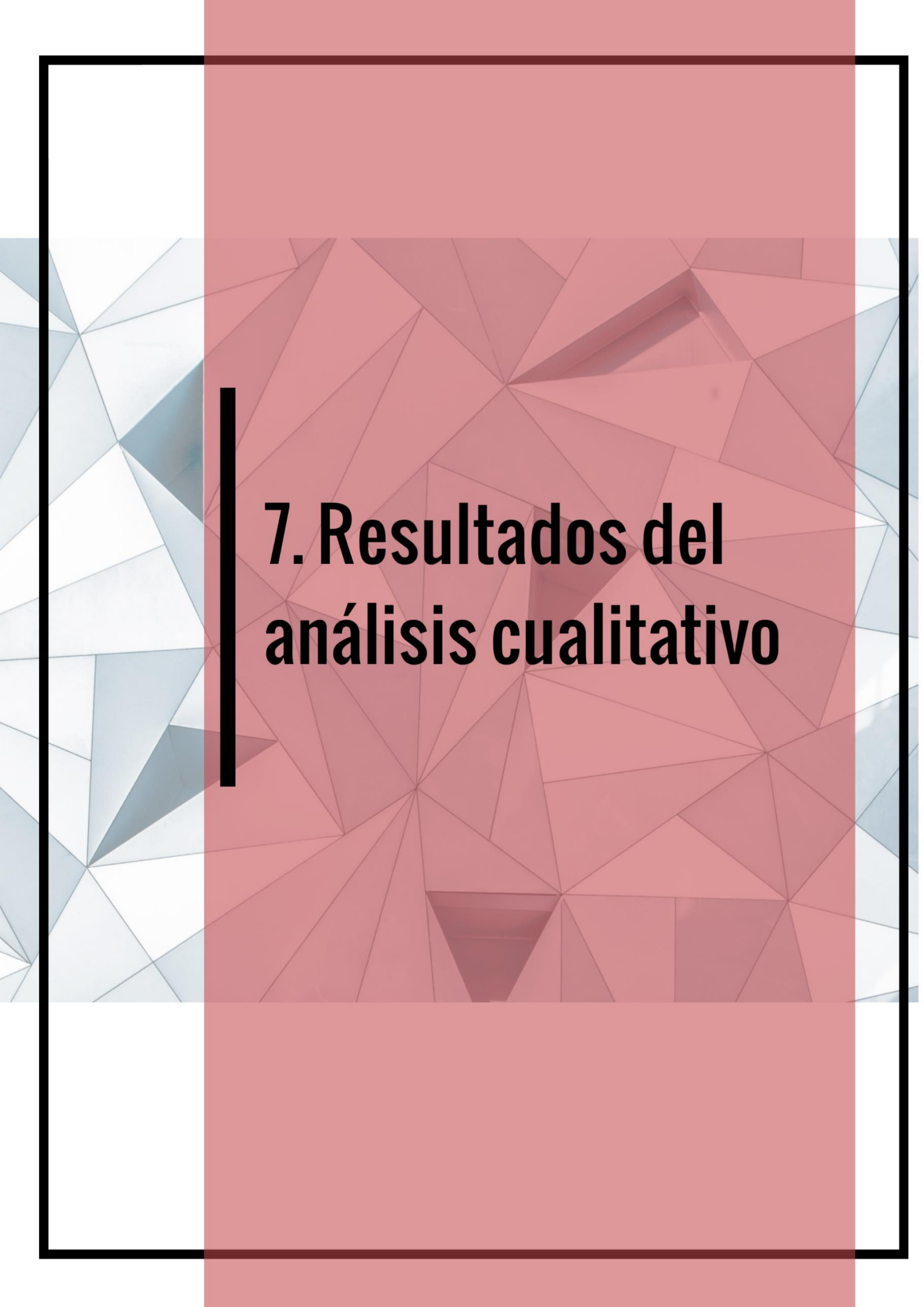
El hecho de que tanto el apoyo por parte del profesor como la retroalimentación percibida tengan una relación negativa con la eficacia resulta contradictorio con la relación positiva presentada por la adaptación de la instrucción realizada por los docentes en el aula, que en cierta manera también representa apoyo y atención a los alumnos por parte del profesor. Sin embargo, el apoyo familiar también se relaciona de manera negativa con la eficacia. El comportamiento negativo inesperado de estas tres variables apunta la necesidad de realizar un análisis más profundo de sus posibles causas. Si la variable criterio fuese el rendimiento podría aventurarse que los alumnos con menos rendimiento son los que necesitan más apoyo y retroalimentación, tanto por parte de los docentes como de sus familias. Sin embargo, tanto en los centros eficaces como ineficaces existe una heterogeneidad de rendimiento tan amplia como en la propia muestra, por lo que se hace necesario explorar otras posibles hipótesis, como un ambiente más generalizado de sobreprotección al alumnado en los centros de baja eficacia, o un mayor nivel de madurez y autorregulación entre el alumnado de los centros de alta eficacia que requiera un menor apoyo familiar y docente. Para ello, habría que buscar fuentes alternativas de información, ya que las variables ofrecidas por los cuestionarios de contexto de PISA no ofrecen datos suficientes como para poder dar sentido a estos resultados.

Por último, puede afirmarse que el ajuste del modelo es adecuado, obteniéndose un $R^2=.527$ (Índice de Nagelkerke). Por otro lado, la precisión del modelo predictivo (tabla 6.30) alcanza un 80,52% de clasificaciones correctas de centros de alto y bajo residuo, prediciendo correctamente un 79,59% en el caso de centros de bajo residuo y un 81.48% de los de alto residuo.

Tabla 6.30. Precisión del modelo predictivo para variables de proceso de nivel 1.

		Grupo predicho		Porcentaje correcto
		Baja eficacia	Alta eficacia	
Grupo observado	Baja eficacia	156	40	79,6
	Alta eficacia	35	154	81,5
Porcentaje total				80,5

El cumplimiento del tercer objetivo del presente trabajo de investigación se consigue sólo parcialmente a través de este análisis, ya que las variables de interés primario, que son los procesos a nivel de centro, no muestran ninguna correlación con la variable indicativa de la eficacia, y las variables de primer nivel informan de manera más difusa sobre las políticas y prácticas del centro. Por tanto, el análisis realizado en el siguiente capítulo, en el que se presentan los resultados del estudio cualitativo, servirá para complementar la información sobre este objetivo.



7. Resultados del análisis cualitativo

Resultados del análisis de datos cualitativos

A continuación, se exponen los resultados de la fase de análisis de datos cualitativos, que pretende explorar qué procesos escolares pueden estar relacionados con el valor añadido o la eficacia de los centros educativos.

En primer lugar, cabe recordar que la muestra para esta fase estuvo compuesta por dos centros, seleccionados por su alta o baja eficacia detectadas en la fase cuantitativa anterior. Sin embargo, una vez se había iniciado el proceso de recogida de datos, algunos de los docentes y los integrantes del equipo directivo del centro B (baja eficacia), que es un centro con un porcentaje alto de alumnos de inclusión, señalaron que en la aplicación de PISA 2015 habían participado varios estudiantes con necesidades educativas especiales que habitualmente necesitan adaptaciones para poder seguir las clases con normalidad y realizar las pruebas de evaluación.

En España, los estudiantes con necesidades educativas especiales (NEE) diagnosticadas por especialistas no participan en la prueba PISA, siendo excluidos y sustituidos por otros alumnos en caso de que sean seleccionados en el muestreo aleatorio. Para los estudiantes con NEE diagnosticadas, PISA ofrece una versión del cuestionario adaptada a sus necesidades, denominada *Une Heure* (OECD, 2017a). Sin embargo, los organizadores de la prueba en España (Instituto Nacional de Evaluación Educativa, INEE) consideran que este cuestionario no está bien adaptado a las características de los estudiantes NEE, por lo que se ha decidido no participar en dicha modalidad (L. Cercadillo, comunicación personal, 17 de junio de 2019)²⁰. No obstante, aquellos estudiantes que no tienen un diagnóstico oficial pueden ser seleccionados para participar en la prueba ordinaria, aunque tengan algún tipo de adaptación o apoyo otorgado por parte del centro educativo, lo cual ha sucedido en el centro B.

Por este motivo, aunque a lo largo de este capítulo se presentarán de manera comparada los resultados de la exploración cualitativa en ambos centros seleccionados, la etiqueta de “baja eficacia” asignada al Centro B debería ser considerada a la luz de estos hechos.

²⁰ Esta información se deriva de una consulta realizada a los expertos técnicos del INEE encargados de la gestión y aplicación de las pruebas PISA. En la sección 5.3.5 puede encontrarse más información sobre esta cuestión.

7.1. EL SISTEMA DE CATEGORÍAS

En primer lugar, el sistema de categorías utilizado para el análisis de los datos cualitativos se deriva del trabajo realizado por Lizasoain y Angulo (2014), que consiste en 14 categorías con sus subcategorías correspondientes, y la revisión de dichas categorías expuesta por Lizasoain et al. (2015), que reduce las 14 categorías a 9, quedando por tanto el sistema compuesto por los elementos presentados en la segunda columna de la tabla 7.1.

A lo largo del proceso de análisis de contenido, se llevó a cabo un procedimiento inductivo de categorización en el que algunas subcategorías desaparecieron, otras se combinaron y se añadió una nueva en respuesta a cuestiones planteadas por los sujetos. Aquellas subcategorías que no recibieron mención alguna por parte de ninguno de los seis grupos de informantes se retiraron de la lista final.

Por otro lado, algunas subcategorías de temática parecida se combinaron, al no encontrar en los testimonios de los informantes ninguna característica que permitiera proporcionar información diferencial para cada una de ellas. Un ejemplo de esta práctica es la combinación de las subcategorías originales “Metodologías innovadoras en lenguas” y “Metodologías innovadoras en matemáticas y ciencias” en la nueva subcategoría “Metodologías innovadoras”, ya que los informantes que trataron este tema durante los grupos de discusión no hicieron alusión al área en el que se aplicaban las metodologías mencionadas. Otras subcategorías finales que surgieron de la combinación de dos o más subcategorías originales son “Relación con las familias”, “Implicación con organizaciones externas”, “Organización de la coordinación” o “Evaluación de programas, profesores y centro”.

Por último, la nueva subcategoría incluida está relacionada con las expectativas de los docentes relacionadas con el rendimiento de los estudiantes. Esta subcategoría fue incluida en la categoría general de “Evaluación”, debido a su relación temática con esta. Las subcategorías utilizadas finalmente se reflejan en la tercera columna de la tabla 7.1.

Tabla 7.1. Sistemas de categorías a priori y final para el análisis cualitativo

Categoría	Subcategorías iniciales		Subcategorías finales
1. Proyectos, planes y formación	Planes y proyectos en activo Contenidos y origen de la formación externa Estilos y receptores de la formación Ubicación en el horario Actitud hacia la formación e innovación		Planes y proyectos en activo Contenidos y origen de la formación externa Ubicación en el horario Actitud ante la formación e innovación
2. Metodología y materiales de enseñanza	Metodologías innovadoras en lenguas Metodologías innovadoras en matemáticas y ciencias Metodologías tradicionales y uso del libro de texto Metodologías propias Uso de las TIC Trabajo por competencias Trabajo en valores		Metodologías innovadoras Uso de las TIC Trabajo por competencias
3. Atención a la diversidad y seguimiento del alumnado	Criterios de agrupamiento del alumnado Alumnado NEE Alumnado inmigrante Alumnado con bajo rendimiento y repetidores Otros (absentistas, altas capacidades...) Profesorado orientador, orientación Plan tutorial, tutoría Seguimiento del alumnado y atención individualizada		Criterios de agrupamiento del alumnado Alumnado NEE Alumnado con bajo rendimiento y repetidores Profesorado orientador, orientación Seguimiento del alumnado y atención individualizada
4. Evaluación	Alumnado	Importancia Evaluación externa Criterios de evaluación Evaluación inicial Evaluación integral Evaluación individualizada Auto y co-evaluación	Expectativas con respecto al rendimiento de los alumnos Evaluación de programas, profesores y centro
	Programas, docentes y centros	Tipos de evaluación Valoración de las evaluaciones Uso de la evaluación Herramientas	

Categoría	Subcategorías iniciales	Subcategorías finales
5. Gestión del tiempo	Aprovechamiento de las horas de clase Puntualidad en entradas y salidas, protocolos sobre el control de la puntualidad Organización de los tiempos y horarios del alumnado en extraescolares Más tiempo de horas lectivas del alumnado que las marcadas por ley	Puntualidad en entradas y salidas Organización de tiempos y horarios de alumnos en extraescolares
6. Liderazgo y equipo directivo	Estabilidad en el centro y en la dirección Creencias-filosofía, sentimientos-emociones, clima, implicación y expectativas Estilos de equipo directivo Formación, experiencia, innovación y mejora Liderazgo y presencia de la dirección Nombramiento Proyecto, objetivos y funciones definidas de la dirección	Formación, experiencia, innovación y mejora Proyectos y objetivos del equipo directivo Filosofía, implicación y clima del liderazgo
7. Gestión y organización del centro. Implicación y pertenencia al centro	Modelos de organización Sistemas de calidad Cuidado del profesorado Estabilidad de la plantilla Tipos de coordinación Organización de la coordinación Fines de la coordinación Medios de coordinación	Organización de la coordinación Cuidado del profesorado Estabilidad de la plantilla Implicación, compromiso, entrega y disponibilidad
8. Clima y convivencia	Convivencia Clima Actividades para mejorar el clima Conflictos, resolución	Convivencia Clima Actividades para mejorar el clima Conflictos, resolución
9. Familia y comunidad	Valoración de las familias hacia el centro Grado de implicación de las familias en el centro Implicación pedagógica de las familias Formación para las familias Valoración del centro a las familias Grado de implicación con las organizaciones externas Relación con organizaciones externas: recursos, instalaciones Organizaciones externas: actividades Organizaciones externas: asistencial	Relación con las familias Implicación con organizaciones externas

Fuente: Lizasoain y Angulo (2014) y Lizasoain et al. (2015)

7.2. DESCRIPCIÓN GENERAL DE LOS CENTROS

En primer lugar, se presenta una panorámica general de los datos recogidos en ambos centros a través de las nubes de las palabras más representativas, es decir, aquellos términos que más se mencionaron durante los grupos de discusión (figura 7.1, tabla 7.1). Las palabras vacías de contenido, así como las palabras *alumnos*, *centro*, *profesores* y derivadas fueron retiradas, ya que eran las más repetidas en ambos centros.

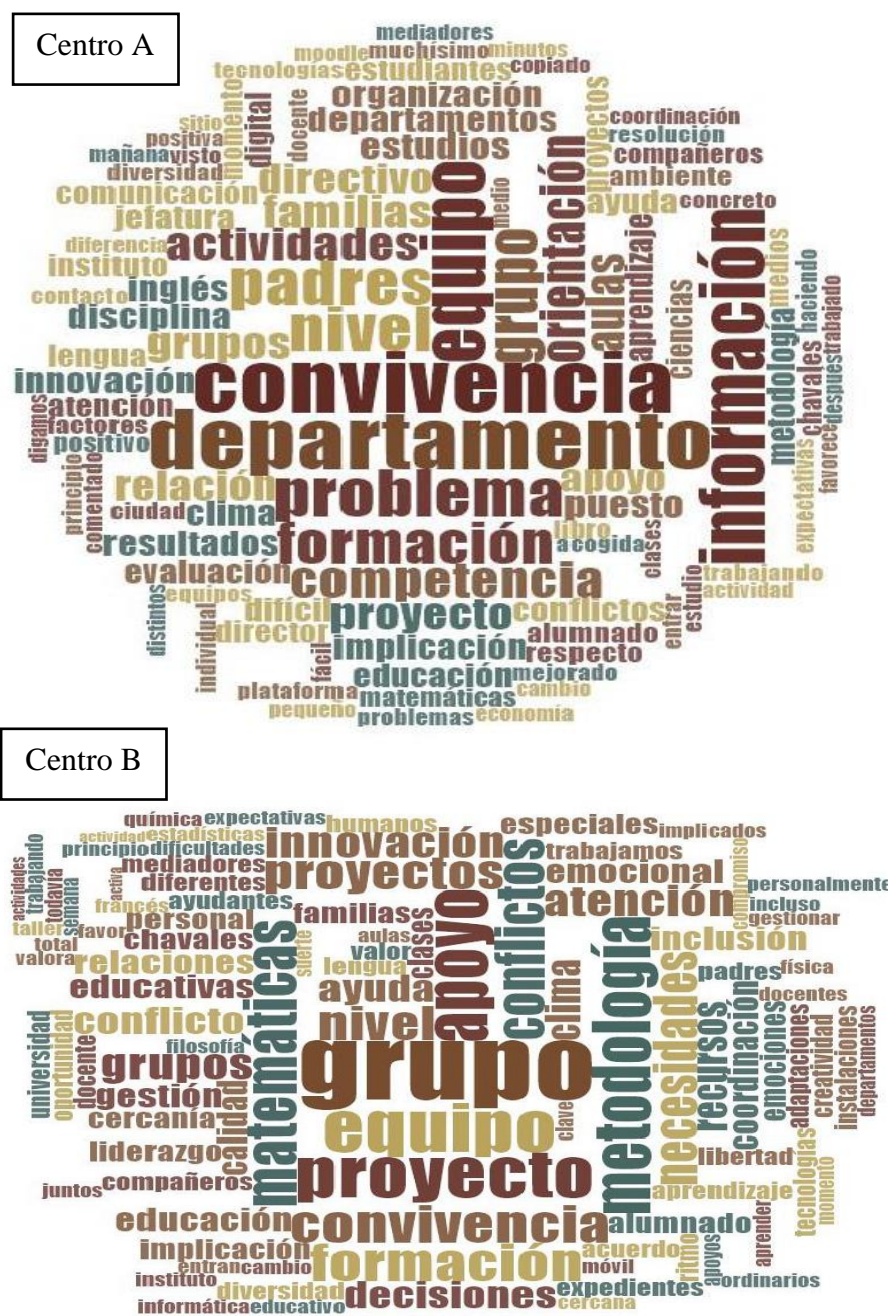


Figura 7.1. Nubes de palabras.

Tabla 7.1. Frecuencia de palabras más mencionadas por centro

Centro A			Centro B		
Palabra	Frecuencia	%	Palabra	Frecuencia	%
Convivencia	33	12,94%	Grupo	20	16,53%
Departamento	33	12,94%	Apoyo	13	10,74%
Equipo	29	11,37%	Equipo	13	10,74%
Información	29	11,37%	Proyecto	13	10,74%
Problema	24	9,41%	Metodología	12	9,92%
Nivel	23	9,02%	Matemáticas	11	9,09%
Padres	23	9,02%	Conflictos	10	8,26%
Grupo	22	8,63%	Convivencia	10	8,26%
Formación	21	8,24%	Formación	10	8,26%
Orientación	18	7,06%	Necesidades	9	7,44%

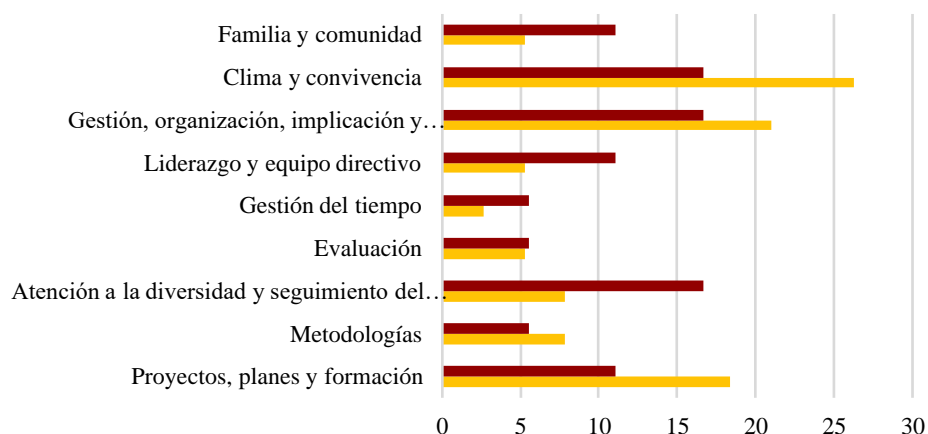
Como puede observarse, en el centro A destacan palabras como convivencia, información, o departamento, todas ellas indicativas de puntos importantes de la organización del centro, como su plan integral de convivencia, acogida e información de los alumnos, o la organización y convivencia de los profesores en función de sus departamentos. En el centro B destacan otras palabras como grupo, equipo, apoyo, necesidades, atención o metodología, que reflejan cuestiones básicas de su filosofía de trabajo con los alumnos: inclusión, trabajo en grupo en las asignaturas y un profesor de apoyo acompañando al profesor de la materia en casi todas las horas de enseñanza.

Otra forma de visualización de las principales diferencias en los temas emergentes en los centros es la exploración de la proporción de tiempo que los informantes (Equipo Directivo – E, docentes noveles – N, o docentes veteranos – V) ha dedicado a hablar de cada categoría (tabla 7.2, figura 7.2), o subcategoría (tabla 7.3, figura 7.3).

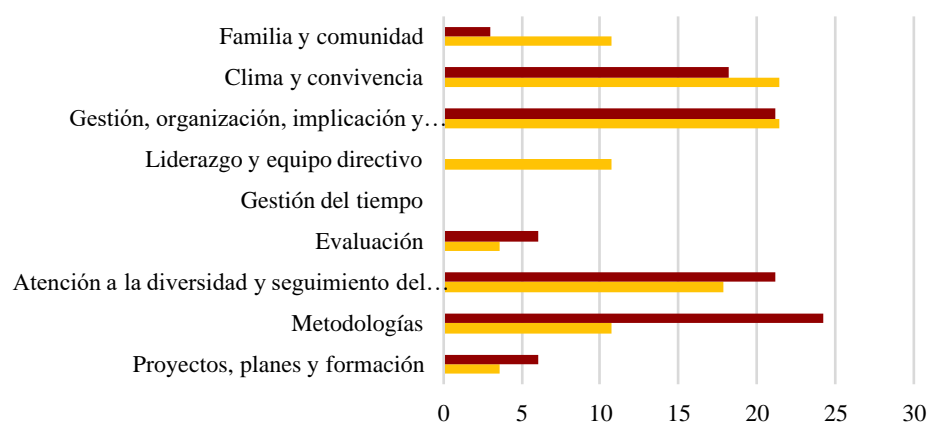
Tabla 7.2. Porcentaje de codificación de categorías según los informantes.

Categoría	A-E	A-N	A-V	B-E	B-N	B-V
1: Proyectos, planes y formación	18.42	7.69	3.57	11.11	11.76	6.06
2: Metodologías	7.89	9.62	10.71	5.56	26.47	24.24
3: Atención a la diversidad y seguimiento del alumnado	7.89	19.23	17.86	16.67	29.41	21.21
4: Evaluación	5.26	5.77	3.57	5.56	5.88	6.06
5: Gestión del tiempo	2.63	1.92	0.00	5.56	0.00	0.00
6: Liderazgo y equipo directivo	5.26	5.77	10.71	11.11	2.94	0.00
7: Gestión, organización, implicación y pertenencia	21.05	9.62	21.43	16.67	8.82	21.21
8: Clima y convivencia	26.32	28.85	21.43	16.67	11.76	18.18
9: Familia y comunidad	5.26	11.54	10.71	11.11	2.94	3.03

Equipo directivo



Docentes veteranos



Docentes noveles

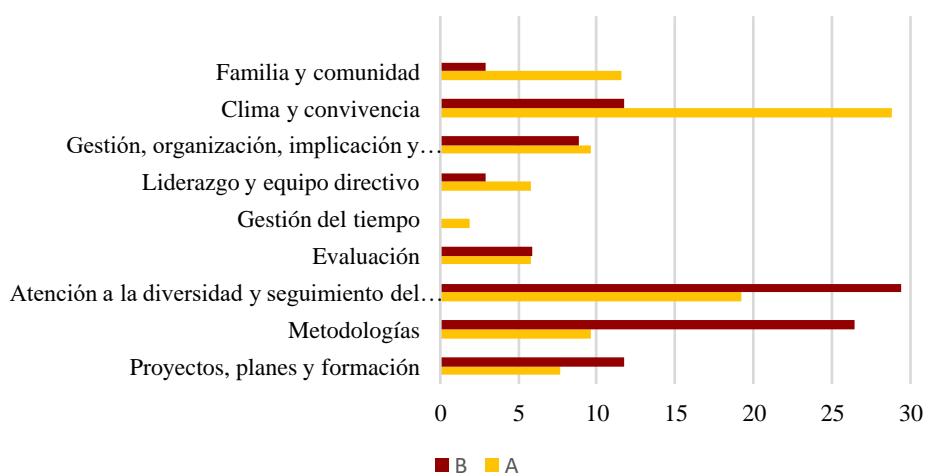


Figura 7.2. Porcentaje dedicado a cada categoría en función del centro y el grupo de informantes.

Tabla 7.3. Porcentaje de codificación de subcategorías según los informantes

Categoría y subcategoría	A-E	A-N	A-V	B-E	B-N	B-V
1 Planes y proyectos en activo	4.65	3.51	0.00	0.00	6.67	3.57
1 Contenidos y origen de la formación externa	4.65	3.51	3.33	5.88	3.33	3.57
1 Actitud hacia la formación e innovación	6.98	0.00	0.00	5.88	3.33	3.57
1 Ubicación en el horario	4.65	0.00	0.00	0.00	0.00	0.00
2 Metodologías innovadoras	0.00	5.26	3.33	0.00	10.00	3.57
2 Trabajo por competencias	0.00	0.00	0.00	0.00	6.67	3.57
2 Uso de las TIC	4.65	5.26	6.67	0.00	3.33	7.14
3 Alumnado con bajo rendimiento y repetidor	0.00	3.51	0.00	0.00	0.00	0.00
3 Alumnado NEE	0.00	3.51	0.00	5.88	16.67	3.57
3 Criterios de agrupamiento del alumnado	2.33	5.26	3.33	5.88	3.33	0.00
3 Profesor orientador, orientación	4.65	0.00	13.33	0.00	0.00	0.00
3 Seguimiento y atención individualizada	0.00	3.51	0.00	11.76	10.00	10.71
4 Evaluación de centro, docentes y programas	2.33	1.75	0.00	5.88	0.00	0.00
4 Expectativas sobre el alumnado	2.33	3.51	3.33	0.00	6.67	7.14
5 Organización de los tiempos y horarios del alumnado en extraescolares	2.33	1.75	0.00	0.00	0.00	0.00
5 Puntualidad en entradas y salidas	0.00	0.00	0.00	5.88	0.00	0.00
6 Filosofía, clima e implicación del liderazgo	4.65	3.51	10.00	5.88	3.33	0.00
6 Formación-experiencia, innovación y mejora	0.00	0.00	0.00	5.88	0.00	0.00
6 Proyectos y objetivos de la dirección	0.00	1.75	3.33	0.00	0.00	0.00
7 Cuidado del profesorado	0.00	0.00	6.67	0.00	0.00	0.00
7 Estabilidad de la plantilla	0.00	1.75	0.00	5.88	0.00	0.00
7 Implicación, compromiso, entrega	13.95	0.00	0.00	5.88	3.33	17.86
7 Organización de la coordinación	4.65	7.02	13.33	0.00	6.67	7.14
8 Actividades para mejorar el clima	6.98	3.51	3.33	0.00	3.33	0.00
8 Clima	11.63	12.28	10.00	11.76	6.67	3.57
8 Conflictos, resolución	4.65	7.02	3.33	5.88	6.67	7.14
8 Convivencia	9.30	12.28	6.67	0.00	0.00	14.29
9 Implicación con organizaciones externas	2.33	5.26	3.33	5.88	0.00	0.00
9 Relación con las familias	2.33	5.26	6.67	5.88	0.00	3.57

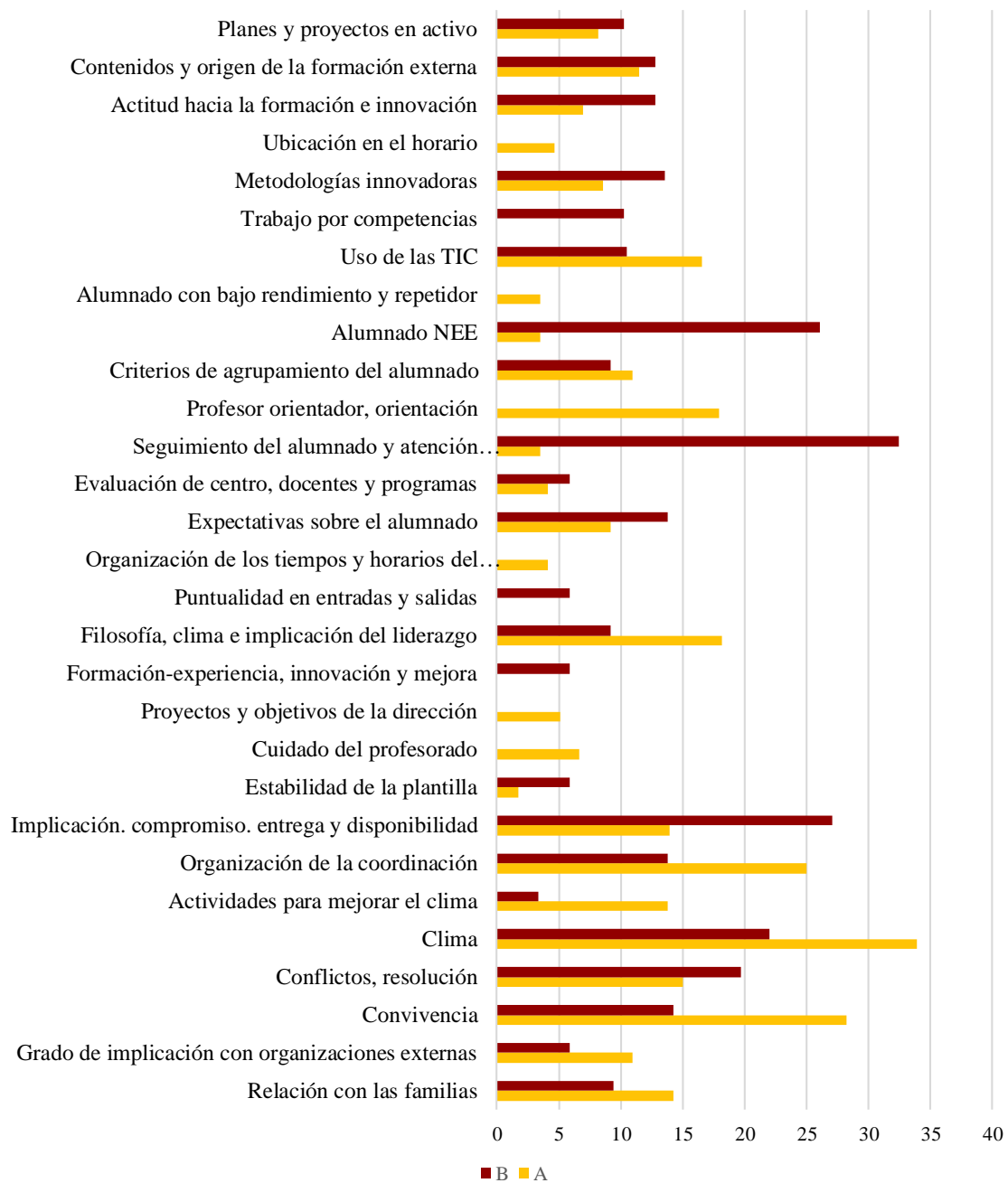


Figura 7.3. Porcentaje dedicado a cada subcategoría en función del centro.

A través de estos análisis se aprecia que en el centro A se dedicó más tiempo a hablar sobre el tema del clima y la convivencia, así como de la atención a la diversidad, centrándose principalmente en los temas de agrupamiento del alumnado y la orientación. En el centro B, en cambio, el tema principal es la atención a la diversidad, especialmente relacionada con el seguimiento y la atención individualizada de los alumnos y los estudiantes con necesidades educativas especiales, y el segundo tema en importancia es el de las metodologías.

Por último, antes de comenzar con la descripción pormenorizada de las categorías, cabe señalar que la agrupación de los recursos (grupos de discusión) en función de su similitud en la codificación, estudiada a través del coeficiente de Jaccard²¹ da como resultado una organización pertinente, siendo más similares entre sí los recursos de los informantes que pertenecen al mismo centro (A o B), y dentro de estos, existiendo más similitud entre los testimonios de los docentes (noveles y veteranos) que entre estos y el equipo directivo (figura 7.4).

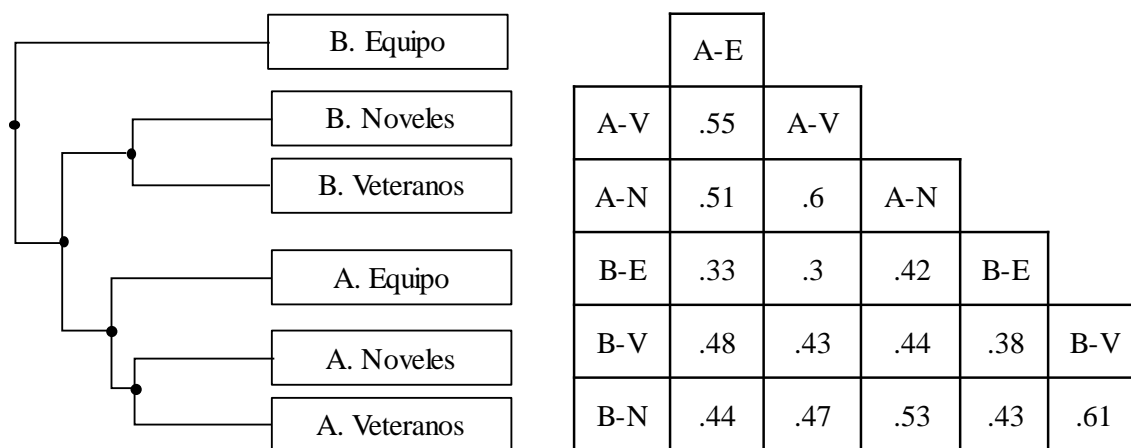


Figura 7.4. Similitud de codificación de los recursos en función del centro y el informante

7.3. COMPARACIÓN DE CATEGORÍAS

En esta sección se presentan las principales opiniones expresadas por los participantes de ambos centros, organizadas en función de las categorías y subcategorías a las que pertenecen. Se exponen también algunos extractos de los grupos de discusión para ejemplificar, matizar, o ampliar la información sobre las categorías.

7.3.1. Planes, proyectos y formación

Planes y proyectos en activo

Ambos centros refieren diversos proyectos en los que se encuentran trabajando actualmente, aunque se aprecia una diferencia en su temática y sus objetivos. Los proyectos del centro A están enfocados al trabajo de contenidos o competencias de corte más académico, como el programa Aula-Empresa, el plan de Fomento de la Lectura, el Plan de Acompañamiento, que consiste en un apoyo extra para alumnos con bajo rendimiento, u otros proyectos que no nombran específicamente, pero vinculan con

²¹ Coeficiente de similitud con un rango entre 0 (ninguna similitud) y 1 (similitud total).

departamentos específicos como Economía o Ciencias. Por otra parte, también tienen un proyecto de Innovación y Formación del profesorado, cuya organización y temas se tratarán con más profundidad en la categoría correspondiente.

El centro B, por su parte, menciona proyectos cuyos objetivos se dirigen a fomentar la inclusión y el desarrollo de competencias de carácter transversal. El trabajo relacionado con la inclusión se evidencia a través de proyectos como el de Patios Dinámicos, que propone juegos inclusivos a los alumnos durante los recreos, o un proyecto de colaboración que han puesto en marcha con ASPRODES (una asociación para personas con discapacidad intelectual), a través del cual hay visitas de los usuarios de la asociación a los alumnos del centro y viceversa. Por otro lado, también han establecido un proyecto llamado Aula de Paz, que consiste en la habilitación de un aula cuyo uso está relacionado con el desarrollo del control y las habilidades emocionales de los alumnos, y está a disponibilidad de todos los docentes.

Contenido y origen de la formación externa

En el centro A, el plan de Innovación y Formación del Profesorado está centrado en tres líneas principales: las nuevas tecnologías en educación (uso de la plataforma Moodle, seguridad en la red), implantación de la FP dual, y aplicación de las nuevas tecnologías educativas en las ramas de Ciencias Sociales y Humanidades.

Además, tanto los docentes como el equipo directivo resaltan la implicación a nivel individual en la formación docente, al margen de las líneas del centro:

Luego a nivel individual cada uno hace su formación, ya sabéis que los Centros de Profesores ofertan formación y, nos enteramos siempre al final del curso, cuando te pasan la memoria de cada departamento, en todo lo que se han formado los profesores ese año. (Jefa de Estudios, centro A)

En el centro B parece existir una organización más uniforme de la formación docente desde el propio centro educativo. Sus líneas principales versan sobre nuevas tecnologías y educación emocional, dentro de las cuales se tratan temas distintos cada año. Adicionalmente, también se realizan cursos puntuales a nivel de todo el centro sobre otros temas de interés, como los Trastornos del Espectro Autista.

Ubicación en el horario

Esta subcategoría fue tratada únicamente en el centro A, en el que desde el Equipo Directivo se comentó que los docentes que participan en proyectos de innovación tienen que hacerlo en su tiempo libre, y reclaman más comprensión por parte de las autoridades educativas en cuanto a la implicación horaria que suponen estos proyectos.

Aunque se comentará en el apartado correspondiente, en el centro B hablaron también de trabajo docente realizado fuera del horario laboral, pero se referían a la coordinación necesaria para sacar adelante las adaptaciones curriculares de los alumnos y para ponerse de acuerdo entre los docentes de asignatura y los de apoyo que comparten horas en el aula.

Actitudes hacia la innovación y la formación

En esta categoría cabe analizar por separado las actitudes hacia la innovación y hacia la formación, ya que en algunos casos se han expresado de manera diferente.

En el centro A se perciben actitudes muy positivas hacia la formación, tanto la correspondiente a las líneas propias del centro, como a aquella que los docentes buscan por su cuenta. Esta implicación es especialmente destacable debido a la naturaleza del centro; al tratarse de un centro educativo público situado en una capital de provincia, los docentes que acceden a él lo hacen habitualmente como destino definitivo, por lo que la motivación extrínseca de acumular méritos no constituye un factor para ellos.

En el centro B, sin embargo, las actitudes hacia la formación son un tanto negativas en el caso de los docentes veteranos, que expresan hastío con el exceso de cursos que han acumulado a lo largo de sus carreras, y una cierta indiferencia en el caso de los docentes noveles, quienes no expresaron particular interés por el tema:

El profesorado yo creo que lo intenta [participar en la formación], pero ni más ni menos que en otros colegios. (PT novel, centro B)

La actitud hacia la innovación se percibió positiva en ambos centros, cuyos integrantes destacaron la gran implicación de los docentes en los proyectos propuestos. Cabe destacar aquí un aspecto relacionado con la categoría 1.1 (proyectos en activo), y es que mientras que en el centro A la implicación se orienta en su mayoría hacia proyectos basados en ideas externas al centro (metodologías nuevas, olimpiadas relacionadas con

asignaturas, etc.), en el centro B se percibe una mayor creatividad y voluntad de generar proyectos propios:

Cada profesor va proponiendo sus proyectos, otros profesores se le unen, y en ese sentido es una innovación basada en la propia creatividad de los profesores, no tanto a partir de la formación o de una metodología nueva, sino de la propia creatividad de los profesores, que van sacando proyectos que luego igual se pueden enmarcar en esos marcos teóricos como el aprendizaje y servicio, o el aprendizaje basado en problemas, pero no surgen por las metodologías en sí mismas. (Profesor de Matemáticas novel, centro B)

7.3.2. Metodologías

Uso de metodologías innovadoras

La metodología es una de las grandes diferencias encontradas entre ambos centros. En el centro A existe una libertad total para elegir la metodología docente, ya que el Equipo Directivo sostiene la filosofía de que no se debe obligar a los docentes a adoptar enfoques metodológicos concretos con los que pueden no sentirse cómodos, prefiriendo en su lugar promover el empleo de metodologías innovadoras a través de la formación. La percepción general es que no se perciben como un centro particularmente innovador:

Yo no lo concibo como un centro puntero en ese sentido, pero es justamente por ese argumento de que cada profesor hace lo que quiere, eso es ya de por sí contradictorio a ser puntero, porque si no tendríamos que tener grupos de trabajo que hicieran algo, (...) poner a principio de curso unos objetivos comunes en algún sentido... (Profesora de alemán veterana, centro A)

La principal referencia que hacen a las metodologías innovadoras es el uso del ordenador y el cañón en el aula, tanto para proyectar recursos como para poder realizar consultas en función de los intereses de los estudiantes:

Puedes usar diferentes recursos, vídeos... y si en algún momento algún alumno te pregunta sobre lo que está ocurriendo en este momento, sobre el Brexit, por ejemplo, y dices “ah, pues vamos a ver qué ha salido últimamente en la BBC”, y ellos lo ven mucho más atractivo, porque no es sólo centrarse en los contenidos gramaticales, sino también en las cuestiones culturales. (Profesora de inglés novel, centro A)

En cambio, en el centro B sí existe una línea metodológica innovadora general marcada por el trabajo en grupo, el cual se aplica en la práctica totalidad de las materias, encontrándose todas las aulas organizadas en mesas grupales en lugar de pupitres individuales. Asimismo, este es uno de los tópicos que se mencionaron con mayor frecuencia a lo largo de las sesiones con los tres grupos de informantes, siempre de manera positiva y valorándolo como un instrumento más para trabajar la inclusión dentro del centro:

Otra de las claves del éxito del centro es el trabajo en grupo, estar sentados en equipo, que todo el mundo tenga que participar sean cuales sean sus capacidades y limitaciones. Todo el mundo tiene algo que aportar, algo que decir dentro del grupo, y creo que es muy importante sobre todo porque es una enseñanza para su futuro. (PT novel, centro B)

Uso de las TIC

Un tema común en ambos centros es la explotación que se hace de los recursos tecnológicos del centro, los cuales son escasos en opinión de los informantes, estando los que hay disponibles constantemente en uso.

Por otro lado, así como en el Centro A se percibe un uso cotidiano de las TIC a través de los recursos disponibles en el aula ordinaria (ordenador y cañón), en el Centro B se reporta el uso de las tecnologías como algo más puntual. Las actividades que mencionan en este sentido se basan en llevar a los estudiantes al aula de informática, o en realizar actividades extraordinarias, como por ejemplo a través del uso de la aplicación *Kahoot!*.

Trabajo por competencias

El trabajo centrado en el desarrollo de competencias específicas sólo se ha tratado en los grupos de discusión del Centro B, en el que se trabaja de manera transversal el desarrollo de la competencia emocional. Este tema se fomenta tanto desde la formación del profesorado, siendo una de las líneas prioritarias de formación en el centro, como a través de proyectos de innovación (ver Proyecto Aula de Paz en la subcategoría de Proyectos en activo) y el trabajo transversal desde las asignaturas:

Se insiste mucho en la gestión de emociones, en ayudarles a gestionar sus emociones ya sea por un conflicto que tienen, o porque son demasiado

impulsivos (...) desde todas las diferentes áreas trabajamos mucho con eso. (PT novel, centro B)

7.3.3. Atención a la diversidad y seguimiento del alumnado

Las discrepancias encontradas en esta categoría, en la que en general se ha recogido mucha más información en el centro B, destacan una de las principales diferencias entre ambos centros, que es el perfil del alumnado al que atienden. Mientras en el centro A tienen un alumnado mayoritariamente ordinario, en el centro B tienen un alto porcentaje de alumnos con algún tipo de necesidad educativa, incluso más alto del que se refleja en la matrícula del centro:

El tipo de alumnado que solicita venir al centro es porque no encuentra en los centros en los que está respuesta a sus necesidades. Y eso no quiere decir que no vengan alumnos ordinarios, que también vienen, pero sí viene un porcentaje muy elevado de alumnos con necesidades educativas especiales, y en algunos casos te dicen que son ordinarios hasta que vienen al centro y te das cuenta de que son NEE, pero claro los padres hacen la trampa porque saben que, si no, no entra por el cupo de plazas. (Director, centro B)

Criterios de agrupamiento del alumnado

Ambos centros transmiten un gran cuidado a la hora de elaborar los agrupamientos de los estudiantes en 1º de ESO con el objetivo de realizar grupos homogéneos entre sí, de manera que no haya ninguno con una proporción mayor de estudiantes con necesidades educativas especiales o algún tipo de dificultad del aprendizaje. La orientadora del centro A menciona que la razón de esta filosofía para el agrupamiento de los estudiantes es generar un *efecto arrastre* a través del cual los buenos estudiantes impulsen el aprendizaje de aquellos con dificultades o con rendimientos más bajos.

Alumnado con necesidades educativas especiales

La naturaleza del centro B como un centro inclusivo explica en cierta manera que en este punto se hayan expresado enfoques diferentes. En el centro A, la atención del alumnado con necesidades educativas especiales se menciona de manera superficial, en general relacionándola con aspectos organizativos del centro, como la existencia de apoyos extraordinarios para los estudiantes con NEE, o la reducción del número máximo de alumnos en aulas donde hay alumnos NEE.

En el centro B, sin embargo, la atención a los estudiantes con necesidades se percibe como uno de los pilares fundamentales que guía muchas de las decisiones organizativas y pedagógicas del centro. Por un lado, está la medida de incorporar profesionales de PT a casi todas las materias, algo que se hace en todos los cursos de Educación Secundaria Obligatoria y que redundará en los alumnos con necesidades, y que todos los docentes (tanto los PT como los de materia) relatan como enriquecedor y positivo para alumnos y profesores. Por otro lado, el Equipo Directivo hace especial hincapié en que la organización anual del centro (horarios, apoyos, agrupamiento...) se realiza en torno al alumnado de nueva admisión, para poder ajustar los horarios y los recursos a las necesidades de los nuevos estudiantes:

Director: En este centro no se toman decisiones hasta que sabemos qué tipo de alumnado tenemos: repetidores, NEEs, etc., porque los horarios y todo dependen del tipo de alumnado, los apoyos, etc. Entonces nosotros no mandamos los horarios hasta casi prácticamente el día antes de empezar el curso, porque no sabemos cuántas horas va a tener el profesor.

Jefe de Estudios: A nosotros nos marca mucho el perfil del alumnado, cada alumno es diferente, el ordinario por supuesto, y el NEE más todavía, entonces la organización depende casi de cuestiones individuales. (Equipo Directivo, centro B)

Alumnado con bajo rendimiento, repetidores

Sólo los docentes del centro A han hecho referencia a la atención ofrecida a este tipo de alumnado. En el centro existen varios mecanismos para atender a aquellos estudiantes de la ESO que presentan un rendimiento bajo, como el ya mencionado Plan de Acompañamiento, dirigido a la reducción del fracaso y el abandono escolar, que consiste en clases extraordinarias de apoyo por las tardes. Además, para aquellos estudiantes que no se prevé que alcancen el título de la ESO, o que muestran interés por seguir sus estudios en Ciclos Formativos de Grado Medio, el centro pone a su disposición clases extraordinarias para la preparación de las pruebas de acceso a dichos Ciclos. Ambos programas son valorados muy positivamente por parte de los docentes, ya que los perciben motivadores para los estudiantes y opinan que es una manera productiva de mantenerlos dentro del sistema educativo por su propia voluntad.

Profesorado orientador, orientación

El tema del Departamento de Orientación se ha tratado únicamente en el centro A. Tanto los dos grupos de docentes como los miembros del Equipo Directivo han alabado la labor de la orientadora del centro en relación con cuestiones como la promoción de la convivencia, la atención personalizada al alumnado (por motivos académicos o personales), la mediación entre los estudiantes y las familias, y otras iniciativas para mejorar el funcionamiento del centro:

El Equipo de Orientación se está implicando mucho, se está implicado en problemas personales, no solo con los alumnos y con los grupos, sino también en que ayuda a mejorar el funcionamiento del Instituto. Por ejemplo, lo de hacer calendarios para que los chicos prevean bien sus fechas de exámenes, todo eso salió por parte de una idea del Equipo de Orientación. (Profesora de Economía veterana, centro A)

Seguimiento del alumnado y atención individualizada

Mientras que en el centro A no se percibe un enfoque de atención individualizada al alumnado, en el centro B se constituye como una cuestión fundamental de la identidad institucional. Todos los sujetos entrevistados han destacado cómo el hecho de tener dos profesores en el aula de manera casi permanente permite asistir de manera individual a los estudiantes, no sólo a los que tienen necesidades educativas especiales, sino a todos los alumnos en el aula:

Yo estoy dando apoyos en la ESO, estamos para los alumnos con necesidades educativas especiales, pero también ayudamos a cualquier alumno que en un momento preciso pueda necesitar una ayuda, un refuerzo individualizado, una adaptación de material, etc. (PT novel, centro B)

Los docentes de materia expresan también que la disponibilidad de profesionales de PT en sus aulas es un apoyo para ellos mismos, y les permite brindar una mejor atención a los alumnos.

Además, como ya se vio en la subcategoría de atención a los alumnos con NEE, en el centro B hay una filosofía de adaptar los agrupamientos y los apoyos a las necesidades de todos los estudiantes, y los docentes destacan que en el centro se respeta mucho el ritmo de aprendizaje de cada alumno.

7.3.4. Evaluación

En ninguno de los dos centros se dedicó mucho tiempo a hablar de los mecanismos de evaluación de los estudiantes, aunque en el centro B sí se comentó que los exámenes no eran la forma prioritaria de evaluación y no tenían un excesivo peso en las notas de los alumnos.

Expectativas sobre el alumnado

El tema que sí se discutió de manera más extensa fue el de las expectativas de los docentes sobre el rendimiento de los estudiantes, el cual, al igual que algunos de los temas discutidos anteriormente, estuvo mediado en gran parte por el tipo de alumnado que acude a cada uno de los centros.

En el centro A, los docentes vinculan sus expectativas del rendimiento de sus alumnos con la motivación por el aprendizaje que estos presentan. En el momento que se plantea el tema, los participantes expresan que se anima a todos los estudiantes a que estudien y se les intenta inculcar la importancia de terminar los estudios de educación secundaria. Sin embargo, tanto en el grupo de noveles como en el de veteranos mencionan la falta de motivación de algunos grupos de alumnos, la cual afecta a sus expectativas sobre el rendimiento de dichos grupos.

Debido a su carácter inclusivo, en el centro B en cambio perciben las altas expectativas sobre el logro de todos sus estudiantes como una seña de la institución:

PT: si de algo nos sentimos orgullosos en este colegio es cuando el día de la graduación ves a chicos con adaptaciones curriculares, o a algunos que les hemos tenido que adaptar contenidos graduando su ritmo de aprendizaje, cuando han pasado con la ayuda de todo el mundo, y los ves graduándose en 2º de bachillerato... es más, les presentan a la PAU y aprueban, y es que es el mayor orgullo. La gente buena sale para adelante en cualquier sitio, pero la gente con dificultades a la que les has dado la oportunidad para madurar y aprender a su ritmo...

Profesor de Geografía e Historia: Sus éxitos son nuestros éxitos, y sus fracasos también son cosa nuestra, por eso es un orgullo ver lo que van consiguiendo. Es una alegría y esas son nuestras medallas. (Docentes veteranos, centro B)

Tanto los docentes noveles como veteranos de este centro hacen hincapié en la importancia de esperar el máximo logro de todos sus estudiantes y de darles la oportunidad de conseguir las cosas a su propio ritmo.

Evaluación de programas, profesores y centros

Los procesos de evaluación de programas y profesores en cada uno de los centros se realizan bajo enfoques diferentes. El Equipo Directivo del centro A reporta que en su centro se evalúa todo lo que se hace (programas, planes, actividades, charlas...). Cada vez que se realiza una actividad o programa, el docente encargado tiene que rellenar un formulario de evaluación y entregarlo al Equipo Directivo:

Todo proyecto cuando se acaba se evalúa, todos tienen su parte de evaluación donde se recoge cómo se va a evaluar. Sí que es cierto que a lo mejor en los años anteriores nos costaba mucho a lo mejor recoger toda la información y poder llegar a conclusiones (...) queremos ir añadiendo cosas poco a poco, porque el problema no es que no se evalúe, se evalúa todo, pero se trata de poder sacar conclusiones de esa evaluación y poder transmitírselas al profesorado, y animar a que se sigan implicando. (Jefa de Estudios, centro A)

Sin embargo, tal como puede observarse en el fragmento anterior, admiten que este proceso es susceptible de mejora, ya que se recoge tanta información que es complicado analizarla y poder poner en marcha las mejoras oportunas, por lo que están llevando a cabo una optimización del proceso de recogida de evaluaciones a través de su informatización.

En el centro B se refieren fundamentalmente a dos procesos evaluativos a nivel institucional: las encuestas de satisfacción de alumnos y familias, y las encuestas de desempeño del profesorado. La finalidad de estas últimas es la mejora de la actividad de los docentes en función de sus resultados, aunque desde el Equipo Directivo planean este aspecto como una tarea a medio plazo en el futuro:

Es un proceso que empezó hace 5 o más años, pero yo creo que no se pueden tomar decisiones sobre encuestas de desempeño de los profesores hasta que pasen por lo menos 10 años, para ver los efectos a lo largo del tiempo, yo el año que viene quiero sentarme con los profesores para decirles “esto es lo que opinan los alumnos, esto es lo que piensan las familias, a ver qué piensas tú” y si el

profesor es lo suficientemente empático y cercano para entender que esa es una conversación constructiva pues bien, si no... (Director, centro B)

Por tanto, se percibe una diferencia en el foco principal de la evaluación del centro, que en el A se sitúa en los proyectos y actividades, y en el B se centra más en el desempeño de los docentes, aunque en ambos centros se percibe una voluntad de utilizar los resultados de la evaluación para mejorar el funcionamiento de la institución.

7.3.5. Gestión del tiempo

Puntualidad en entradas y salidas

Sobre este punto no se recogió mucha información, aunque el Equipo Directivo del centro B comentó que el timbre que marca el comienzo y final de cada una de las clases se respeta siempre de manera adecuada y que no hay problemas de puntualidad ni de organización en las entradas y salidas de las aulas.

Organización de tiempos y horarios de los alumnos en actividades extraescolares

Sobre este tema, los docentes del centro A refieren la gran cantidad de actividades que se realizan fuera del horario escolar, hecho que la orientadora del centro atribuye, además de a la implicación de los profesores, al buen funcionamiento del centro:

El hecho de que todo funcione más o menos hace que nos multipliquemos con actividades extraescolares, complementarias, culturales... aquí tenemos un clima de actividades extras que es para darnos un premio a la productividad.
(Orientadora novel, centro A)

7.3.6. Liderazgo y Equipo Directivo

Formación, experiencia, innovación y mejora

Este aspecto fue tratado solamente en el centro B, donde el equipo directivo expresó la importancia de formar a los gestores del centro y expuso el proceso de formación para el liderazgo que se sigue en el centro, que consiste en más de un año y medio de acciones formativas específicas en una organización externa.

Proyectos y objetivos del equipo directivo

En el centro A, tanto los docentes noveles como los veteranos reconocen como uno de los objetivos principales del equipo directivo la mejora de la convivencia, la disciplina y

la resolución de conflictos con los estudiantes de una manera productiva, dialogada y cercana.

Filosofía, implicación y clima del liderazgo

La mayor parte de opiniones sobre filosofía del liderazgo del centro y la percepción de los docentes sobre este tema se recogió en el centro A, en el que los participantes expresaron una filosofía de liderazgo horizontal y compartido, desde los jefes de departamento hasta el propio director:

Profesor de Lengua: Yo soy jefe de departamento y me da la sensación de que en departamento trabajamos todos en todas las cosas, es verdad que el jefe tiene que encargarse de algunas cosas, de hacer la memoria, la programación, (...) pero lo que es el grueso del trabajo en mi departamento lo repartimos todos, a mí no me da la sensación de tener una posición de liderazgo con respecto de mis compañeros.

Profesor de Inglés: Incluso el director del centro, pues no lo ves como un director digamos... una figura jerárquica, es lo más asequible y es el primero que echa una mano, y los demás del quipo pues igual.

Profesor de Lengua: A ver, sabemos que son nuestro equipo directivo, pero luego es un trabajo muy solidario y compartido. (Docentes noveles, centro A)

En el centro B, la conversación sobre el liderazgo surgió fundamentalmente en torno a la toma de decisiones, la cual tanto el Equipo Directivo como los docentes expresaron que en general era compartida y consensuada, aunque a veces es necesario tomar la decisión desde los órganos de liderazgo:

Yo quiero hacer un comentario sobre lo del liderazgo, yo creo que hay veces que hay decisiones que se tienen que tomar desde arriba, y desde abajo no podemos opinar, y otras cosas que nos atañen a veces se hace un claustro o una reunión informativa, nadie te impone el 100% de todas las decisiones, porque al final hay cosas que se tienen que decidir entre todos, aunque la última palabra se tenga arriba. (Profesora de Matemáticas novel, centro B)

7.3.7. Gestión y organización del centro

Organización de la coordinación

El tema sobre el que más se trató en ambos centros en relación con esta subcategoría fue la coordinación entre los docentes. En el centro A hay una opinión generalizada de que los docentes se organizan prioritariamente por departamentos, y la coordinación interdepartamental es menor. Ambos grupos de docentes expresan que uno de los factores que puede afectar a este aspecto son los espacios y tiempos de convivencia de los que disponen, ya que los recreos son cortos y la cultura del centro se caracteriza más por acudir a las salas propias de los departamentos que a la sala general de profesores, que queda algo más retirada. Por otro lado, los docentes veteranos son los que perciben de manera más acusada una actitud individualista por parte del profesorado del centro:

No hay gran equipo... mi departamento es pequeño (...) y entre nosotras bien, pero podría haber más equipo en general, o sea yo reconozco que somos muy individualistas. (Profesora de Alemán veterana, centro A)

Sin embargo, sí se destaca la coordinación entre distintos departamentos a la hora de poner en marcha proyectos concretos.

En el centro B, en cambio, el foco se sitúa fundamentalmente en la coordinación entre los docentes que comparten aula (profesionales de PT y docentes de materia), la cual expresan que es fluida y constante, ya que comparten muchos espacios y tiempos a lo largo del curso. Por otro lado, también describen una buena coordinación entre los profesores que imparten la misma materia en el mismo curso con el objetivo de ajustar los ritmos de cada una de las clases. Esto se debe a la particular configuración de los docentes de materia en este centro:

Profesora de Matemáticas: Cuando hablamos de coordinación, no sólo la hay entre los dos profes que conviven en el aula, por curso también estamos todos coordinados para que vayan al mismo ritmo y sea equitativo.

Profesor de Matemáticas: Es que en matemáticas tenemos un tema de organización, y es que dentro del mismo curso a cada línea [A, B o C] le da un profesor distinto, y eso requiere por nuestra parte la coordinación de objetivos y temas, tenemos que estar en constante diálogo. (Docentes noveles, centro B)

Cuidado del profesorado

Este tema se trató fundamentalmente en el centro A, donde varios docentes relataron sentirse muy bien atendidos y apoyados por el equipo directivo, cuestión respaldada por uno de los miembros del equipo directivo, quien explicó que las puertas de la Jefa de Estudios general siempre están abiertas para los docentes y que esta realiza una buena atención individualizada de sus necesidades.

Estabilidad de la plantilla

La diferente titularidad de ambos centros (público y concertado) hace esperable que la información sobre la estabilidad de la plantilla sea diferente. Al tratarse de un centro público situado en una capital de provincia, el centro A supone el último destino de la mayoría de sus docentes después de haber pasado por otros centros públicos de manera temporal. Por este motivo, la plantilla de este centro es muy estable.

En el centro B, sin embargo, las horas asignadas a los docentes cada año dependen del número de estudiantes que se matriculen y de su perfil (ordinarios, con NEE, etc.), por lo que existe una mayor inestabilidad de la plantilla, y en opinión del Equipo Directivo puede generar algunos conflictos con los docentes:

De ahí es de donde surgen los conflictos, porque cuando hay vacas gordas pues genial, porque todo el mundo tiene horas, pero los años malos se generan conflictos... que es el mayor conflicto con el personal, no el currículum ni la forma de trabajar, sino la estabilidad. Que es comprensible, pero claro dependemos del número de alumnos. (Director, centro B)

Implicación, compromiso, entrega y disponibilidad

Sobre este punto se aprecia poca diferencia entre los centros, ya que, tanto los docentes como los equipos directivos de ambos destacan la gran implicación de todos los miembros de la institución tanto en la docencia ordinaria como en los proyectos del centro, incluso en aquellas cuestiones que requieren una inversión de tiempo fuera del horario laboral (actividades, visitas, coordinación entre docentes que comparten aula en el caso del centro B, etc.):

Tenemos profesores que ya llegan al centro muy muy muy motivados a participar en todo tipo de proyectos, hasta el punto de que tenemos muchos premios de participación, aparte de concursos de todo tipo y olimpiadas de todas

las asignaturas (...) Vamos, que se mueven todos los departamentos, el departamento de Lengua no para, además de hacer convenios con la biblioteca tiene muchísimas cosas, el de Ciencias, el de Filosofía, bueno en todos... (Jefa de Estudios, Centro A)

Pasamos mucho tiempo a parte de las clases fuera, que eso también lo destaco del cole, la implicación que hay por parte de los profesores, porque todas las adaptaciones que hacemos las hacemos fuera del horario laboral. (PT novel, centro B)

7.3.8. Clima y convivencia

Este fue uno de los puntos a los que se hizo una mayor cantidad de referencias a lo largo de las sesiones con los grupos de discusión, lo cual revela la importancia que el clima y la convivencia tienen para ambos centros. Sin embargo, se encuentran algunos matices diferentes en cuanto al enfoque de estos elementos, que se irán viendo a lo largo de las siguientes subcategorías, pero que pueden definirse brevemente para ofrecer contexto para su interpretación. En general, en el centro A el clima y la convivencia, así como las actividades para mejorarlas, se tratan desde un punto de vista institucional, ya que los participantes hablan sobre los planes de convivencia y acogida del alumnado, la coordinadora de convivencia, o de cómo estos elementos son uno de los objetivos prioritarios del Equipo Directivo y de los docentes. En el centro B, sin embargo, se acercan al tema del clima desde una perspectiva más emocional, aludiendo a que es la cercanía entre los docentes, los alumnos y las familias y la calidad humana de la comunidad educativa lo que hace que se genere un buen clima en el centro.

Convivencia

Aunque ambos centros destacan la importancia de la convivencia en el centro y la implicación de todos los miembros de la comunidad educativa en su desarrollo, en el centro A se señala claramente como uno de los pilares institucionales:

Lo de la convivencia no es casual, un clima de convivencia positivo no se consigue de repente porque estamos en un buen barrio, porque llegan alumnos buenos, o los padres son excepcionales. Se consigue porque es el primer objetivo del centro, del equipo directivo, el departamento y los tutores, el que haya un

buen clima de convivencia y los conflictos con los chicos estén resueltos.

(Orientadora novel, centro A)

Uno de los elementos que se señalan en este centro como fundamentales son las acciones de acogimiento que se realizan en 1º de ESO para integrar a los estudiantes en la institución.

En el centro B, la convivencia se caracteriza más por su vinculación con otros temas que se tratarán más adelante, como el clima del centro o la resolución de conflictos, pero no se trata en sí misma.

Clima

En el centro A, la mayoría de las menciones del clima del centro se centran en el ambiente general de tranquilidad y ausencia de problemas de disciplina en el centro, relacionándolos con el bienestar de los profesores y la capacidad para dar sus clases a gusto. En el centro B, el tema central relacionado con el clima es la cercanía de las relaciones establecidas entre docentes y alumnos, cuestión que resaltan los tres grupos de informantes, y que es tema de conversación habitual con familias y estudiantes, sobre todo con aquellos que han tenido experiencias en otros centros y destacan estas relaciones de cercanía como una particularidad del centro.

Actividades para mejorar el clima

Como se ha mencionado anteriormente, una de las actividades fundamentales para mejorar el clima y la convivencia en el centro A es el plan de acogimiento a los estudiantes de nuevo ingreso en 1º de ESO. Este plan comprende acciones de diversa naturaleza, desde contactar con los centros de Educación Primaria de los cuales reciben la mayor parte de su alumnado para tener un primer contacto con los estudiantes, hasta realizar sesiones de bienvenida (tanto en junio como en septiembre antes de comenzar el curso) con estudiantes y familias, en las que se les enseña el centro y se explican las normas de convivencia y la filosofía institucional. La importancia del plan de acogida con respecto a la buena convivencia que existe en el centro fue mencionada por todos los grupos de discusión, y su impacto positivo en el ambiente del centro fue apoyado por todos los participantes.

Resolución de conflictos

En este punto, ambos centros destacan un enfoque de resolución de conflictos basado en el diálogo a la hora de resolver los conflictos con los estudiantes, aunque los procedimientos difieren entre ellos. En el centro A, el procedimiento se organiza en torno a la coordinadora de convivencia y las jefas de estudios, que son las encargadas de tratar con los estudiantes implicados en conflictos. Además, se plantean también acciones de prevención con los alumnos de 1º y 2º de ESO a través de información sobre las normas del centros y acciones de convivencia para reducir la cantidad de conflictos que se producen:

Profesora de Inglés: [cuando hay un conflicto] se remite a los afectados a jefatura de estudios, y jefatura decide si hace falta el apoyo de orientación, o aplicar alguna medida, pero no suele haber ningún problema de comunicación, los jefes de estudio en ese sentido son ejemplares y se resuelve todo de una buena manera.

Profesor de Lengua: Lo resuelven todo a base del diálogo, la cuestión es fomentar el diálogo entre las personas afectadas en ese conflicto, lo cual me parece fundamental.

Profesor de Inglés: Hay una profesora que se dedica a ello.

Profesora de Inglés: Sí, hay una profesora que es la coordinadora de convivencia, que hace la labor de mediación, y hace una función preventiva con los pequeños de 1º y 2º. En cualquier nivel puede intervenir, pero con los pequeños yo creo que hace una labor preventiva fundamental. (Docentes noveles, centro A)

En el centro B han puesto en marcha un protocolo de alumnos mediadores y ayudantes para que sean ellos los que gestionen los conflictos entre los compañeros. Esto no quiere decir que el personal de centro no se implique en la resolución de los conflictos, pero utilizan la figura de los alumnos mediadores como una primera vía para resolverlos, ya que entienden que en algunos casos el hecho de tratar con iguales puede resultar más cercano que tratar directamente con el personal docente. Estos alumnos están formados por psicólogos para ejercer su papel de manera adecuada:

PT1: A la hora de resolver conflictos hay un protocolo, y hay unos alumnos que son alumnos ayudantes, que cuando hay un conflicto entre dos personas y no son

capaces de resolverlo, están los mediadores que se encargan de intentar solucionar el conflicto, porque a veces se soluciona mejor entre iguales que si entra el profesor...

Profesor de Matemáticas: Una puntualización sobre los alumnos mediadores, son alumnos de 3º y 4º que están formados por una persona especializada en psicología, que viene y les da un curso de formación. Luego, en todos los cursos hay alumnos ayudantes, que a falta de mediadores también están ahí para echar una mano.

PT2: Yo creo que los ayudantes están para cosas más del día a día y cuando hay un conflicto más grave se llama a los mediadores. (Docentes noveles, centro B)

7.3.9.Familia y comunidad

Relación con las familias

Los dos centros expresan tener buena relación con las familias, aunque el tema se ha tratado de forma mucho más extensa en el centro A, cuyos profesionales describen sus relaciones con las familias de los estudiantes como positivas y constructivas. Uno de los factores que consideran que puede influir en esta relación es el proceso de acogimiento en 1º de ESO, en el cual también se invita a las familias a acudir al centro y conocer sus espacios y personal, y se les transmiten las normas de funcionamiento de la institución, lo cual aumenta su implicación.

Desde el Equipo Directivo se realiza la observación de que la relación con las familias se desarrolla a título personal más que a través del AMPA, que no tiene mucha actividad en este centro:

Sí, la relación es buena, lo que pasa que las familias es cada una más a nivel individual, el AMPA no es mayoritaria (...), quizá porque están bien atendidos a nivel individual... las puertas están abiertas, vienen y no hace falta respetar el horario de tutorías, entonces eso favorece el clima de entendimiento. (Director, centro A)

En el centro B mencionan que lo que más valoran las familias en las encuestas anuales de satisfacción es la relación de cercanía con los profesores.

Implicación con organizaciones externas

Una de las principales relaciones institucionales mencionada por los profesionales de ambos centros es la Universidad, aunque dicha colaboración se configura de maneras distintas. En el centro A acercan a los estudiantes a la Universidad de Salamanca no sólo a través de las jornadas de puertas abiertas, sino también mediante convenios con algunas facultades (Ciencias, Traducción e Interpretación) para que los alumnos puedan acudir a realizar algunas actividades prácticas. La colaboración del centro B con la Universidad se realiza a través de estudiantes de universidades, usualmente privadas, que acuden al centro a realizar las prácticas de sus respectivas titulaciones, aunque también resaltan su historial de colaboraciones con la Facultad de Educación de la USAL.

Sobre relaciones con otras organizaciones, en el centro A destacan su colaboración con el SACYL para la impartición de cursos de educación sexual a lo largo de la ESO y también con el Ayuntamiento de Salamanca a través de la Fundación Ciudad de Saberes, con la que diversos profesores han realizado actividades relacionadas con sus respectivas asignaturas. Además, también realizan un contacto rutinario con los centros de Educación Primaria de los cuales suelen recibir alumnos, para recibir información de dichos alumnos y poder realizar un primer contacto con ellos al final del 6º curso de primaria:

A final de curso en junio hacemos una entrevista la orientadora y yo a los centros mayoritarios que son [nombre de dos colegios de primaria de la ciudad], estos son los centros adscritos al Instituto, y ahí niño por niño nos dicen cuáles son sus fortalezas y debilidades, el amparo familiar que tienen, que eso es importantísimo, etcétera. Tenemos unos ítems sobre cómo son los niños intelectualmente, cómo son emocionalmente y cuál es su amparo familiar, porque esto es importantísimo a la hora de medrar. (Jefa de Estudios de ESO, centro A)

7.4. RESUMEN

Después de haber descrito de manera pormenorizada los resultados de los grupos de discusión con respecto a los temas principales de interés, se presenta una tabla a modo de resumen con las principales características definitorias de los dos centros en cada una de las categorías estudiadas (tabla 7.4).

Tabla 7.4. Resumen de los resultados de la fase cualitativa

Categoría	Centro A	Centro B
Planes, proyectos y formación	Proyectos relacionados con lo académico Alta implicación en la formación Innovación basada en proyectos externos o metodologías	Proyectos relacionados con la inclusión Actitudes menos entusiastas o neutras hacia la formación Innovación interna basada en creatividad de los profesores
Metodologías	Libertad metodológica Uso rutinario de cañones y ordenadores en el aula	Trabajo en grupo como línea metodológica general Uso puntual de las TIC Trabajo de la competencia emocional
Atención a la diversidad y seguimiento	Mayor desarrollo de la orientación Programas para alumnos con bajo rendimiento	Alumnado NEE marca la organización del centro PT en clase Seguimiento individualizado del alumnado
Evaluación	Expectativas más ligadas a motivación de los estudiantes Evaluación centrada en mejora de proyectos y actividades	Expectativas altas para todos Evaluación centrada en satisfacción de familias y desempeño de profesores
Liderazgo y equipo directivo	Liderazgo compartido, relaciones horizontales	Algunas decisiones consensuadas, relaciones más verticales
Gestión y organización	Coordinación entre docentes organizada por departamentos Estabilidad de la plantilla	Coordinación entre docentes organizada por compartir aula o asignatura en el mismo curso Plantilla inestable dependiente de la matrícula de alumnado
Clima y convivencia	Convivencia como objetivo fundamental del centro Resolución de conflictos basada en el diálogo con docentes Plan de acogimiento en 1º de ESO	Convivencia basada en relaciones de cercanía (emocional) Resolución de conflictos mediada por alumnos
Familia y comunidad	Relaciones positivas y constructivas con las familias, plan de acogimiento, poca fuerza del AMPA Numerosas colaboraciones externas para hacer actividades	Colaboraciones externas más centradas en tareas de gestión o participación de estudiantes universitarios en el centro a través de prácticum

Como se puede observar en la tabla 7.4, aunque los resultados obtenidos no pueden relacionarse de manera clara con la eficacia de los centros, sí se han encontrado algunas cuestiones que parecen ir en la línea de los estudios existentes, como la alta implicación en la formación, la promoción de la convivencia, la estabilidad de la plantilla y las buenas relaciones con las familias en el centro de alta eficacia, o unos niveles menores de colaboración con organizaciones externas, estabilidad de la plantilla, actitudes hacia la formación y distribución del liderazgo en el centro de baja eficacia.

En definitiva, nos encontramos ante dos centros con unos principios y una filosofía propia que difieren entre sí. Cada uno tiene un pilar principal muy bien definido; en el centro A la preocupación fundamental de todo el personal es la convivencia, cuestión que promueven a través del comprensivo programa de acogimiento a los alumnos de nuevo ingreso y sus familias, los protocolos de resolución de conflictos y la relación cercana de todo el Equipo Directivo y docente con los estudiantes. Por otro lado, el centro B está completamente volcado en la inclusión de estudiantes con necesidades específicas, por lo que sus esfuerzos organizativos y pedagógicos se centran en introducir la figura del profesional de apoyo en casi todas las aulas y promover el trabajo en grupo para que todos los estudiantes puedan expresar y aportar lo que saben.

Además de estas dos cuestiones, se perciben otras diferencias que se reflejan en varias categorías, y que también podrían ser indicativas de la cultura diferente de ambos centros. Una de ellas es la apertura del centro a ideas y organizaciones exteriores, la cual se percibe en mayor medida en el centro A. En este centro, las actitudes hacia la formación externa son altamente positivas, y se reportan muchas más colaboraciones con organizaciones externas para realizar actividades fuera del centro, o bien motivadas por cuestiones externas (proyectos relacionados con iniciativas como olimpiadas, o programas de emprendimiento, por ejemplo). En el centro B, sin embargo, se percibe una mayor tendencia hacia la organización interna, ejemplificada en la práctica generalizada de basar los proyectos y las innovaciones en ideas propias de los docentes, y no en corrientes o metodologías externas, las actitudes indiferentes o ligeramente negativas hacia la formación externa, o las relaciones establecidas con organizaciones externas, que se suceden en mayor medida dentro del centro (acogida de alumnos de prácticas procedentes de la universidad, por ejemplo).

Por otro lado, la diferencia de la titularidad de los centros queda patente en cuestiones como la estabilidad de la plantilla, que en el centro público es mucho mayor que en el

concertado. Otra de las cuestiones que podrían estar mediadas por la titularidad de los centros, aunque su relación puede ser menos directa que la estabilidad de la plantilla, es la finalidad de los procesos de evaluación a nivel de centro. En el centro A, el foco principal de la evaluación son los proyectos y las actividades, con el objetivo de mejorarlas, mientras que en el centro B el objetivo principal de la evaluación es conocer la satisfacción de los usuarios (alumnos y familias) y el desempeño de los trabajadores.

En conclusión, se puede generar un perfil diferenciado de cada uno de estos centros, pero la relación de los elementos identificados con el nivel de eficacia del centro no es clara. Uno de los motivos fundamentales de este hecho es la reducida muestra con la que se cuenta, la cual no permite extraer conclusiones generales de ningún tipo. Por otro lado, el hecho de haber escogido Castilla y León como Comunidad Autónoma de referencia tiene como consecuencia que el rango de residuos de los centros es más reducido que en la muestra general, es decir, existen diferencias menos extremas entre los centros de alta y baja eficacia. Por este motivo, quizá esa diferencia menor en los residuos suponga mayores similitudes entre los centros, y por tanto conclusiones menos claras acerca de las cuestiones que los caracterizan de manera diferencial.



8. Discusión y conclusiones

Discusión y conclusiones

El objetivo general establecido para este trabajo de investigación fue la detección de aquellas prácticas escolares que pudiesen fomentar (o dificultar) la eficacia escolar en los centros educativos de Educación Secundaria en Castilla y León. La definición de eficacia escolar tomada como modelo para este estudio es la formulada por Murillo (2003), en la que se establecen tres pilares fundamentales: la equidad, la educación integral, y el valor añadido de los centros con respecto al rendimiento.

Para dar cumplimiento al objetivo general de la tesis, así como para atender a los tres elementos centrales de la eficacia escolar, se determinaron tres objetivos específicos que fueron cumpliéndose a través de la aplicación de diversas técnicas de investigación. En primer lugar, el estudio de los factores de contexto relacionados con el rendimiento de los estudiantes en las tres competencias evaluadas por PISA (matemáticas, comprensión lectora y ciencias) se realizó a través de la aplicación de modelos multinivel, encontrando que las variables que más impacto tienen en el rendimiento son el nivel socioeconómico (a nivel de estudiante y de centro), el género, el curso y la repetición, el número de cambios de centro a lo largo de la escolarización del estudiante, y la condición de inmigrante de 1ª generación (sólo en el caso de matemáticas y ciencias). Este análisis se relaciona simultáneamente con los tres elementos de la eficacia, ya que examina de qué manera los factores socioeconómicos y demográficos de los estudiantes afectan a su rendimiento (equidad), se explora el rendimiento en tres competencias de naturaleza diferente (educación integral), y se determina a través de los residuos de segundo nivel cuál es el aporte específico de los centros al aprendizaje de los estudiantes (valor añadido).

El segundo objetivo específico consistió en seleccionar aquellos centros que presentasen una alta o baja eficacia, lo cual se realizó a través del estudio de los residuos de segundo nivel generados por los modelos multinivel.

El tercer objetivo, relacionado fundamentalmente con el elemento del valor añadido, se cumplió a través de la aplicación de dos técnicas de investigación. Por un lado, se trabajó con las variables de proceso ofrecidas por los cuestionarios de contexto de PISA para estudiar su relación con la eficacia de los centros. Este análisis se realizó primero con variables de centro (liderazgo, participación de los docentes, etc.), sin embargo, no arrojó ninguna relación significativa con la variable indicativa de la eficacia. Después se replicó con medidas de variables de proceso de los estudiantes agregadas a nivel de centro

(adaptación de la instrucción, clima de disciplina, apoyo del profesorado, etc.), encontrando algunas variables con relación significativa (positiva o negativa) con la eficacia de los centros. Para complementar este análisis, se realizó también un estudio cualitativo de casos en dos centros de la Comunidad Autónoma de Castilla y León, en los que se entrevistó a docentes y equipo directivo para recoger su visión de aquellos factores organizativos y pedagógicos del centro que pudiesen estar influyendo en el rendimiento de los estudiantes y en la eficacia del centro.

En este capítulo se presentan las principales reflexiones derivadas tanto de los resultados obtenidos en ambos estudios como de las revisiones de la literatura presentadas en la parte teórica de este documento. Asimismo, se discuten las limitaciones teóricas y metodológicas que presentan los estudios, y las posibles líneas de futuro que pueden derivarse como resultados de estas dificultades. Todos estos elementos se estructurarán alrededor de los tres pilares de la eficacia anteriormente mencionados.

8.1. EQUIDAD

El concepto de equidad en la educación supone que todos los estudiantes de un sistema educativo disponen de las mismas oportunidades de aprendizaje independientemente de su género, nivel socioeconómico familiar, estatus migratorio, etc. La equidad no significa que todos los estudiantes obtengan los mismos niveles de rendimiento, sino que el rendimiento que obtengan no esté relacionado con sus circunstancias demográficas o socioeconómicas, sobre las cuales los estudiantes no tienen control alguno (OECD, 2018).

Por tanto, para examinar la equidad de un sistema educativo, es necesario analizar hasta qué punto este tipo de características están asociadas con el rendimiento de sus estudiantes, entendiendo que el objetivo es que su efecto sea lo más pequeño posible para poder caracterizar a un sistema como equitativo. Para llevar a cabo esta tarea, la prueba PISA proporciona ciertas características de identificación de los estudiantes que permiten esclarecer cuál es el efecto que tienen sobre su nivel competencial. A lo largo del capítulo 6 se ha visto cómo las variables de nivel socioeconómico, curso, género, estatus migratorio y número de cambios de centro a lo largo del historial académico tienen un peso sustancial en el desempeño de los estudiantes en las tres competencias principales evaluadas por PISA. A continuación se desgranarán estos resultados, comparándolos con la literatura existente y explorando sus implicaciones.

8.1.1. Nivel socioeconómico

El impacto del nivel socioeconómico de la familia sobre el rendimiento de los estudiantes ha sido ampliamente estudiado y confirmado desde los resultados del Informe Coleman (1966), tanto en el contexto español (Cordero, Pedraja y Simancas, 2015; Gil-Flores & García-Gómez, 2015, Tourón et al., 2018) como en el internacional (Acosta & Hsu, 2014a; Cheung et al., 2014; Hwang, Choi, Bae & Shin, 2018; Willms, 2010). Entre las variables introducidas tanto a nivel individual como agregadas a nivel de centro en los modelos multinivel (ESCS, repetición, estatus migratorio, género), el nivel socioeconómico es el único que aparece como significativo en ambos niveles consistentemente en los tres modelos. Este hecho revela la gran importancia que tiene en el rendimiento de un estudiante no sólo el nivel socioeconómico de su familia, sino también el nivel de las familias de sus compañeros/as (nivel promedio del centro).

Las implicaciones de esta información de cara a la equidad de los sistemas educativos se concretarían en dos cuestiones, una relacionada con las prácticas educativas de los centros y otra con las políticas a nivel de sistema. Por un lado, es importante que los centros compensen el efecto del nivel socioeconómico de sus estudiantes a través de políticas de apoyo o compensación, no sólo asegurando un trato igualitario a todos los estudiantes, sino implementando prácticas de discriminación positiva que brinden un mayor apoyo a aquellos alumnos procedentes de entornos desfavorecidos, promoviendo así una verdadera igualdad de oportunidades educativas (Bolívar, 2005).

Sin embargo, también debe haber políticas a nivel de sistema educativo que aseguren que los centros son socialmente diversos, ya que la segregación de los centros educativos por nivel socioeconómico beneficia únicamente a los estudiantes de los centros con mayor ESCS y aumenta notablemente el impacto que esta variable tiene en el rendimiento, contribuyendo así a una mayor brecha entre los estudiantes desfavorecidos y aquellos que tienen niveles más altos de renta familiar (OECD, 2019). Una de las políticas que se relaciona habitualmente con la segregación social de los estudiantes es la libertad de elección de centro por parte de las familias, la cual constituye un tema ampliamente debatido y estudiado. Algunos de los partidarios de la libre elección de escuela argumentan que la introducción del factor de competición por los alumnos en un sistema de elección libre hará que los centros se motiven para innovar y mejorar para atraer a más alumnos, mejorando de esta manera la calidad del sistema educativo (Ball, 1993). Sin embargo, las prácticas de libre elección han demostrado en algunos estudios su relación

con una mayor segregación, y por tanto una mayor inequidad en el sistema educativo (Windle, 2016). Aquellos países en los que existe un mayor porcentaje de estudiantes que asisten a la escuela en función de su residencia muestran una mayor distribución de los estudiantes con nivel socioeconómico bajo, mientras que en los países en los que el lugar de residencia no es un criterio de acceso a la escuela, los estudiantes de ESCS bajo tienen a concentrarse en los mismos centros educativos (Givord, 2019).

Otras formas de segregación, como la separación de estudiantes en función de su rendimiento académico en distintos tipos de clase o de escuela desde edades tempranas (*early tracking*), también amplían la brecha de rendimiento entre los estudiantes en función de su nivel socioeconómico, aumentando el impacto que tiene esta variable sobre el desarrollo cognitivo de los estudiantes (Le Donne, 2014; Van de Werfhorst & Mijs, 2010).

Además de asegurar una composición heterogénea dentro los centros educativos, otras acciones que promueven la equidad educativa para los alumnos de entornos desfavorecidos son también aquellas cuestiones básicas que caracterizan la filosofía de un sistema educativo con vocación equitativa: comprensividad, gratuidad, currículum común, y centros y profesores formalmente equivalentes (Bolívar, 2005). Todas estas medidas se encaminan a garantizar el acceso a la educación en iguales condiciones para todos los estudiantes, lo cual, unido a las medidas de compensación para alumnos desfavorecidos que pueden implementarse en los centros, constituyen mecanismos sistémicos encaminados a garantizar el éxito educativo de todos los estudiantes de una manera equitativa.

8.1.2. **Género**

El género es otra de las variables que habitualmente se estudia en investigaciones sobre el impacto de factores contextuales en el rendimiento de los estudiantes. Aunque el efecto de esta variable difiere en función del país y de la competencia evaluada, en general los estudiantes varones suelen mostrar niveles competenciales más altos en matemáticas y ciencias, mientras que sus compañeras demuestran una mayor competencia en comprensión lectora. En el caso del presente estudio, esta ventaja es muy inferior a la presentada por los alumnos en las materias de la rama científica, por lo que nos centraremos en estas últimas.

Muchos estudios recientes, centrados fundamentalmente en las materias científicas (matemáticas y ciencias) reportan un descenso en la brecha de rendimiento en función del género (OECD, 2016; Lindberg, Hyde, Petersen & Linn, 2010; Zhu, Kaiser & Cai, 2018), lo cual apunta a que estas diferencias no se deben a unas habilidades innatas diferenciales entre alumnos y alumnas, sino a factores relacionados con la socialización de los estudiantes (familiares, educativos, roles de género etc.) que favorecen el aprendizaje de los varones (o que interfieren en el desarrollo académico y personal de las mujeres). De esta forma, esta tendencia al cierre de la brecha demuestra que los esfuerzos por reducir la desigualdad de género en el aprendizaje de las ciencias y las matemáticas desde los centros, los sistemas educativos y las sociedades en general están logrando sus objetivos e indica la necesidad de seguir trabajando en este aspecto.

Sin embargo, algunas investigaciones apuntan a que el hecho de que las puntuaciones medias de hombres y mujeres se estén acercando no quiere decir que no sigan existiendo desigualdades. El estudio de la distribución de los géneros por niveles de rendimiento en las pruebas a gran escala sugiere que, aunque las diferencias en puntuaciones medias se reduzcan, la desigualdad en los niveles extremos de rendimiento sigue siendo patente, habiendo una mayor proporción de hombres que de mujeres en los niveles más altos de rendimiento²² (Stoet & Geary, 2013; Zhou, Fan, Wei & Tai, 2017). Además, las alumnas también muestran niveles más bajos que sus compañeros en cuestiones como el interés por la ciencia y las matemáticas o la autoeficacia percibida en estas materias

Por otro lado, los informes de la OCDE (2016) indican que, aún para el mismo nivel de rendimiento en ciencias, las expectativas de trabajar en un campo relacionado con la ciencia son significativamente mayores entre los hombres. Por este motivo, se hace necesario invertir esfuerzos para reducir la brecha de género en dos frentes diferentes, aunque relacionados: mejorar el rendimiento de las alumnas en ciencias y matemáticas y motivarlas a relacionarse con la ciencia y generar actitudes positivas hacia ella.

Un reciente informe de la UNESCO (2017) sobre la situación de las mujeres con respecto a la educación matemática y científica proporciona una comprensiva explicación de los factores que afectan su rendimiento e interés más bajo por estas materias, así como actuaciones para mejorarlas desde el sistema educativo. Este informe comienza

²² Lo mismo sucede en los niveles más bajos de rendimiento, ya que en general las alumnas muestran una menor dispersión en sus puntuaciones.

descartando las teorías de las diferencias biológicas entre los sexos como base para la disparidad de rendimiento existente, centrándose en otro tipo de cuestiones, como la influencia de la sociedad, el entorno cercano, el centro educativo, o la propia psicología en el rendimiento e intereses de las mujeres por los campos STEM (figura 8.1).



Figura 8.1. Condicionantes del desempeño y los intereses de las alumnas por materias STEM. Elaboración propia a partir de UNESCO (2017).

A partir de estos factores, se pueden realizar una serie de recomendaciones para la mejora de las condiciones en las cuales las mujeres se relacionan con las matemáticas y las ciencias. En el ámbito de interés de esta tesis, que es el centro educativo, se realizan una serie de recomendaciones que podrían servir para aumentar la implicación y el interés de las alumnas en las materias científicas. Cuestiones como garantizar la contratación de profesoras de ciencias, capacitar a los docentes de ciencias para que comprendan los factores que influyen en el aprendizaje de las alumnas y puedan trabajar sobre ellos, mejorar las prácticas educativas (lenguaje y enseñanza inclusivos, más aprendizajes prácticos, adaptación de la instrucción a intereses de las alumnas), o facilitar el acceso a una orientación académica y profesional sensible a las necesidades de las alumnas se proponen con el objetivo de mejorar la inclusión de las mujeres en los ámbitos científicos (UNESCO, 2017).

La relevancia y el impacto de estas medidas trascienden el rendimiento académico de las alumnas y se manifiestan también en la composición del cuerpo de trabajadores y trabajadoras del ámbito STEM, en el que existe una patente falta de mujeres (Zhou et al., 2017). Por este motivo, el hecho de que un sistema educativo invierta esfuerzos en generar ambientes de aprendizaje de las ciencias que inviten a las mujeres a participar podría redundar también en un mayor acceso de trabajadoras a este tipo de profesiones.

8.1.3. Estatus migratorio

Otra de las posibles fuentes de desigualdad en el rendimiento de los estudiantes que tiene que ver con la equidad de los sistemas educativos es el estatus migratorio del alumnado. Los resultados generales de PISA 2015 apuntan tres cuestiones fundamentales: los estudiantes de origen inmigrante tienen más del doble de posibilidades de estar por debajo del nivel 2 (de 6) de rendimiento en ciencias, los inmigrantes de 2ª generación tienen mejores puntuaciones que los de 1ª generación, pero significativamente peores que sus compañeros nativos, y el hecho de asistir a un centro educativo con un porcentaje elevado de inmigrantes no muestra impacto en el rendimiento una vez se ha controlado el efecto del nivel socioeconómico medio del centro (OECD, 2016). Los modelos multinivel realizados para el presente trabajo de investigación apoyan la primera y la tercera cuestión plenamente, ya que la variable de inmigración a nivel estudiante aparece en dos de los tres modelos (ciencias y matemáticas) como significativa, pero su agregado a nivel de centro no aparece en ninguno de los tres. Sin embargo, en España no se cumple la segunda cuestión, ya que el rendimiento de los inmigrantes de 2ª generación no muestra diferencias significativas con el de los estudiantes nativos.

Resultaría conveniente explorar de manera más profunda las características del sistema educativo español que contribuyen a la integración educativa de los estudiantes de 2ª generación, y estudiar cuáles son las necesidades específicas de los inmigrantes de 1ª generación a las que no se está atendiendo de manera adecuada.

8.1.4. Repetición de curso

El hecho de haber repetido curso aparece consistentemente como una de las variables contextuales que más influyen sobre el nivel de los alumnos en las tres competencias estudiadas. Aunque no puede establecerse la direccionalidad de esta relación (es decir, si los alumnos con niveles competenciales más bajos son los que repiten, o si el hecho de repetir un curso es lo que genera niveles competenciales más bajos), no puede negarse su

relevancia, ya que en la muestra española hay una alta prevalencia de la repetición de curso (aproximadamente un 30% de los estudiantes participantes habrían repetido un curso).

Con respecto a este punto, cabría también cuestionar la equivalencia del nivel competencial de los estudiantes con su rendimiento académico ya que, como se menciona en la literatura (Carabaña, 2015, 2016; Jornet, 2016), PISA no está basada en los currículos nacionales sino en constructos transversales desarrollados en paralelo al conocimiento curricular. Por tanto, si como postulan estos autores la adquisición de competencias está menos ligada a la instrucción académica que el rendimiento, su relación con la repetición de curso cobraría más relevancia a la hora de interpretar el impacto de la repetición de curso en el desarrollo de los estudiantes y la equidad del sistema educativo.

Independientemente de esta distinción, la investigación educativa ha demostrado en repetidas ocasiones que la repetición de curso no resulta una medida efectiva. Los estudiantes que repiten no muestran una mejora en su rendimiento, e incluso presentan un peor rendimiento que el que habrían obtenido si hubiesen pasado de curso (Jimerson, 2001; Fruehwirth, Navarro and Takahashi, 2016). Por otro lado, los datos también apuntan a una desigualdad en la aplicación de medidas de repetición de curso en función de las variables de contexto de los estudiantes (mayor porcentaje de repetidores entre alumnos desfavorecidos e inmigrantes) (OECD, 2016), lo cual constituye otro punto a favor de su revisión.

Esta información promueve una reflexión entorno a los motivos de prevalencia de las medidas de repetición de curso en nuestro país, así como una consideración de la pertinencia de otro tipo de medidas más beneficiosas para los estudiantes de bajo rendimiento, como acciones de detección precoz y apoyo para este tipo de alumnado (Choi, Gil, Mediavilla & Valbuena, 2016). Los datos relativos al porcentaje de alumnos repetidores ofrecidos por PISA parecen mostrar una tendencia positiva en las políticas educativas a nivel internacional en este sentido, ya que este porcentaje ha disminuido significativamente en la mayor parte de sistemas educativos entre las aplicaciones de 2009 y 2015 (OECD, 2016), aunque sería conveniente investigar si la repetición de curso se está sustituyendo por otro tipo de medidas de apoyo, y en ese caso, cuál es su eficacia.

8.1.5. Cambios de centro a lo largo del historial académico

El número de cambios de centro que un estudiante ha experimentado a lo largo de su escolarización (siendo estos cambios siempre dentro de la misma etapa educativa, no entre etapas) es una variable que ha resultado significativa en los tres modelos explicativos del rendimiento calculados en el capítulo 6. Sin embargo, el impacto de esta variable sobre el rendimiento de los estudiantes en pruebas estandarizadas ha sido mucho menos estudiado que las variables anteriores. En su estudio con los datos españoles de ciencias en PISA 2015, Tourón et al. (2018) también encuentran una diferencia significativa en función del número de cambios de centro que han realizado los estudiantes, siendo este número mayor entre los estudiantes con peor rendimiento.

Debido a la falta de investigaciones centradas en las circunstancias por las que los estudiantes cambian de centro y su relación con el rendimiento académico, no puede establecerse ni siquiera una dirección de causalidad en la correlación entre estas dos variables, es decir, no hay indicios que señalen si los alumnos obtienen peor rendimiento por cambiar de centro o cambian de centro a causa de su bajo rendimiento (o a causa de otros motivos que a su vez puedan influir en el rendimiento). Por tanto, tampoco es posible aventurar posibles prácticas educativas que promuevan el ajuste escolar de los estudiantes que cambian de centro, ya que en función de la relación que se establezca entre las dos variables, las medidas habrían de ser muy diferentes (mejorar los planes de acogida de nuevo alumnado en el primer caso expuesto, o establecer medidas de retención y apoyo al alumnado con bajo rendimiento en el segundo caso, por ejemplo).

Lo que queda patente es que esta variable indica una línea de profundización en la investigación sobre el rendimiento y la eficacia escolar.

8.1.6. Limitaciones

Una de las limitaciones principales con respecto a los análisis de los factores contextuales que afectan al rendimiento de los estudiantes se descubrió gracias a los grupos de discusión realizados en los centros educativos de alta y baja eficacia durante la fase de indagación cualitativa. La problemática descubierta está relacionada con la participación de alumnos con necesidades educativas especiales (NEE) en las pruebas estandarizadas, unida a la falta de transparencia de los documentos y bases de datos de PISA a este respecto.

Una de las bases fundamentales para la aplicación de pruebas estandarizadas a gran escala es la representatividad de la muestra en función de las características de identificación de los participantes, las cuales deben estar debidamente informadas en la base de datos junto con la información del rendimiento. En PISA, esta representatividad se garantiza a través de los procesos de muestreo y de los pesos muestrales asignados a los estudiantes (OECD, 2017b).

Sin embargo, en lo que respecta a la participación de estudiantes con NEE, la documentación de PISA no es del todo transparente, y se hizo necesario acudir a otras fuentes para poder comprender el alcance de la participación de estos alumnos en la prueba, y sus consecuencias en la configuración de las bases de datos y su análisis (para más información sobre este punto, ver sección 5.3.5).

Tras las consultas a la literatura y a personal experto en la materia, la conclusión es clara: en España los alumnos con NEE diagnosticadas son excluidos de participar en la prueba, por lo que no existe una base de datos separada con sus resultados, como sucede en otros países. Sin embargo, hay algunos alumnos sin diagnóstico oficial que son percibidos como alumnos con necesidades por parte de su centro y, si son seleccionados para participar en la prueba, deben hacerlo, incluyéndose sus datos en la base de datos general. En algunos centros con un alto porcentaje de alumnos de inclusión, como en el centro B de la fase cualitativa, este hecho puede suponer una disminución en sus puntuaciones promedio que es imposible de controlar, ya que estos estudiantes no se encuentran identificados a través de ninguna variable.

La información descubierta a través de la entrevista con el centro B y las consultas realizadas con expertos destacan la necesidad de mejorar sustancialmente la transparencia de los documentos de PISA, ya que todos estos datos deberían ser claros y accesibles para cualquier persona interesada. Por otro lado, también promueve una reflexión acerca del tratamiento de los alumnos con necesidades educativas (diagnosticadas o no) en la aplicación de la prueba. La existencia de una base de datos separada y privada, sólo disponible bajo demanda y no descrita en la documentación oficial, resulta un problema tanto a nivel de transparencia como a nivel de equidad. En la base de datos general se informa de otras variables que pueden influir en el rendimiento, como el género, el nivel socioeconómico o el estatus migratorio, permitiendo así un análisis y seguimiento de su evolución a lo largo de los años. Sin embargo, la condición de estudiante con necesidades educativas especiales se trata de manera separada y, aunque se proporcione a los

investigadores bajo demanda, esto genera un número mucho más bajo de estudios que tengan en cuenta esta variable (ver capítulo 2).

Por otro lado, se cuestiona también la pertinencia de la definición estricta de NEE realizada por la OCDE, ya que esta permite la participación de estudiantes que pueden encontrarse con dificultades a la hora de participar en la prueba, y que sin embargo no son identificados de ninguna manera en la base de datos.

Aunque el estudio del rendimiento de los estudiantes NEE no es el objetivo de este trabajo de investigación, la posibilidad de considerar esta variable en el análisis habría mejorado el ajuste de la medida de eficacia de los centros, y habría proporcionado información sobre otra de las facetas fundamentales en la equidad del sistema educativo, por lo que no disponer de ella se considera una limitación relevante.

8.1.7. Líneas futuras de investigación

El estudio del impacto de los factores de contexto sobre el rendimiento de los estudiantes se ha realizado en esta tesis a través de modelos de intercepto aleatorio y pendientes fijas (ver sección 6.1), ya que el objetivo era conocer el impacto promedio de estos factores en el rendimiento de los estudiantes españoles. Sin embargo, el uso de modelos de pendientes aleatorias permitiría un estudio más profundo de dicho impacto a través de la exploración de las pendientes de los centros. Esta corriente, denominada en la literatura como eficacia diferencial (Thomas, Sammons, Mortimore & Smees, 1997), se centra en el estudio del rendimiento diferencial de distintos grupos de alumnos en función de diversas variables de agrupación (género, nivel socioeconómico, estatus migratorio, etc.). Así, a través de este enfoque de investigación se pueden identificar centros cuya brecha de rendimiento en función de estas características sea significativamente pequeña (o grande) e indagar qué características escolares pueden estar relacionadas con este hecho. Aunque existen algunos estudios sobre los efectos aleatorios en función de los centros, estos tienden a estar más centrados en el análisis de la magnitud de las diferencias en la eficacia de los centros en función de variables de interés, pero no sobre qué variables pueden estar influyendo en dichas diferencias (Reynolds et al., 2014), por lo que se estima como una línea de trabajo innovadora que puede generar resultados de relevancia para la disciplina de la IEE.

Por otro lado, se considera también necesario profundizar en el estudio de algunas de las variables contextuales que se encuentran menos analizadas, como el número de

cambios de centro, o cuyos resultados plantean otras preguntas de investigación, como el estatus migratorio. Este último apunta a la existencia de una influencia negativa de la condición de inmigrante de primera generación, sin embargo, no ocurre lo mismo con los inmigrantes de segunda generación. La falta de variables relacionadas con las prácticas educativas – a nivel de centro o de sistema educativo – dirigidas a la atención de este tipo de alumnado supone la necesidad de acudir a otro tipo de fuentes de información para abordar los motivos de esta diferencia.

8.2. EDUCACIÓN INTEGRAL

El concepto de educación integral supone una visión holística de la formación del estudiante en la que se tiene en cuenta no sólo su aprendizaje académico, sino también su adquisición de valores, el desarrollo de su autoestima o su formación emocional, cuestiones que llevan asociadas sus propios resultados de aprendizaje.

La multidimensionalidad de estos resultados de aprendizaje de los estudiantes se aborda en este estudio a través del uso de las tres competencias principales evaluadas por PISA con un doble objetivo: comprobar la estabilidad de los efectos del contexto sobre el rendimiento en las tres competencias, y generar una medida más sólida de la eficacia como base para la selección de centros de alta y baja eficacia. Las competencias matemática, lectora y científica evalúan aspectos diferentes del desarrollo académico de los estudiantes y se relacionan con procesos cognitivos distintos, por lo que el estudio conjunto de las tres habilidades puede ofrecer una idea más completa del nivel de aprendizajes que están adquiriendo los estudiantes de un sistema educativo. Además, las relaciones de los factores de contexto con una de las variables de resultado no han de tomarse como una representación fidedigna de la relación establecida con las demás variables de resultado (Tsai, Smith & Hauser, 2017).

A través del estudio simultáneo de las tres competencias se ha podido constatar la influencia constante de algunas de las variables contextuales en el rendimiento de los alumnos, como el ESCS (individual y escolar), el curso, la repetición y el número de cambios de centro realizados dentro de una misma etapa escolar, las cuales surgieron como significativas en los tres modelos. Sin embargo, otras variables muestran un comportamiento diferencial en función de la variable criterio de estudio. El género, por ejemplo, muestra una ventaja sustancial a favor de los estudiantes varones tanto en ciencias como en matemáticas, mientras que en la competencia lectora las alumnas tienen

ventaja sobre sus compañeros, aunque su magnitud es mucho menor. El hecho de que el idioma hablado en el hogar sea diferente de la lengua vehicular en la que el alumno estudia resulta significativo sólo en el modelo explicativo de la comprensión lectora, cuestión que cobra sentido teniendo en cuenta el fuerte componente lingüístico de ambas variables.

Conocer el perfil diferenciado de cada una de las competencias en función de las variables contextuales que más influyen en su desarrollo constituye una base fundamental a la hora de establecer prácticas y políticas que fomenten una educación integral eficaz para todo el alumnado.

8.2.1. Limitaciones

Una vez discutido el impacto positivo que el estudio de las tres competencias estudiadas por PISA puede tener en la práctica educativa, cabe destacar que todas ellas abordan sin distinción un dominio cognitivo o académico del desarrollo del alumnado, obviando todas aquellas cuestiones que tienen que ver con la adquisición de valores o el desarrollo de resultados no cognitivos.

A pesar de que la definición de eficacia escolar utilizada para este trabajo de investigación concibe el desarrollo integral como una de las bases de la educación, la inclusión de factores de producto o resultado no cognitivos (autorregulación, desarrollo emocional, autoconcepto, etc.) es menos común que la utilización del rendimiento académico como medida del resultado del sistema educativo. Sin embargo, como apuntan algunos autores (Creemers & Kyriakides, 2008; Murillo, 2007), el desarrollo cognitivo académico no es (o no debería ser) el único objetivo de los sistemas educativos, quienes deben comprometerse con el desarrollo psico-socio-educativo de todos los estudiantes (Jornet, 2016), cuestión que evidentemente trasciende los resultados puramente académicos. Por ello, para poder evaluar esta característica, los estudios de eficacia escolar deberían incluir variables de tipo no cognitivo.

La preferencia por las variables de resultado de corte cognitivo patente en la literatura empírica sobre eficacia escolar plantea una reflexión sobre la naturaleza misma de la disciplina. La búsqueda del desarrollo integral se encuentra tanto en la base de los sistemas educativos como en la de la propia disciplina, atendiendo tanto a las definiciones (Creemers & Kyriakides, 2008; Murillo, 2005) como a los modelos integrados que han tratado de explicar las relaciones entre los distintos elementos educativos. Estos modelos sitúan como resultados deseables del sistema educativo cuestiones no cognitivas como

las competencias metacognitivas (Creemers, 1994), el desarrollo afectivo (Creemers & Kyriakides, 2008), la satisfacción de los estudiantes (Muñoz-Repiso et al., 1995) o los logros sociales y afectivos (Murillo 2007).

Tomando estas cuestiones como referencia, cabe evaluar la deriva mostrada por la disciplina de la IEE en las últimas décadas, en las que ha aumentado el estudio de las variables de rendimiento (ver capítulo 2), y reflexionar sobre las posibles causas de esta elección de variables de interés. Quizá la proliferación de datos procedentes de evaluaciones a gran escala como PISA, cuya concepción fundamental de la calidad de un sistema educativo es la excelencia (Jornet, 2016), ha podido desplazar el interés de parte de la comunidad científica desde enfoques centrados en la equidad y la educación integral hacia una visión más orientada hacia el rendimiento y la excelencia.

Se comprende el atractivo de las variables de rendimiento ofrecidas por las evaluaciones a gran escala debido a su mayor fiabilidad y validez, mientras que la medida de los resultados no cognitivos se sostiene en autoinformes que pueden presentar desventajas metodológicas (Chi et al., 2018; Petko, et al., 2016; Sandilands, et al., 2014; Tourón et al., 2018). Sin embargo, aunque la excelencia de los sistemas educativos es un tema relevante, no debe sustituir por completo el objetivo social más amplio que representa el desarrollo de la educación integral para todos los alumnos.

8.2.2. Líneas futuras de investigación

Una de las futuras líneas de investigación relacionadas con la formación integral del alumnado se deriva directamente de la limitación anteriormente planteada. Esta línea consiste en añadir al estudio de las tres competencias otras medidas de resultados no cognitivos, como el autoconcepto o la autoeficacia, para comprobar la influencia de los factores contextuales en su desarrollo, y qué variables procesuales pueden resultar beneficiosas para su promoción.

Por otro lado, el ámbito cognitivo de los estudiantes en este trabajo se evalúa a partir de las tres competencias, ya que esto proporciona una perspectiva más holística del rendimiento del alumnado. Sin embargo, desde un punto de vista metodológico, cada competencia se analiza por separado, y son los resultados de dichos análisis los que se utilizan para seleccionar los centros y estimar la eficacia total de los centros. Una posible línea futura para mejorar esta parte del proceso sería el empleo de modelos de ecuaciones estructurales MIMIC (múltiples indicadores y múltiples causas), los cuales permiten la

introducción de diversas medidas de rendimiento en el mismo modelo para calcular una variable latente (el rendimiento académico general) y comprobar el impacto de los factores de contexto en dicha variable (Tsai, Smith & Hauser, 2017), en lugar de construir un modelo separado para cada competencia.

La principal ventaja de este tipo de modelo con respecto a la práctica de los modelos separados es que “el modelo MIMIC permite probar las restricciones de igualdad impuestas en las cargas factoriales de cada competencia, estableciendo así si dichas cargas son estadísticamente diferentes entre sí en lugar de comparar los coeficientes a ojo” (Tsai et al., 2017, p. 84). Además, al estimar una variable latente de rendimiento general, podría suponer un entendimiento conceptual más completo de los factores que explican la variación en dicha variable latente (Finch & French, 2011).

Debido a estas posibles ventajas, se plantea como investigación futura una comparación entre el funcionamiento del procedimiento de análisis separado de las competencias presentado en esta tesis y la aplicación de modelos MIMIC, para comprobar cuáles son las diferencias entre ambos y valorar las aportaciones positivas de este nuevo enfoque.

8.3. VALOR AÑADIDO

Según la definición estricta de valor añadido, que se refiere al aporte del centro educativo al progreso de los estudiantes, el presente estudio no podría categorizarse dentro de esta corriente, ya que se basa en datos transversales. Sin embargo, hay algunos autores que realizan un uso más amplio del término para referirse al estudio de los efectos escolares en el rendimiento (Martínez Abad, Lizasoain, castro & Joaristi, 2017), por lo que en esta sección se discutirán los resultados relacionados con la magnitud de los efectos escolares encontrada, así como de los factores de proceso asociados a dicho efecto.

En el presente estudio, la variación atribuible a las diferencias entre escuelas, calculada a través del Coeficiente de Correlación Intraclase se sitúa entre el 12% y el 13% dependiendo de la competencia, y desciende hasta alrededor de un 5% tras controlar los efectos de las variables de contexto. Para evaluar la magnitud de estos datos se hace necesario compararlos con los resultados del CCI obtenidos del análisis de otros países, los cuales varían notablemente en la literatura. Examinando los artículos que forman parte de la SLR presentada en el capítulo 2, los valores más bajos encontrados entre los estudios que aplican modelos multinivel se corresponden con los países nórdicos, que muestran

un ICC de entre 6% y 13% (Sortkær & Reimer, 2018). Otros países con CCI relativamente bajos son Canadá y Estados Unidos, con 22% y 23% respectivamente (Milford, Ross & Anderson, 2006). En el otro lado del espectro, se encuentran países con CCI considerables, como Argentina (60%), México y Chile (45%) (Milford et al, 2006), Serbia con un CCI del 45% (Lazarevic & Orlic, 2018), Corea y Hong Kong con un 40% (Kim & Law, 2012), o Grecia con un 39% (Cheema, 2017).

Teniendo en cuenta estas referencias, se puede concluir que España es uno de los países con menor proporción de varianza en las puntuaciones de los estudiantes explicada a nivel de centro educativo, situándose al nivel de los países nórdicos. Mientras que este dato puede tomarse como una medida positiva de la equidad del sistema educativo (el acudir a un centro educativo u otro afecta en promedio un 13% al rendimiento del alumno), también significa que existe menos variabilidad que atribuir a las variables de proceso diferenciales de los centros, sobre todo teniendo en cuenta que, una vez aislado el efecto del contexto, la variabilidad atribuible a las prácticas escolares se reduce hasta el 5%.

Por su parte, las variables de proceso asociadas a la eficacia de los centros se exploraron en este trabajo a través de dos vías: un análisis cuantitativo de regresión logística, y un análisis cualitativo basado en estudio de casos. El análisis cuantitativo arrojó resultados poco satisfactorios en relación con el cumplimiento del objetivo de investigación correspondiente. A nivel de centro, ninguna de las variables introducidas (liderazgo, participación de los docentes en la toma de decisiones, desarrollo profesional del personal, desarrollo del currículo, etc.) resultó mostrar una relación significativa con la eficacia de los centros. Teniendo en cuenta la larga tradición de estudios sobre eficacia escolar que encuentran diferencias entre los centros de alta y baja eficacia en función de sus prácticas organizativas y pedagógicas (ver capítulo 3), quizás el foco de esta cuestión ha de situarse sobre la manera en la que los cuestionarios de contexto se planifican y aplican en PISA.

En el caso de las variables de proceso a nivel de estudiante, sí se encontraron algunas relaciones significativas con la variable de eficacia de los centros, aunque hubo algunos resultados llamativos en cuanto a la fortaleza y el signo de algunas de estas relaciones. Como ya se comentó en la sección 6.3, el análisis arrojó resultados aparentemente contradictorios sobre el papel del apoyo educativo y la atención personalizada, ya que las prácticas de adaptación de la instrucción a las necesidades del alumnado presentan una

relación positiva con la eficacia, mientras que el apoyo familiar y el apoyo y la retroalimentación ofrecidos durante el docente durante las clases mientras una relación negativa con esta variable.

El último análisis realizado sobre la naturaleza del aporte de las escuelas a la eficacia educativa fue el estudio cualitativo de casos en dos centros identificados como de alta y baja eficacia dentro de la Comunidad Autónoma de Castilla y León. Este estudio tiene como objetivo ampliar y complementar los análisis cualitativos sobre este mismo tema expuestos anteriormente a través de una exploración más profunda de las variables de procesos organizativos y pedagógicos de los centros.

Los dos principales estudios que se tomaron como referencia para este análisis fueron los llevados a cabo por Murillo (2007) y Lizasoain et al. (2013, 2015), debido a su estructura (selección de centros de alta y baja eficacia en función del rendimiento de los estudiantes y análisis posterior de los factores de proceso a través de métodos cualitativos).

A pesar de que los resultados obtenidos no pueden relacionarse de manera clara con la eficacia de los centros debido a diversos motivos, como una muestra muy reducida, una Comunidad Autónoma con pocas diferencias entre centros, o un muestreo poco transparente de alumnos con necesidades educativas especiales (ver 8.1.5), sí se han encontrado algunas cuestiones que parecen ir en la línea de los estudios existentes.

El centro A muestra características ligadas con la alta eficacia en los dos estudios mencionados. Los docentes presentan actitudes positivas y una alta motivación con respecto a la formación, con un gran nivel de demanda y participación a pesar de no necesitar méritos para el traslado por ser un centro de destino final para la mayoría de los docentes. También se percibe un alto grado de implicación en los proyectos educativos dirigidos al alumnado, tanto por parte de los docentes como del Equipo Directivo, y existe una arraigada cultura de evaluación de sus procesos y resultados con el objetivo de mejorarlos. Los responsables de la orientación y convivencia tienen un gran peso en el centro, siendo valorados como figuras relevantes y positivas para el clima y la resolución de conflictos por parte de todos los informantes. Esta cuestión se refleja también en la planificación de la convivencia en el centro, la cual constituye uno de los principales pilares institucionales, siendo promovida a través de un extenso plan de convivencia y también del plan de acogida del nuevo alumnado.

El liderazgo se percibe como repartido a todos los niveles (Equipo Directivo, Jefaturas de Estudios, Jefes de Departamento, etc.), aunque las funciones de los individuos que ejercen los cargos están claramente identificadas. Por último, el centro exhibe una gran apertura a colaboraciones con organizaciones externas tanto para la organización de actividades como para facilitar la transición entre primaria y secundaria de los alumnos de nuevo ingreso, y las relaciones establecidas con las familias se caracterizan como cordiales, constructivas y sistemáticas.

Por otro lado, este centro muestra también algunas características más comúnmente asociadas con los centros de baja eficacia que marcan sus posibles áreas de mejora para un futuro. Una de ellas es la ausencia de procesos sistematizados de trabajo en equipo por parte del profesorado. En ambos grupos de discusión realizados con docentes, los participantes expresaron una mayor inclinación por la organización individual del trabajo, o una organización basada en los departamentos de las distintas áreas, pero declararon que no contaban con tiempos de coordinación y trabajo en equipo para poder intercambiar información y colaborar con compañeros de otras áreas fuera del tratamiento de casos concretos de interés (alumnos con dificultades, cuestiones relacionadas con la tutoría, etc.). La organización de la atención a la diversidad y el seguimiento individualizado de los estudiantes constituye otra de las posibles áreas de mejora. Ninguno de los participantes mencionó en detalle la configuración de la atención a la diversidad en el centro, y las pocas medidas que se relataron en este sentido eran actividades extraescolares dirigidas a los estudiantes con bajo rendimiento. Este hecho, quizá motivado por el bajo porcentaje de alumnos con necesidades educativas especiales, puede apuntar a la necesidad de revisar los procesos de atención individualizada a las dificultades de aprendizaje dentro de las propias aulas.

Por su parte, el centro B presenta también algunas de las características que tanto Murillo (2007) como Lizasoain et al. (2013, 2015) identifican con los centros de alta eficacia. La mayoría de ellas tienen que ver con la atención a la diversidad y la inclusión, cuestiones fundamentales a la hora de caracterizar este centro. En este sentido, tanto las prácticas organizativas como pedagógicas se encuentran totalmente ajustadas a las necesidades de los estudiantes, cuestión que se refleja en la existencia de dos profesores por aula en casi todas las materias, el seguimiento individualizado de todos los estudiantes, o las acciones de atención a la diversidad realizadas dentro del aula. Además, todo el personal del centro demuestra unas expectativas altas del rendimiento de los

estudiantes, y también expresan una voluntad compartida de formar al alumnado de manera integral, promoviendo, por ejemplo, cuestiones como la educación emocional.

Sin embargo, el centro expone también algunas áreas de mejora. En general, no se percibe entre el profesorado una actitud particularmente positiva hacia las acciones formativas realizadas en el centro, lo cual podría afectar a su motivación y aprovechamiento de las mismas. En cuanto a la convivencia, a pesar de que todos los informantes coinciden en que el clima del centro es muy positivo, no se percibe una gran sistematización de su promoción, cuestión que se ejemplifica en la manera en la que los participantes se refieren en mayor medida a la existencia de relaciones cercanas que a medidas concretas para favorecer la convivencia.

Adicionalmente, se observa que el liderazgo en el centro tiene un carácter menos distribuido y compartido que en el centro A, así como una menor participación de los docentes en las decisiones pedagógicas y organizativas. Por último, la apertura del centro a la participación de organizaciones externas no es muy alta, ya que la mayoría de los proyectos que se desarrollan (así como las acciones formativas) son de carácter interno y sólo implican a agentes educativos pertenecientes al centro educativo.

8.3.1. Limitaciones

El uso de datos de PISA para evaluar la magnitud de los efectos de los centros educativos, así como la relevancia de las variables de contexto y proceso, conlleva una serie de limitaciones causadas por el propio diseño de la prueba.

Una de estas principales limitaciones metodológicas es la falta de datos longitudinales (Carabaña, 2015; Han, 2018; Rutkowski et al., 2017), o en su defecto de datos sobre el rendimiento previo de los estudiantes (Dumay & Dupriez, 2014; Willms, 2010), lo cual supone la imposibilidad de establecer modelos de valor añadido que permitan el estudio del progreso de los estudiantes, mostrando así un análisis del aprendizaje como un elemento dinámico. PISA sólo ofrece datos de una cohorte de estudiantes en un momento determinado de su escolarización, por lo que la medida que se hace de su rendimiento y las variables asociadas a él proporcionan una visión mucho más estática de la calidad del sistema educativo.

Otra falta clave de información que acusa la comunidad investigadora son los datos a nivel de aula (Lafontaine et al., 2015; Sandilands et al., 2014; Scheerens et al., 2015).

Como se ha reconocido en diversos estudios, las características de los docentes y de los procesos de enseñanza-aprendizaje llevados a cabo en el aula son altamente relevantes para el estudio de los resultados de los estudiantes (Chapman et al., 2016; Scheerens & Bosker, 1997). Sin embargo, aunque en 2015 PISA ha introducido información del profesorado, esta no supone una equivalencia con el nivel de aula, ya que los alumnos no se encuentran identificados en función del docente que les atiende.

En relación con los cuestionarios de contexto, algunos autores muestran insatisfacción con la manera en la que estos cuestionarios se construyen y contestan, alegando una menor fiabilidad de las puntuaciones con respecto a las obtenidas en las pruebas competenciales (Dolin, 2007), una baja correlación entre las respuestas de estudiantes y familias en preguntas parecidas (Rutkowski & Rutkowski, 2010), o una validez cuestionable debido a la dependencia de autoinformes (Cheema, 2017; Chi et al., 2018; Tourón et al., 2018). Además, como todas las medidas de percepción basadas en autoinformes, estos cuestionarios se ven afectados por sesgos cognitivos como la deseabilidad social, el efecto halo, o las tendencias centrales o extremas de respuesta (de las Cuevas & González, 1992), por lo que, sumado a las consideraciones anteriores, podría cuestionarse hasta qué punto la información proporcionada por los estudiantes, docentes y directores de los centros representa la realidad de manera fidedigna, o se ve en cambio afectada por sesgos conscientes o inconscientes. Este hecho, junto con el bajo ICC de los modelos finales, podría dificultar el establecimiento de relaciones significativas entre estas variables y el nivel de eficacia de los centros, ayudando a explicar parcialmente los resultados inconcluyentes de la aplicación de las técnicas de regresión logística con este fin.

8.3.2. Líneas futuras de investigación

En primer lugar, dado que la reducida muestra de centros sobre la que se ha realizado el estudio cualitativo limitan el alcance de las conclusiones que pueden extraerse de dicho análisis, una de las cuestiones que se plantean es la ampliación de esta muestra de cara a futuras investigaciones. Por otro lado, podría considerarse también la adición de otros grupos de informantes que puedan complementar las opiniones de los docentes y directivos, como las familias, los estudiantes, o los inspectores encargados del centro educativo. Cada uno de estos colectivos tiene un punto de vista diferente acerca de las prácticas y políticas educativas del centro, y puede aportar información valiosa que sirva para caracterizar al centro en mayor detalle.

Adicionalmente, teniendo en cuenta la falta de relaciones significativas con la eficacia mostrada por las variables a nivel de centro y aparente incongruencia de los resultados arrojados por las variables a nivel de estudiante, se estima necesario buscar otras fuentes de información que puedan ayudar a dar sentido a estos resultados. Además de la mencionada extensión de los estudios de caso a otros centros y/o informantes, cabe también la posibilidad de emplear otro tipo de técnicas basadas en cuestionarios para explorar las variables de proceso que hayan sido validadas y muestren un mejor ajuste con la realidad de los centros.

8.4. CONCLUSIONES

Los resultados del presente trabajo de investigación aportan información sobre dos cuestiones fundamentales relacionadas con el sistema educativo español: la equidad y la eficacia.

En el caso de la equidad, estudiada a través de la influencia de ciertas variables contextuales de estudiantes y centros sobre el nivel competencial del alumnado, se ponen de relevancia una serie de conclusiones principales:

- El nivel socioeconómico tiene una fuerte influencia sobre el rendimiento, y dicha influencia funciona a dos niveles diferentes: el nivel socioeconómico del propio estudiante y el nivel promedio de sus compañeros.
- El género tiene una relación considerable con el rendimiento en matemáticas y ciencias, presentando una ventaja a favor de los estudiantes varones. En el caso de la comprensión lectora, las mujeres presentan un mayor rendimiento, pero la diferencia entre ambos géneros es más pequeña.
- Los estudiantes que han repetido curso obtienen calificaciones significativamente inferiores a las de sus compañeros en las tres competencias estudiadas.
- Los estudiantes que han realizado un cambio de centro dentro de la misma etapa educativa presentan un nivel competencial inferior al de los demás estudiantes. Si han realizado dos o más cambios, la diferencia es aún mayor.
- El hecho de ser inmigrante de 1ª generación supone una disminución en el nivel competencial tanto de matemáticas como de ciencias, mientras que los inmigrantes de 2ª generación no muestran diferencias significativas con respecto a sus compañeros nativos.

- Los estudiantes en cuyos hogares se habla una lengua distinta a aquella en la que se desarrolla la instrucción presentan una competencia en comprensión lectora inferior a la de sus compañeros.

Las implicaciones para la práctica y política educativa que tienen estos hallazgos ya han sido exploradas en este capítulo, aunque cabe recordar su relevancia a la hora de asegurar una educación de calidad y en igualdad de oportunidades para todo el alumnado.

Por otro lado, se han estudiado también aquellos procesos educativos relacionados con la eficacia de los centros, aunque los resultados a este respecto han sido menos concluyentes. Sin embargo, se extraen algunas conclusiones generales que pueden servir como base para investigaciones futuras de profundización:

- Las variables de proceso a nivel de centro educativo recogidas por los cuestionarios de contexto de PISA no permiten realizar distinciones entre los centros de alta y baja eficacia.
- Las variables de proceso referidas a estudiantes producen algunos resultados que se estiman incongruentes a la luz de los datos disponibles, como el comportamiento dispar de variables aparentemente relacionadas (adaptación a la instrucción y apoyo del profesorado o apoyo emocional familiar).
- Los resultados del estudio de casos a través de técnicas cualitativas han coincidido en algunos factores de eficacia (o ineficacia) con la literatura existente, aunque se considera necesaria la ampliación de la muestra de centros.

Como conclusión final, la consecución de los objetivos planteados para esta tesis es satisfactoria, teniendo en cuenta la información y los recursos de los que se ha dispuesto a lo largo del proceso de investigación. Sin embargo, a raíz de la reflexión sobre los resultados, sus limitaciones y sus implicaciones, surgen prometedoras líneas de investigación futura, las cuales se espera que ayuden a profundizar en el funcionamiento del sistema educativo y permitan un mayor entendimiento de los mecanismos que promueven una educación eficaz para todos los estudiantes.

8.5. PUBLICACIONES DERIVADAS DEL TRABAJO DE INVESTIGACIÓN

Durante el desarrollo de la presente tesis doctoral se ha realizado una labor de diseminación de los principales hallazgos, dando lugar a las siguientes publicaciones.

Artículos

- Gamazo, A., Martínez-Abad, F., Olmos-Migueláñez, S., & Rodríguez-Conde, M. J. (2018). Evaluación de factores relacionados con la eficacia escolar en PISA 2015. Un análisis multinivel / Assessment of factors related to school effectiveness in PISA 2015. A multilevel analysis. *Revista de Educación*, (379), 56–84. doi:10.4438/1988-592X-RE-2017-379-369

Así mismo, se han realizado otros **tres artículos** relacionados con el trabajo de la tesis que se encuentran actualmente en proceso de revisión.

Capítulos de libro

- Martínez-Abad, F., Torrijos-Fincias, P., Gamazo, A., & Conde, M. J. (2018). Assessment of Information Literacy and Its Relationship With Learning Outcomes. En F. García-Peñalvo (Ed.), *Global Implications of Emerging Technology Trends* (pp. 1-18). Hershey, Estados Unidos: IGI Global. doi:10.4018/978-1-5225-4944-4.ch001

Comunicaciones en congresos

- Martínez-Abad, F., Gamazo, A., Rodríguez-Conde, M.J., & Hernández-Ramos, J.P. (2019). *Factores del profesorado asociados a la eficacia escolar en PISA 2015*. Comunicación presentada en el congreso AIDIPE 2019, celebrado en Madrid.
- Martínez-Abad, F., Gamazo, A.; Rodríguez-Conde, M.J. (2018). Big Data in Education: Detection of ICT Factors Associated with School Effectiveness with Data Mining Techniques. En *Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality*. pp. 145 - 150. ACM, 10/2018.
- Gamazo, A., Martínez-Abad, F., Olmos-Migueláñez, S., & Rodríguez-Conde, M. J. (2018). *The Relationship Between School-Level PISA Factors and School*

Effectiveness: A Study with PISA 2015 Data from Spain. Comunicación presentada en el congreso ECER 2018, celebrado en Bolzano.

- Martínez-Abad, F., Gamazo García, A., & Rodríguez-Conde, M.J. (2018). *Uso y autopercepción de factores asociados a la eficacia escolar en PISA 2015.* Comunicación presentada en el congreso JUTE 2018, celebrado en San Sebastián.
- Martínez-Abad, F., Gamazo García, A., & Rodríguez-Conde, M.J. (2018). *Identificación de factores asociados a la eficacia escolar en educación secundaria a partir de las pruebas PISA 2015.* Comunicación presentada en el congreso CIP 2018, celebrado en Buenos Aires.
- García-Llorente, H.J., Martínez-Abad, F., Gamazo, A & Rodríguez-Conde, M.J. (2017). *Subjective and Objective Assessment of Information Literacy in Adolescents from a Spanish Region with a High Performance in PISA.* Comunicación presentada en el congreso ECER 2017, celebrado en Copenhague.
- Gamazo, A. (2017) *Promoting equity in the assessment of education systems: a study on factors related to school effectiveness.* Comunicación presentada en el “Three minute research work presentation meeting” en el congreso ISATT 2017, celebrado en Salamanca.
- Gamazo, A., Olmos-Migueláñez, S., & Martínez-Abad, F. (2016) Multilevel models for the assessment of school effectiveness using PISA scores. En *Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality.* pp. 1161 - 1166. ACM, 11/2016.
- Gamazo, A., Martínez-Abad, F., Olmos-Migueláñez, S., & Rodríguez-Conde, M. J. (2016) Characterising High- And Low-Performing Schools. The Use Of Hierarchical Linear Models To Analyse PISA Scores. Comunicación presentada en el congreso ECER 2016, celebrado en Dublín.

Durante el desarrollo de la tesis también se han realizado **otros trabajos relacionados con el ámbito de la investigación educativa** que se detallan a continuación:

Artículos

- Blanco, J. L., Miguel, V., Delhaxhe, A., Vansteenkiste, W., Vázquez, E., Crespo, M., Gallego, J.; Gamazo, A.; ... & Valle, N. (2017). Información y estudios sobre

educación como objeto de cooperación territorial (II): la cooperación en Europa a través de la Red Eurydice como modelo de cooperación entre las administraciones educativas a través de la Red Española REDIE. *Journal of Supranational Policies of Education (JOSPOE)*, (2017), 31-68.

Capítulos de libro

- Olmos-Migueláñez, S.; Torrecilla-Sánchez, E.M.; Gamazo, A. (2017) Student Involvement in E-assessment: A Practical Case of Empowerment Through Assessment. En J.P. Bakken (Ed.) *Classrooms. Volume 1: Assessment Practices for Teachers and Student Improvement Strategies* (pp.109-128). Nueva York, Estados Unidos: Nova Publishing.

Comunicaciones en congresos

- Rodríguez-Conde, M.J., Olmos-Migueláñez, S., Gamazo, A., & O'Hara, J. (2018) Evaluation in education and Guidance: a perspective from 2018. En *Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality*. pp. 129 - 133. ACM, 10/2018.
- Torrecilla-Sánchez, E.M., Gamazo, A., Rodríguez-Conde, M.J., & Olmos-Migueláñez, S. (2017). *Percepción de la Orientación Universitaria de los estudiantes Egresados de Pedagogía*. Comunicación presentada en el CIOU 2017, celebrado en Zaragoza.
- Torrijos-Fincias, P., Gamazo, A., Torrecilla-Sánchez, E.M., Martínez-Abad, F., Hernández-Ramos, J.P., & Rodríguez-Conde, M.J. (2017). *Impact of Emotional Education programmes on the training of school teachers*. Comunicación presentada en el congreso ECER 2017, celebrado en Copenhague.
- Olmos-Migueláñez, S., Torrecilla-Sánchez, E.M., Gamazo, A., Ibarra-Sáiz, M.S., & Rodríguez-Gómez, G. (2017) Design of an assessment system for the improvement of competences among Pedagogy students. En *Proceedings of the 5th International Conference on Technological Ecosystems for Enhancing Multiculturality*. ACM, 10/2017.
- Gamazo, A., Olmos-Migueláñez, S., Torrecilla-Sánchez, E.M., Sánchez-Prieto, J.C. (2017). ¿Qué competencias desarrollan los grados del ámbito de la educación? Una evaluación del nivel competencial de los estudiantes. En *Actas XVIII Congreso Internacional de Investigación Educativa. Interdisciplinariedad y Transferencia*. pp. 1997 - 2008. AIDIPE2017, 06/2017.

- Olmos-Migueláñez, S., Torrecilla-Sánchez, E.M., & Gamazo, A. (2017) *Involvement of pre-service teachers in e-assessment activities. An empirical study on the correlation between self- and peer assigned grades*. Comunicación presentada en el congreso EuroSoTL 2017, celebrado en Lund.
- Gamazo, A., Olmos-Migueláñez, S., Torrecilla-Sánchez, E.M., Sánchez-Prieto, J.C. (2016) Assessing the development of student generic competences in education-related university degrees. En *Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality*. pp. 199 - 204. ACM, 11/2016.



Referencias

- Acosta, S. T., & Hsu, H. (2014a). Negotiating diversity: An empirical investigation into family, school and student factors influencing New Zealand adolescents' science literacy. *Educational Studies*, 40(1), 98-115. doi:10.1080/03055698.2013.830243
- Acosta, S., & Hsu, H. (2014b). Shared academic values: Testing a model of the association between Hong Kong parents' and adolescents' perception of the general value of science and scientific literacy. *Educational Studies*, 40(2), 174-195. doi:10.1080/03055698.2013.866889
- Adams, R., & Wu, M. (2003). *PISA 2000 technical report*. París, Francia: OECD.
- Addey, C. (2017). Golden relics & historical standards: How the OECD is expanding global education governance through PISA for development. *Critical Studies in Education*, 58(3), 311-325. doi:10.1080/17508487.2017.1352006
- Aitkin, M., & Longford, N. (1986). Statistical modelling issues in school effectiveness studies. *Journal of the Royal Statistical Society: Series A (General)*, 149(1), 1-43. doi:10.2307/2981882
- Arikan, S., van de Vijver, F. J. R., & Yagmur, K. (2017). PISA mathematics and reading performance differences of mainstream european and turkish immigrant students. *Educational Assessment, Evaluation and Accountability*, 29(3), 229-246. doi:10.1007/s11092-017-9260-6
- Ball, S. J. (1993). Education markets, choice and social class: The market as a class strategy in the UK and the USA. *British Journal of Sociology of Education*, 14(1), 3-19. doi:10.1080/0142569930140101
- Balzer, C., & Martens, K. (2007). All bark and no bite? The implementation styles of the European Union and the Organization for Economic Cooperation and Development in education policy. In J. Joachim, B. Reinalda & B. Verbeek (Eds.), *International organizations and implementation* (pp. 110-124). Londres, Reino Unido: Routledge.
- Barnard-Brak, L., Lan, W. Y., & Yang, Z. (2018). Differences in mathematics achievement according to opportunity to learn: A 4pL item response theory examination. *Studies in Educational Evaluation*, 56, 1-7. doi:10.1016/j.stueduc.2017.11.002

- Bell, A., Fairbrother, M., & Jones, K. (2019). Fixed and random effects models: Making an informed choice. *Quality & Quantity*, 53(2), 1051-1074. doi:10.1007/s11135-018-0802-x
- Berberoglu, G. (2005). An analysis of the Programme for International Student Assessment 2000 (PISA 2000) mathematical literacy data for Brazilian, Japanese and Norwegian students. *Studies in Educational Evaluation*, 31(4), 283-314. doi:10.1016/j.stueduc.2005.11.006
- Birnbaum, A. (1968). Some latent trait models and their use in inferring a student's ability. In F. M. Lord, & M. R. Novick (Eds.), *Statistical theories of mental test scores* (pp. 397-479). Reading, Estados Unidos: Addison-Wesley.
- Boedeker, P. (2017). Hierarchical linear modeling with maximum likelihood, restricted maximum likelihood, and fully bayesian estimation. *Practical Assessment, Research & Evaluation*, 22(2), 1-19.
- Bolívar, A. (2005). Equidad educativa y teorías de la justicia. *REICE: Revista Electrónica Iberoamericana Sobre Calidad, Eficacia y Cambio En Educación*, 3(2), 42-69.
- Brookover, W., Beady, C., Flood, P., Schweitzer, J., & Wisenbaker, J. (1979). *School social systems and student achievement: Schools can make a difference*. Nueva York, Estados Unidos: Praeger.
- Brookover, W. B., & Schneider, J. M. (1975). Academic environments and elementary school achievement. *Journal of Research and Development in Education*, 9, 82-91.
- Brookover, W. B., & Lezotte, L. W. (1979). *Changes in school characteristics coincident with changes in student achievement: Executive summary*. East Lansing, Estados Unidos: Michigan State University.
- Carabaña, J. (2015). *La inutilidad de PISA para las escuelas*. Madrid: Catarata.
- Carabaña, J. (2016). El informe Coleman, 50 años después. *Revista De Sociología De La Educación-RASE*, 9(1), 9-21.
- Caro, D. H., Sandoval-Hernández, A., & Lüdtke, O. (2013). Cultural, social, and economic capital constructs in international assessments: An evaluation using exploratory structural equation modeling. *School Effectiveness and School Improvement*, 25(3), 433-450. doi:10.1080/09243453.2013.812568

- Caso-Niebla, J., & Hernández-Guzmán, L. (2010). Variables que inciden en el rendimiento académico de adolescentes mexicanos. *Revista Latinoamericana De Psicología*, 39(3), 487-501.
- Castejón, J. (1996). *Determinantes del rendimiento académico y de los centros educativos: Factores y modelos*. Alicante: Editorial Club Universitario.
- Castejón, J. L. (1994). Estabilidad de diversos índices de eficacia de centros educativos. *Revista de Investigación Educativa, RIE*, 12(24), 45-60.
- Castro, M., & Lizasoain, L. (2012). Las técnicas de modelización estadística en la investigación educativa: Minería de datos, modelos de ecuaciones estructurales y modelos jerárquicos lineales. *Revista Española De Pedagogía*, 70(251), 131-148.
- Chaparro, A. A., González Barbera, C., & Caso Niebla, J. (2016). Familia y rendimiento académico: Configuración de perfiles estudiantiles en secundaria. *Revista Electrónica de Investigación Educativa*, 18(1), 53-68.
- Chapman, C., Muijs, D., Reynolds, D., Sammons, P., & Teddlie, C. (2015). *The Routledge international handbook of educational effectiveness and improvement: Research, policy, and practice*. Nueva York, Estados Unidos: Routledge.
- Cheema, J. R. (2018). Effect of math-specific self-efficacy on math literacy: Evidence from a greek survey. *Research in Education*, 102(1), 13-36. doi:10.1177/0034523717741914
- Cheung, K., Sit, P., Soh, K., Jeong, M., & Mak, S. (2014). Predicting academic resilience with reading engagement and demographic variables: Comparing Shanghai, Hong Kong, Korea, and Singapore from the PISA perspective. *The Asia-Pacific Education Researcher*, 23(4), 895-909. doi:10.1007/s40299-013-0143-4
- Chi, S., Liu, X., Wang, Z., & Won Han, S. (2018). Moderation of the effects of scientific inquiry activities on low SES students' PISA 2015 science achievement by school teacher support and disciplinary climate in science classroom across gender. *International Journal of Science Education*, 40(11), 1284-1304. doi:10.1080/09500693.2018.1476742
- Chiao, C., & Chiu, C. (2018). The mediating effect of ICT usage on the relationship between students' socioeconomic status and achievement. *The Asia-Pacific Education Researcher*, 27(2), 109-121. doi:10.1007/s40299-018-0370-9

- Chiu, M. M. (2015). Family inequality, school inequalities, and mathematics achievement in 65 countries: Microeconomic mechanisms of rent seeking and diminishing marginal returns. *Teachers College Record*, 117(1), 010305.
- Choi, A., Gil, M., Mediavilla, M., & Valbuena, J. (2016). *Double toil and trouble: Grade retention and academic performance*. (IEB Working Paper No. 2016/07). doi:10.2139/ssrn.2745970
- Clarke, P., Crawford, C., Steele, F., & Vignoles, A. F. (2010). *The choice between fixed and random effects models: Some considerations for educational research*. (IZA Discussion Papers No. 5287). Bonn, Alemania: Institute for the Study of Labor (IZA). Recuperado de <http://hdl.handle.net/10419/46182>
- Coleman, J. S. (1966). *Equality of educational opportunity*. Washington, Estados Unidos: National Center for Education Statistics.
- Condrón, D. J. (2011). Egalitarianism and educational excellence: Compatible goals for affluent societies? *Educational Researcher*, 40(2), 47-55. doi:10.3102/0013189X11401021
- Cordero Ferrera, J. M., Pedraja Chaparro, F., & Simancas Rodríguez, R. (2015). Factores del éxito escolar en condiciones socioeconómicas desfavorables. *Revista De Educación*, 370, 172-198. doi:10.4438/1988-592X-RE-2015-370-302
- Creemers, B. P. (1994). *The effective classroom*. Londres, Reino Unido: Cassell.
- Cruz-Benito, J., García-Peñalvo, F. J., & Therón, R. (2019). Analyzing the software architectures supporting HCI/HMI processes through a systematic review of the literature. *Telematics and Informatics*, 38, 118-132. doi:10.1016/j.tele.2018.09.006
- Darmawan, I., & Keeves, J. P. (2006). Accountability of teachers and schools: A value-added approach. *International Education Journal*, 7(2), 174-188.
- De la Orden Hoz, A., & Jornet Meliá, J. M. (2012). La utilidad de las evaluaciones de sistemas educativos: El valor de la consideración del contexto. *Bordón. Revista De Pedagogía*, 64(2), 69-88.
- de las Cuevas, C., & González, J. L. (1992). Autoinformes y respuestas sesgadas. *Anales De Psiquiatría*, 8(9), 362-366.

- Dolin, J. (2007). PISA - an example of the use and misuse of large-scale comparative test. In S. T. Hopmann, G. Brinek & M. Retzl (Eds.), *PISA zufolge PISA - PISA according to PISA: Hält PISA was es verspricht - does PISA keep what it promises?* (pp. 93-126). Viena, Austria: Lit-Verlag.
- Dolin, J., & Krogh, L. B. (2010). The relevance and consequences of PISA science in a Danish context. *International Journal of Science and Mathematics Education*, 8(3), 565-592. doi:10.1007/s10763-010-9207-6
- Dumay, X., & Dupriez, V. (2014). Educational quasi-markets, school effectiveness and social inequalities. *Journal of Education Policy*, 29(4), 510-531. doi:10.1080/02680939.2013.850536
- Edmonds, R. (1979). Effective schools for the urban poor. *Educational Leadership*, 37(1), 15-24.
- Egger, M., Smith, G.D., Schneider, M. y Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315(7109), 629-634. doi:10.1136/bmj.315.7109.629
- Eickelmann, B., Gerick, J., & Koop, C. (2017). ICT use in mathematics lessons and the mathematics achievement of secondary school students by international comparison: Which role do school level factors play? *Education and Information Technologies*, 22(4), 1527-1551. doi:10.1007/s10639-016-9498-5
- Ercikan, K., Chen, M. Y., Lyons-Thomas, J., Goodrich, S., Sandilands, D., Roth, W., & Simon, M. (2015). Reading proficiency and comparability of mathematics and science scores for students from English and non-English backgrounds: An international perspective. *International Journal of Testing*, 15(2), 153-175. doi:10.1080/15305058.2014.957382
- Fernández-Cano, A. (2016). Una crítica metodológica a las evaluaciones PISA. *RELIEVE-Revista Electrónica De Investigación y Evaluación Educativa*, 22(1), 1-17. doi:10.7203/relieve.22.1.8806
- Finch, W. H., & French, B. F. (2011). Estimation of MIMIC model parameters with multilevel data. *Structural Equation Modeling*, 18(2), 229-252. doi:10.1080/10705511.2011.557338

- Fruehwirth, J. C., Navarro, S., & Takahashi, Y. (2016). How the timing of grade retention affects outcomes: Identification and estimation of time-varying treatment effects. *Journal of Labor Economics*, 34(4), 979-1021. doi:10.1086/686262
- Gabriel, F., Signolet, J., & Westwell, M. (2018). A machine learning approach to investigating the effects of mathematics dispositions on mathematical literacy. *International Journal of Research & Method in Education*, 41(3), 306-327. doi:10.1080/1743727X.2017.1301916
- Gamazo, A., Martínez-Abad, F., Olmos-Migueláñez, S., & Rodríguez-Conde, M. J. (2018). Evaluación de factores relacionados con la eficacia escolar en PISA 2015. Un análisis multinivel. *Revista de Educación*, 379, 56-84. doi: 10.4438/1988-592X-RE-2017-379-369
- García Sánchez, F., Therón, R., & Gómez-Isla, J. (2019). Visual literacy in New Media: Systematic Review and Mapping of the Literature. *Education in the Knowledge Society*, 20, 1-44. doi:10.14201/eks2019_20_a6
- Gil-Flores, J., & García-Gómez, S. (2017). The importance of teaching practices in relation to regional educational policies in explaining PISA achievement. *Revista de Educación*, 378, 52-77. doi:10.4438/1988-592X-RE-2017-378-361
- Givord, P. (2019). *How are school-choice policies related to social diversity in schools?* (PISA in Focus No. 96). doi:10.1787/2d448c77-en
- Goetz, J. P., & LeCompte, M. D. (1988). *Etnografía y diseño cualitativo en investigación educativa*. Madrid: Ediciones Morata.
- Goldstein, H. (1987). *Multilevel models in education and social research*. Nueva York, Estados Unidos: Oxford University Press.
- González Barbera, C., Caso Niebla, J., Díaz López, K., & López Ortega, M. (2012). Rendimiento académico y factores asociados: Aportaciones de algunas evaluaciones a gran escala. *Bordón. Revista De Pedagogía*, 64(2), 51-68.
- Gorard, S. (2003). What is Multi-level modelling for? *British Journal of Educational Studies*, 51(1), 46-63. doi:10.1111/1467-8527.t01-2-00224
- Greger, D. (2012). When PISA does not matter? the case of the Czech Republic and Germany. *Human Affairs*, 22(1), 31-42. doi:10.2478/s13374-012-0004-5

- Grisay, A., De Jong, J., Gebhardt, E., Berezner, A., & Halleux-Monseur, B. (2007). Translation equivalence across PISA countries. *Journal of Applied Measurement*, 8(3), 249.
- Gupta, V., Hanges, P. J., & Dorfman, P. (2002). Cultural clusters: Methodology and findings. *Journal of World Business*, 37(1), 11-15. doi:10.1016/s1090-9516(01)00070-0
- Han, S. W. (2018). School-based teacher hiring and achievement inequality: A comparative perspective. *International Journal of Educational Development*, 61, 82-91. doi:10.1016/j.ijedudev.2017.12.004
- Hayes, A. F. (2006). A primer on multilevel modeling. *Human Communication Research*, 32(4), 385-410. doi:10.1111/j.1468-2958.2006.00281.x
- He, J., Van, d. V., & Kulikova, A. (2017). Country-level correlates of educational achievement: Evidence from large-scale surveys. *Educational Research and Evaluation*, 23(5-6), 163-179. doi:10.1080/13803611.2017.1455288
- Hill, H. C. (2017). The Coleman report, 50 years on: What do we know about the role of schools in academic inequality? *The Annals of the American Academy of Political and Social Science*, 674(1), 9-26. doi:10.1177/0002716217727510
- Hu, X., Gong, Y., Lai, C., & Leung, F. K. (2018). The relationship between ICT and student literacy in mathematics, reading, and science across 44 countries: A multilevel analysis. *Computers & Education*, 125, 1-13. doi:10.1016/j.compedu.2018.05.021
- Hwang, J., Choi, K. M., Bae, Y., & Shin, D. H. (2018). Do teachers' instructional practices moderate equity in mathematical and scientific literacy?: An investigation of the PISA 2012 and 2015. *International Journal of Science and Mathematics Education*, 16(1), 25-45. doi:10.1007/s10763-018-9909-8
- Jakobi, A. P., & Teltemann, J. (2011). Convergence in education policy? A quantitative analysis of policy change and stability in OECD countries. *Compare: A Journal of Comparative and International Education*, 41(5), 579-595. doi:10.1080/03057925.2011.566442
- Jencks, C. S., Smith, M., Ackland, H., Bane, M. J., Cohen, D., Gintis, H., . . . Michelson, S. (1972). *Inequality: A reassessment of the effect of the family and schooling in America*. Nueva York, Estados Unidos: Basic books.

- Jensen, A. (1969). How much can we boost IQ and scholastic achievement. *Harvard Educational Review*, 39(1), 1-123. doi:10.17763/haer.39.1.13u15956627424k7
- Jimerson, S. R. (2001). A synthesis of grade retention research: Looking backward and moving forward. *The California School Psychologist*, 6(1), 47-59. doi:10.1007/BF03340883
- Joaristi, L., Lizasoain, L., & Azpillaga, V. (2014). Detección y caracterización de los centros escolares de alta eficacia de la Comunidad Autónoma del País Vasco mediante modelos transversales contextualizados y modelos jerárquicos lineales. *Estudios Sobre Educación*, 27, 37-61. doi:10.15581/004.27.37-61
- Jornet Meliá, J. M. (2016). Análisis metodológico del proyecto PISA como evaluación internacional. *RELIEVE. Revista Electrónica De Investigación y Evaluación Educativa*, 22(1), 1-26. doi:10.7203/relieve22.1.8293
- Kalaycioglu, D. B. (2015). The influence of socioeconomic status, self-efficacy, and anxiety on mathematics achievement in England, Greece, Hong Kong, the Netherlands, Turkey, and the USA. *Educational Sciences: Theory and Practice*, 15(5), 1391-1401. doi:10.12738/estp.2015.5.2731
- Kanes, C., Morgan, C., & Tsatsaroni, A. (2014). The PISA mathematics regime: Knowledge structures and practices of the self. *Educational Studies in Mathematics*, 87(2), 145-165. doi:10.1007/s10649-014-9542-6
- Keeves, J. P., Hungi, N., & Afrassa, T. (2005). Measuring value added effects across schools: Should schools be compared in performance? *Studies in Educational Evaluation*, 31(2-3), 247-266. doi:10.1016/j.stueduc.2005.05.012
- Kim, D. H., & Law, H. (2012). Gender gap in maths test scores in South Korea and Hong Kong: Role of family background and single-sex schooling. *International Journal of Educational Development*, 32(1), 92-103. doi:10.1016/j.ijedudev.2011.02.009
- Kim, S. (2018). ICT for children of immigrants: Indirect and total effects via self-efficacy on math performance. *Journal of Educational Computing Research*, 55(8), 1168-1200. doi:10.1177/0735633117699954

- Kitchenham, B.A., Budgen, D. & Brereton, P.O. (2011). Using mapping studies as the basis for further research – A participant-observer case study. *Information and Software Technology*, 53(6), 638-651. doi:10.1016/j.infsof.2010.12.011
- Kitchenham, B. & Charters, S. (2007). *Guidelines for performing Systematic Literature Reviews in Software Engineering. Version 2.3* (EBSE-2007-01) School of Computer Science and Mathematics, Keele University. Recuperado de <https://goo.gl/L1VHcw>
- Kleinbaum, D. G., Dietz, K., Gail, M., Klein, M., & Klein, M. (2002). *Logistic regression*. Nueva York, Estados Unidos: Springer.
- Kreiner, S., & Christensen, K. B. (2014). Analyses of model fit and robustness. A new look at the PISA scaling model underlying ranking of countries according to reading literacy. *Psychometrika*, 79(2), 210-231. doi:10.1007/s11336-013-9347-z
- Kyriakides, L., & Creemers, B. P. (2008). Using a multidimensional approach to measure the impact of classroom-level factors upon student achievement: A study testing the validity of the dynamic model. *School Effectiveness and School Improvement*, 19(2), 183-205. doi:10.1080/09243450802047873
- Lafontaine, D., Baye, A., Vieluf, S., & Monseur, C. (2015). Equity in opportunity-to-learn and achievement in reading: A secondary analysis of PISA 2009 data. *Studies in Educational Evaluation*, 47, 1-11. doi:10.1016/j.stueduc.2015.05.001
- Lazarević, L. B., & Orlić, A. (2018). PISA 2012 mathematics literacy in Serbia: A multilevel analysis of students and schools. *Psihologija*, 51(4), 413-432. doi:10.2298/PSI170817017L
- Le Donne, N. (2014). European variations in socioeconomic inequalities in students' cognitive achievement: The role of educational policies. *European Sociological Review*, 30(3), 329-343. doi:10.1093/esr/jcu040
- Lee, E. S., & Forthofer, R. N. (2005). *Analyzing complex survey data*. Thousand Oaks, Estados Unidos: Sage Publications.
- Lee, V. E. (2000). Using hierarchical linear modeling to study social contexts: The case of school effects. *Educational Psychologist*, 35(2), 125-141. doi:10.1207/s15326985ep3502_6

- LeRoy, B. W., Samuel, P., Deluca, M., & Evans, P. (2018). Students with special educational needs within PISA. *Assessment in Education: Principles, Policy & Practice*, 1-11. doi:10.1080/0969594X.2017.1421523
- Levine, D. U., & Lezotte, L. W. (1990). Unusually effective schools: A review and analysis of research and practice. *School Effectiveness and School Improvement*, 1(3), 221-224. doi:10.1080/0924345900010305
- Lezotte, L., Edmonds, R., & Ratner, G. (1974). *A final report: Remedy for school failure to equitable basic school skills*. East Lansing, Estados Unidos: Michigan State University.
- Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin*, 136(6), 1123-1135. doi:10.1037/a0021276
- Lingard, B., Martino, W., & Rezai-Rashti, G. (2013). Testing regimes, accountabilities and education policy: Commensurate global and national developments. *Journal of Educational Policy*, 28(5), 539-556. doi:10.1080/02680939.2013.820042
- Liou, P., & Hung, Y. (2015). Statistical techniques utilized in analyzing PISA and TIMSS data in science education from 1996 to 2013: A methodological review. *International Journal of Science and Mathematics Education*, 13(6), 1449-1468. doi:10.1007/s10763-014-9558-5
- Liu, X., & Whitford, M. (2011). Opportunities-to-learn at home: Profiles of students with and without reaching science proficiency. *Journal of Science Education and Technology*, 20(4), 375-387. doi:10.1007/s10956-010-9259-y
- Lizasoain, L., Angulo, A., Azpillaga, V., Bartau, I., Damborenea, M. D., del Frago, R., . . . Valadez, C. (2015). *La eficacia escolar en los centros del País Vasco. Informe final 2011-2015*. País Vasco: Universidad del País Vasco - Instituto Vasco de Evaluación e Investigación Educativa.
- Lizasoain, L., Angulo, A., Azpillaga, V., Damborenea, M. D., del Frago, R., Etxeberría, F., . . . Valadez, C. (2013). *Caracterización y buenas prácticas de los centros escolares de alto valor añadido. Informe final fase I (año 2012)*. País Vasco: Universidad del País Vasco - Instituto Vasco de Evaluación e Investigación Educativa.

- Luu, K., & Freeman, J. G. (2011). An analysis of the relationship between information and communication technology (ICT) and scientific literacy in Canada and Australia. *Computers & Education, 56*(4), 1072-1082. doi:10.1016/j.compedu.2010.11.008
- Maas, C. J., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology, 1*(3), 86-92. doi:10.1027/1614-2241.1.3.86
- Martens, K. (2007). Martens, K. (2007). How to become an influential actor: The 'comparative turn' in OECD education policy. In K. Martens, A. Rusconi & K. Leuze (Eds.), *New arenas of education governance* (pp. 40-56). Londres, Reino Unido: Palgrave Macmillan.
- Martinez Abad, F., & Chaparro Caso López, Alicia A. (2017). Data-mining techniques in detecting factors linked to academic achievement. *School Effectiveness and School Improvement, 28*(1), 39-55. doi:10.1080/09243453.2016.1235591
- Martínez Abad, F., Lizasoain Hernández, L., Castro Morera, M., & Joaristi Olariaga, L. M. (2017). Selección de escuelas de alta y baja eficacia en baja california (México). *Revista Electrónica De Investigación Educativa, 19*(2), 38-53. doi:10.24320/redie.2017.19.2.960
- Maslowski, R., Scheerens, J., & Luyten, H. (2007). The effect of school autonomy and school internal decentralization on students' reading literacy. *School Effectiveness and School Improvement, 18*(3), 303-334. doi:10.1080/09243450601147502
- McConney, A., & Perry, L. B. (2010a). Science and mathematics achievement in Australia: The role of school socioeconomic composition in educational equity and effectiveness. *International Journal of Science and Mathematics Education, 8*(3), 429-452. doi:10.1007/s10763-010-9197-4
- McConney, A., & Perry, L. B. (2010b). Socioeconomic status, self-efficacy, and mathematics achievement in Australia: A secondary analysis. *Educational Research for Policy and Practice, 9*(2), 77-91. doi:10.1007/s10671-010-9083-4
- Meggiolaro, S. (2018). Information and communication technologies use, gender and mathematics achievement: Evidence from Italy. *Social Psychology of Education, 21*(2), 497-516. doi:10.1007/s11218-017-9425-7
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Thousand Oaks, Estados Unidos: Sage Publishing.

- Milford, T., Ross, S. P., & Anderson, J. O. (2010). An opportunity to better understand schooling: The growing presence of PISA in the Americas. *International Journal of Science and Mathematics Education*, 8(3), 453-473. doi:10.1007/s10763-010-9201-z
- Ministerio de Educación, Cultura y Deporte. (2016). *PISA 2015 programa para la evaluación internacional de los alumnos, informe español*. Madrid: Secretaría General Técnica.
- Mortimore, P., Sammons, P., Stoll, L., Lewis, D., & Ecob, R. (1989). A study of effective junior schools. *International Journal of Educational Research*, 13(7), 753-768. doi:10.1016/0883-0355(89)90026-8
- Mundy, K. (2007). Global governance, educational change. *Comparative Education*, 43(3), 339-357. doi:10.1080/03050060701556281
- Muñoz-Repiso, M., Cerdán, J., Murillo, F. J., Calzón, J., Castro, M., Egido, I., . . . Lucio-Villegas, M. (1995). *Calidad de la educación y eficacia de la escuela. estudio sobre la gestión de los recursos educativos*. Madrid: CIDE.
- Muñoz-Repiso, M., Murillo, F. J., Barrio, R., Brioso, M. J., Hernández, M. L., & Pérez-Albo, M. J. (2000). *La mejora de la eficacia escolar: Un estudio de casos*. Madrid: CIDE.
- Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *ETS Research Report Series*, 16(2), 159-177. doi:10.1002/j.2333-8504.1992.tb01436.x
- Murillo, F. J. (2003). *La investigación sobre eficacia escolar en Iberoamérica: Revisión internacional del estado del arte*. Bogotá, Colombia: Convenio Andrés Bello.
- Murillo, F. J. (2005). *La investigación sobre la eficacia escolar*. Barcelona: Octaedro.
- Murillo, F. J. (2007). *Investigación iberoamericana sobre eficacia escolar*. Bogotá, Colombia: Convenio Andrés Bello.
- Murillo, F. J. (2008). Hacia un modelo de eficacia escolar. estudio multinivel sobre los factores de eficacia de las escuelas españolas. *REICE: Revista Iberoamericana Sobre Calidad, Eficacia y Cambio En Educación*, 6(1), 4-28.
- Niemann, D., Martens, K., & Teltemann, J. (2017). PISA and its consequences: Shaping education policies through international comparisons. *European Journal of Education*, 52(2), 175-183. doi:10.1111/ejed.12220

- Oliveri, M. E., & von Davier, M. (2014). Toward increasing fairness in score scale calibrations employed in international large-scale assessments. *International Journal of Testing*, 14(1), 1-21. doi:10.1080/15305058.2013.825265
- Organisation for Economic Co-operation and Development (OECD). (1967). *Methods and statistical needs for educational planning*. Paris, Francia: OECD Publishing.
- Organisation for Economic Co-operation and Development (OECD). (1999). *Measuring student knowledge and skills: A new framework for assessment*. Paris, Francia: OECD Publishing.
- Organisation for Economic Co-operation and Development (OECD). (2016). *PISA 2015 results (volume I): Excellence and equity in education*. Paris, Francia: OECD Publishing.
- Organisation for Economic Co-operation and Development (OECD). (2017). *PISA 2015 assessment and analytical framework: Science, reading, mathematics, financial literacy and collaborative problem solving*. Paris, Francia: OECD Publishing. doi:10.1787/9789264281820-en
- Organisation for Economic Co-operation and Development (OECD). (2017). *PISA 2015 technical report*. Paris, Francia: OECD Publishing.
- Organisation for Economic Co-operation and Development (OECD). (2018). *Equity in education: Breaking down barriers to social mobility*. Paris, Francia: OECD Publishing. doi:10.1787/9789264073234-en
- Organisation for Economic Co-operation and Development (OECD). (2019). *Balancing school choice and equity: An international perspective based on PISA*. Paris, Francia: OECD Publishing. doi:10.1787/2592c974-en
- Özdemir, B., & Gelbal, S. (2014). Investigating factors that affect Turkish students' academic success with canonical commonality analysis according to PISA 2009 results. *Egitim Ve Bilim*, 39(175), 41-57. doi:10.15390/EB.2014.3025
- Özdemir, C. (2016). Equity in the Turkish education system: A multilevel analysis of social background influences on the mathematics performance of 15-year-old students. *European Educational Research Journal*, 15(2), 193-217. doi:10.1177/1474904115627159

- Papadopoulos, G. S. (1994). *Education 1960-1990: The OECD perspective*. París, Francia: OECD Publishing.
- Perry, L. B., & McConney, A. (2013). School socioeconomic status and student outcomes in reading and mathematics: A comparison of Australia and Canada. *Australian Journal of Education, 57*(2), 124-140. doi:10.1177/0004944113485836
- Perry, L. B., & McConney, A. (2010). Does the SES of the school matter? an examination of socioeconomic status and student achievement using PISA 2003. *Teachers College Record, 112*(4), 1137-1162.
- Perry, L., & McConney, A. (2010). School socio-economic composition and student outcomes in Australia: Implications for educational policy. *Australian Journal of Education, 54*(1), 72-85. doi:10.1177/000494411005400106
- Petko, D., Cantieni, A., & Prasse, D. (2017). Perceived quality of educational technology matters: A secondary analysis of students' ICT use, ICT-related attitudes, and PISA 2012 test scores. *Journal of Educational Computing Research, 54*(8), 1070-1091. doi:10.1177/0735633116649373
- Purkey, S. C., & Smith, M. S. (1983). Effective schools: A review. *The Elementary School Journal, 83*(4), 427-452. doi:10.1086/461325
- Rabe-Hesketh, S., & Skrondal, A. (2008). *Multilevel and longitudinal modeling using Stata*. Texas, Estados Unidos: STATA Press.
- Rasch, G. (1961). On general laws and the meaning of measurement in psychology. Paper presented at the *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, (Vol. 4, pp. 321-333).
- Raudenbush, S. W., & Bryk, A. S. (1992). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, Estados Unidos: Sage Publications.
- Raudenbush, S. W., & Willms, J. (1995). The estimation of school effects. *Journal of Educational and Behavioral Statistics, 20*(4), 307-335. doi:10.3102/10769986020004307
- Raudenbush, S., & Bryk, A. S. (1986). A hierarchical model for studying school effects. *Sociology of Education, 59*(1), 1-17. doi:10.2307/2112482

- Reynolds, D., Hargreaves, A., & Blackstone, T. (1980). Review symposium of Rutter et al.'s fifteen thousand hours. *British Journal of Sociology of Education*, *1*, 207-219. doi:10.1080/0142569800010206
- Reynolds, D., Sammons, P., De Fraine, B., Van Damme, J., Townsend, T., Teddlie, C., & Stringfield, S. (2014). Educational effectiveness research (EER): A state-of-the-art review. *School Effectiveness and School Improvement*, *25*(2), 197-230. doi:10.1080/09243453.2014.885450
- Rodríguez-Santero, J., & Gil-Flores, J. (2018). Contextual variables associated with differences in educational performance between European Union countries / Variables contextuales asociadas a las diferencias de rendimiento educativo entre los países de la Unión Europea. *Cultura y Educación*, *30*(4), 605-632. doi:10.1080/11356405.2018.1522024
- Rutkowski, D., Rutkowski, L., Wild, J., & Burroughs, N. (2018). Poverty and educational achievement in the US: A less-biased estimate using PISA 2012 data. *Journal of Children and Poverty*, *24*(1), 47-67. doi:10.1080/10796126.2017.1401898
- Rutkowski, L., & Rutkowski, D. (2010). Getting it 'better': The importance of improving background questionnaires in international large-scale assessment. *Journal of Curriculum Studies*, *42*(3), 411-430. doi:10.1080/00220272.2010.487546
- Rutter, M., Maughan, B., Mortimore, P., & Ouston, J. (1979). *Fifteen thousand hours: Secondary schools and their effects on children*. Cambridge, Estados Unidos: Harvard University Press.
- Sammons, P., Hillman, J., & Mortimore, P. (1995). *Key characteristics of effective schools. A review of school effectiveness research*. Londres, Reino Unido: Institute of Education, University of London.
- Sanders, W. L., & Wright, S. P. (2008). *A response to Amrein-Beardsley (2008) "methodological concerns about the education Value-Added assessment system"*
- Sandilands, D. (., Barclay McKeown, S., Lyons-Thomas, J., & Ercikan, K. (2014). An investigation of school-level factors associated with science performance for minority and majority francophone students in Canada. *Canadian Journal of Science, Mathematics and Technology Education*, *14*(2), 135-153. doi:10.1080/14926156.2014.903316

- Schachner, M. K., He, J., Heizmann, B., & Van de Vijver, Fons JR. (2017). Acculturation and school adjustment of immigrant youth in six European countries: Findings from the Programme for International Student Assessment (PISA). *Frontiers in Psychology*, 8, 649. doi:10.3389/fpsyg.2017.00649
- Scheerens, J. (1990). School effectiveness research and the development of process indicators of school functioning. *School Effectiveness and School Improvement*, 1(1), 61-80. doi:10.1080/0924345900010106
- Scheerens, J., & Bosker, R. (1997). *The foundations of educational effectiveness*. Oxford, Reino Unido: Pergamon.
- Scheerens, J., & Creemers, B. P. (1996). School effectiveness in the Netherlands: The modest influence of a research programme. *School Effectiveness and School Improvement*, 7(2), 181-195. doi:10.1080/0924345960070205
- Scheerens, J., Luyten, H., van, d. B., & Glas, C. A. W. (2015). Exploration of direct and indirect associations of system-level policy-amenable variables with reading literacy performance. *Educational Research and Evaluation*, 21(1), 15-39. doi:10.1080/13803611.2015.1008520
- Sebastian, J., & Huang, H. (2016). Examining the relationship of a survey based measure of math creativity with math achievement: Cross-national evidence from PISA 2012. *International Journal of Educational Research*, 80, 74-92. doi:10.1016/j.ijer.2016.08.010
- Sexton, P. C. (1961). *Education and income: Inequalities in our public schools*. Nueva York, Estados Unidos: Viking Press.
- Sjøberg, S. (2015). PISA and global educational governance-A critique of the project, its uses and implications. *Eurasia Journal of Mathematics, Science & Technology Education*, 11(1), 111-127. doi:10.12973/eurasia.2015.1310a
- Snijders, T., & Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Thousand Oaks, Estados Unidos: Sage Publications.
- Sortkær, B., & Reimer, D. (2018). Classroom disciplinary climate of schools and gender – evidence from the nordic countries. *School Effectiveness and School Improvement*, 29(4), 511-528. doi:10.1080/09243453.2018.1460382

- Stoet, G., & Geary, D. C. (2013). Sex differences in mathematics and reading achievement are inversely related: Within-and across-nation assessment of 10 years of PISA data. *PLoS One*, 8(3), e57988. doi:10.1371/journal.pone.0057988
- Stringfield, S. C., & Slavin, R. E. (1992). A hierarchical longitudinal model for elementary school effects. In B. P. M. Creemers, & G. J. Reezigt (Eds.), *Evaluation of educational effectiveness* (pp. 35-69). Groninga, Países Bajos: ICO.
- Täht, K., Must, O., Peets, K., & Kattel, R. (2014). Learning motivation from a cross-cultural perspective: A moving target? *Educational Research and Evaluation*, 20(4), 255-274. doi:10.1080/13803611.2014.929009
- Teddlie, C., & Stringfield, S. (1993). *Schools make a difference: Lessons learned from a 10-year study of school effects*. Nueva York, Estados Unidos: Teachers College Press.
- Teltemann, J., & Jude, N. (2019). Assessments and accountability in secondary education: International trends. *Research in Comparative and International Education*, 14 (2), 249-271. doi: 10.1177/1745499919846174
- Thomas, S., Sammons, P., Mortimore, P., & Smees, R. (1997). Differential secondary school effectiveness: Comparing the performance of different pupil groups. *British Educational Research Journal*, 23(4), 451-469. doi:10.1080/0141192970230405
- Tourón, J., López-González, E., Lizasoain, L., García-San Pedro, M. J., & Navarro, E. (2018). Alumnado español de alto y bajo rendimiento en ciencias en PISA 2015: Análisis del impacto de algunas variables de contexto. *Revista De Educación*, 380, 156-184. doi:0.4438/1988-592X-RE-2017-380-376
- Tröhler, D. (2013). The OECD and cold war culture: Thinking historically about PISA. In H. D. Meyer, & A. Benavot (Eds.), *PISA, power, and policy: The emergence of global educational governance* (pp. 141-161). Oxford, Reino Unido: Symposium books.
- Tsai, S., Smith, M. L., & Hauser, R. M. (2017). Families, schools, and student achievement inequality: A multilevel MIMIC model approach. *Sociology of Education*, 90(1), 64-88. doi:10.1177/0038040716683779
- United Nations Educational, Scientific and Cultural Organization (UNESCO). (2017). *Cracking the code: Girls' and women's education in science, technology, engineering and mathematics (STEM)*. París, Francia: UNESCO.

- Van de Werfhorst, H. G., & Mijs, J. J. (2010). Achievement inequality and the institutional structure of educational systems: A comparative perspective. *Annual Review of Sociology*, 36, 407-428. doi:10.1146/annurev.soc.012809.102538
- Weber, G. (1971). *Inner-city children can be taught to read: Four successful schools*. Washington, Estados Unidos: Council for Basic Education.
- Willms, J. D. (2010). School composition and contextual effects on student outcomes. *Teachers College Record*, 112(4), 1008-1037.
- Windle, J. A. (2016). *Making sense of school choice: Politics, policies, and practice under conditions of cultural diversity*. Nueva York, Estados Unidos: Palgrave Macmillan.
- Wingersky, M., Kaplan, B., & Beaton, A. E. (1983). Joint estimation procedures. In A. E. Beaton (Ed.), *Implementing the new design: The NAEP 1983-84 technical report* (pp. 285-292). Princeton, Estados Unidos: Educational Testing Service.
- Wiseman, A. W. (2013). Policy responses to PISA in comparative perspective. In H. D. Meyer, & A. Benavot (Eds.), *PISA, power, and policy: The emergence of global educational governance* (pp. 303-322). Oxford, Reino Unido: Symposium Books.
- Woltman, H., Feldstain, A., MacKay, J. C., & Rocchi, M. (2012). An introduction to hierarchical linear modeling. *Tutorials in Quantitative Methods for Psychology*, 8(1), 52-69. doi:10.20982/tqmp.08.1.p052
- Woods-McConney, A., Oliver, M. C., McConney, A., Maor, D., & Schibeci, R. (2013). Science engagement and literacy: A retrospective analysis for indigenous and non-indigenous students in Aotearoa New Zealand and Australia. *Research in Science Education*, 43(1), 233-252. doi:10.1007/s11165-011-9265-y
- Wu, M. (2005). The role of plausible values in large-scale surveys. *Studies in Educational Evaluation*, 31(2-3), 114-128. doi:10.1016/j.stueduc.2005.05.005
- Yore, L. D., Anderson, J. O., & Chiu, M. (2010). Moving PISA results into the policy arena: Perspectives on knowledge transfer for future considerations and preparations. *International Journal of Science and Mathematics Education*, 8(3), 593-609. doi:10.1007/s10763-010-9211-x

Zhou, Y., Fan, X., Wei, X., & Tai, R. H. (2017). Gender gap among high achievers in math and implications for STEM pipeline. *The Asia-Pacific Education Researcher*, 26(5), 259-269. doi:10.1007/s40299-017-0346-1

Zhu, Y., Kaiser, G., & Cai, J. (2018). Gender equity in mathematical achievement: The case of China. *Educational Studies in Mathematics*, 99(3), 245-260. doi:10.1007/s10649-018-9846-z



Anexos

ANEXO I: Artículos incluidos en la Revisión Sistemática de la Literatura

Año	Autores	Título	Revista
2018	Rodriguez-Santero, Javier; Gil-Flores, Javier	Contextual variables associated with differences in educational performance between European Union countries	Cultura y Educación
2018	Zhu, Yan; Kaiser, Gabriele; Cai, Jinfa	Gender equity in mathematical achievement: the case of China	Educational Studies in Mathematics
2018	Han S.W.	School-based teacher hiring and achievement inequality: A comparative perspective	International Journal of Educational Development
2018	Chiao C., Chiu C.-H.	The Mediating Effect of ICT Usage on the Relationship Between Students' Socioeconomic Status and Achievement	Asia-Pacific Education Researcher
2018	Barnard-Brak L., Lan W.Y., Yang Z.	Differences in mathematics achievement according to opportunity to learn: A 4pL item response theory examination	Studies in Educational Evaluation
2018	Tourón J., López-González E., Lizasoain Hernández L., García San Pedro M.J., Navarro Asencio E.	Spanish High and Low achievers in Science in PISA 2015: Impact analysis of some contextual variables [Alumnado español de alto y bajo rendimiento en ciencias en PISA 2015: Análisis del impacto de algunas variables de contexto]	Revista de Educación
2018	Smith P., Cheema J., Kumi-Yeboah A., Warrican S.J., Alleyne M.L.	Language-based differences in the literacy performance of bidialectal youth	Teachers College Record
2018	Chi, S., Liu, X., Wang, Z., Won Han, S.	Moderation of the effects of scientific inquiry activities on low SES students' PISA 2015 science achievement by school teacher support and disciplinary climate in science classroom across gender	International Journal of Science Education
2018	Hwang, J., Choi, K.M., Bae, Y., Shin, D.H.	Do Teachers' Instructional Practices Moderate Equity in Mathematical and Scientific Literacy?: an Investigation of the PISA 2012 and 2015	International Journal of Science and Mathematics Education
2018	Lazarević, L.B., Orlić, A.	PISA 2012 Mathematics literacy in Serbia: A multilevel analysis of students and schools	Psihologija
2018	Sortkær, B., Reimer, D.	Classroom disciplinary climate of schools and gender—evidence from the Nordic countries	School Effectiveness and School Improvement

Año	Autores	Título	Revista
2017	Hu X., Gong Y., Lai C., Leung F.K.S.	The relationship between ICT and student literacy in mathematics, reading, and science across 44 countries: A multilevel analysis	Computers and Education
2017	Gabriel F., Signolet J., Westwell M.	A machine learning approach to investigating the effects of mathematics dispositions on mathematical literacy	International Journal of Research and Method in Education
2017	Meggiolaro S.	Information and communication technologies use, gender and mathematics achievement: evidence from Italy	Social Psychology of Education
2017	Rutkowski D., Rutkowski L., Wild J., Burroughs N.	Poverty and educational achievement in the US: A less-biased estimate using PISA 2012 data	Journal of Children and Poverty
2017	Kim S.	ICT for Children of Immigrants: Indirect and Total Effects via Self-Efficacy on Math Performance	Journal of Educational Computing Research
2017	Zhou Y., Fan X., Wei X., Tai R.H.	Gender Gap Among High Achievers in Math and Implications for STEM Pipeline	Asia-Pacific Education Researcher
2017	He J., Van de Vijver F.J.R., Kulikova A.	Country-level correlates of educational achievement: evidence from large-scale surveys	Educational Research and Evaluation
2017	Schachner M.K., He J., Heizmann B., Van de Vijver F.J.R.	Acculturation and school adjustment of immigrant youth in six European countries: Findings from the Programme for International Student Assessment (PISA)	Frontiers in Psychology
2017	Arikan S., van de Vijver F.J.R., Yagmur K.	PISA mathematics and reading performance differences of mainstream European and Turkish immigrant students	Educational Assessment, Evaluation and Accountability
2017	Gil-Flores J., García-Gómez S.	The importance of teaching practices in relation to regional educational policies in explaining PISA achievement [Importancia de la actuación docente frente a la política educativa regional en la explicación del rendimiento en PISA]	Revista de Educación
2017	Cheema, J.R.	Effect of math-specific self-efficacy on math literacy: Evidence from a Greek survey	Research in Education
2016	Eickelmann B., Gerick J., Koop C.	ICT use in mathematics lessons and the mathematics achievement of secondary school students by international comparison: Which role do school level factors play?	Education and Information Technologies

Año	Autores	Título	Revista
2016	Petko D., Cantieni A., Prasse D.	Perceived Quality of Educational Technology Matters: A Secondary Analysis of Students ICT Use, ICT-Related Attitudes, and PISA 2012 Test Scores	Journal of Educational Computing Research
2016	Cheung K.-C., Mak S.-K., Sit P.-S., Soh K.-C.	A typology of student reading engagement: Preparing for response to intervention in the school curriculum	Studies in Educational Evaluation
2016	Sebastian J., Huang H.	Examining the relationship of a survey based measure of math creativity with math achievement: Cross-national evidence from PISA 2012	International Journal of Educational Research
2015	Kalaycıoğlu D.B.	The influence of socioeconomic status, self-efficacy, and anxiety on mathematics achievement in England, Greece, Hong Kong, the Netherlands, Turkey, and the USA	Kuram ve Uygulamada Eğitim Bilimleri
2015	Lafontaine D., Baye A., Vieluf S., Monseur C.	Equity in opportunity-to-learn and achievement in reading: A secondary analysis of PISA 2009 data	Studies in Educational Evaluation
2015	Cordero Ferrera J.M	Success factors for educational attainment in unfavorable socioeconomic conditions [Factores del éxito escolar en condiciones socioeconómicas desfavorables]	Revista de Educación
2015	Ercikan K., Chen M.Y., Lyons-Thomas J., Goodrich S., Sandilands D., Roth W.-M., Simon M.	Reading Proficiency and Comparability of Mathematics and Science Scores for Students From English and Non-English Backgrounds: An International Perspective	International Journal of Testing
2015	Scheerens J., Luyten H., van den Berg S.M., Glas C.A.W.	Exploration of direct and indirect associations of system-level policy-amenable variables with reading literacy performance	Educational Research and Evaluation
2015	Chiu M.M.	Family inequality, school inequalities, and mathematics achievement in 65 countries: Microeconomic mechanisms of rent seeking and diminishing marginal returns	Teachers College Record
2014	Özdemir B., Gelbal S.	Investigating factors that affect Turkish students' academic success with canonical commonality analysis according to PISA 2009 results	Eğitim ve Bilim
2014	Acosta S.T., Hsu H.-Y.	Negotiating diversity: an empirical investigation into family, school and student factors influencing New Zealand adolescents' science literacy	Educational Studies

Año	Autores	Título	Revista
2014	Dumay X., Dupriez V.	Educational quasi-markets, school effectiveness and social inequalities	Journal of Education Policy
2014	Sandilands D., Barclay McKeown S., Lyons-Thomas J., Ercikan K.	An Investigation of School-Level Factors Associated With Science Performance for Minority and Majority Francophone Students in Canada	Canadian Journal of Science, Mathematics and Technology Education
2014	Täht K., Must O., Peets K., Kattel R.	Learning motivation from a cross-cultural perspective: A moving target?	Educational Research and Evaluation
2014	Acosta S., Hsu H.-Y.	Shared academic values: Testing a model of the association between Hong Kong parents' and adolescents' perception of the general value of science and scientific literacy	Educational Studies
2013	Perry L.B., McConney A.	School socioeconomic status and student outcomes in reading and mathematics: A comparison of Australia and Canada	Australian Journal of Education
2013	Cheung K.-C., Sit P.-S., Soh K.-C., Ieong M.-K., Mak S.-K.	Predicting Academic Resilience with Reading Engagement and Demographic Variables: Comparing Shanghai, Hong Kong, Korea, and Singapore from the PISA Perspective	Asia-Pacific Education Researcher
2013	Woods-McConney A., Oliver M.C., McConney A., Maor D., Schibeci R.	Science Engagement and Literacy: A Retrospective Analysis for Indigenous and Non-Indigenous Students in Aotearoa New Zealand and Australia	Research in Science Education
2012	Kim D.H., Law H.	Gender gap in maths test scores in South Korea and Hong Kong: Role of family background and single-sex schooling	International Journal of Educational Development
2011	Liu X., Whitford M.	Opportunities-to-Learn at Home: Profiles of Students With and Without Reaching Science Proficiency	Journal of Science Education and Technology
2011	Luu K., Freeman J.G.	An analysis of the relationship between information and communication technology (ICT) and scientific literacy in Canada and Australia	Computers and Education
2011	Condron D.J.	Egalitarianism and educational excellence: Compatible goals for affluent societies?	Educational Researcher
2010	McConney A., Perry L.B.	Socioeconomic status, self-efficacy, and mathematics achievement in Australia: A secondary analysis	Educational Research for Policy and Practice

Año	Autores	Título	Revista
2010	Perry L.B., McConney A.	Does the SES of the school matter? An examination of socioeconomic status and student achievement using PISA 2003	Teachers College Record
2010	Willms J.D.	School Composition and Contextual Effects on Student Outcomes	Teachers College Record
2010	Milford T., Ross S.P., Anderson J.O.	An opportunity to better understand schooling: The growing presence of pisa in The AMERICAS*	International Journal of Science and Mathematics Education
2010	McConney A., Perry L.B.	Science and mathematics achievement in Australia: The role of school socioeconomic composition in educational equity and effectiveness	International Journal of Science and Mathematics Education
2007	Maslowski R., Scheerens J., Luyten H.	The effect of school autonomy and school internal decentralization on students' reading literacy	School Effectiveness and School Improvement
2005	Güzel Ç.I., Berberoğlu G.	An analysis of the Programme for International Student Assessment 2000 (PISA 2000) mathematical literacy data for Brazilian, Japanese and Norwegian students	Studies in Educational Evaluation

ANEXO II: Ítems que componen los índices utilizados en la regresión logística

Tabla AII. 1. Variables a nivel de estudiante

Variable	Ítems que la componen	Escala
Motivación hacia el rendimiento	¿En qué medida estás de acuerdo?: Quiero sacar las notas más altas en todas o casi todas las asignaturas. Quiero poder elegir entre las mejores oportunidades disponibles cuando me titule. Quiero ser el mejor, haga lo que haga. Me considero una persona ambiciosa. Quiero ser uno de los mejores alumnos de la clase.	1-4
Disposición al trabajo colaborativo	¿En qué medida estás de acuerdo?: Sé escuchar. Me gusta ver que mis compañeros de clase tienen éxito. Tengo en cuenta los intereses de los demás. Me gusta tener en cuenta distintos puntos de vista	1-4
Valor del trabajo colaborativo	¿En qué medida estás de acuerdo?: Prefiero más trabajar en equipo que solo. Creo que se toman mejores decisiones en grupo que de forma individual. Creo que el trabajo en equipo aumenta mi propio rendimiento. Me gusta colaborar con mis colegas.	1-4
Autoeficacia en ciencias	¿Serías capaz de realizar sin ayuda las siguientes tareas?: Reconocer el problema científico que subyace en un artículo periodístico sobre un tema de salud. Explicar por qué ocurren terremotos con más frecuencia en unas zonas que en otras. Describir la función de los antibióticos en el tratamiento de una enfermedad. Identificar el problema científico asociado a la eliminación de basuras. Predecir en qué medida los cambios medioambientales afectarán a la supervivencia de determinadas especies. Interpretar la información científica contenida en las etiquetas de los productos alimenticios. Debatir si la aparición de nuevos datos puede hacerte cambiar de opinión sobre que haya vida en Marte. Identificar la mejor de dos explicaciones sobre la formación de la lluvia ácida.	1-4
Creencias epistemológicas	¿En qué medida estás de acuerdo?: Una buena manera de saber si algo es cierto es haciendo un experimento. A veces los conceptos científicos cambian. Las buenas respuestas se basan en los resultados de muchos experimentos diferentes. Es bueno repetir los experimentos más de una vez para asegurarse de los resultados. A veces, los científicos cambian de opinión sobre lo que la ciencia considera como cierto. A veces, los conceptos científicos que figuran en los libros de ciencias cambian.	1-4

Variable	Ítems que la componen	Escala
Actividades relacionadas con las ciencias	¿Con qué frecuencia haces lo siguiente?: Ver programas científicos en televisión. Comprar o tomar prestados libros sobre temas científicos. Visitar páginas web dedicadas a temas científicos. Leer revistas científicas o artículos de periódico sobre ciencia. Asistir a un taller sobre ciencia. Simular fenómenos naturales en programas de ordenador/laboratorios virtuales. Visitar páginas web de organizaciones ecológicas. Seguir noticias de ciencia, medioambientales o de organizaciones ecológicas por blogs y <i>nanoblogging</i> .	1-4
Disfrute por la ciencia	¿En qué medida estás de acuerdo?: Por lo general me divierto aprendiendo temas científicos. Me gusta leer sobre ciencia. Estoy contento trabajando temas científicos. Lo paso bien adquiriendo nuevos conocimientos científicos. Me interesa aprender cosas de ciencia.	1-4
Interés por temas científicos	¿Hasta qué punto te interesan los siguientes temas científicos?: La biosfera (p. ej. los ecosistemas, la sostenibilidad. El movimiento y las fuerzas (p. ej. la velocidad, el rozamiento, las fuerzas magnéticas y gravitatorias). La energía y su transformación (p. ej. la conservación, las reacciones químicas). El universo y su historia. La forma en que la ciencia nos puede ayudar a prevenir las enfermedades.	1-4
Clima de disciplina en el aula de ciencias	¿Con qué frecuencia ocurre esto?: Los alumnos no atienden a lo que dice el profesor. Hay ruido y falta de orden. El profesor tiene que esperar mucho rato a que los alumnos se callen. Los alumnos no pueden trabajar bien. Los alumnos no empiezan a trabajar hasta mucho después de comenzada la clase.	1-4
Enseñanza de ciencias basada en la investigación	¿Con qué frecuencia ocurre esto?: A los alumnos se les da la oportunidad de exponer sus ideas. Los alumnos pasan tiempo en el laboratorio realizando experimentos prácticos. Se les pide a los alumnos que debatan sobre cuestiones científicas. Se les pide a los alumnos que saquen conclusiones del experimento que han realizado. El profesor explica cómo un mismo principio puede aplicarse a varios fenómenos diferentes. A los alumnos se les permite diseñar sus propios experimentos. Hay debates en clase sobre investigaciones. El profesor explica con claridad la importancia de los conceptos científicos en la vida de las personas. Se pide a los alumnos	1-4

Variable	Ítems que la componen	Escala
	que hagan una investigación para comprobar ciertos conceptos.	
Apoyo del profesor en clase de ciencias	¿Con qué frecuencia ocurre esto?: El profesor muestra interés por el aprendizaje de cada alumno. El profesor proporciona ayuda extra cuando los alumnos la necesitan. El profesor ayuda a los alumnos con su aprendizaje. El profesor continúa explicando hasta que los alumnos lo comprenden. El profesor da a los alumnos la oportunidad de expresar sus opiniones.	1-4
Enseñanza de ciencias dirigida por el profesor	¿Con qué frecuencia ocurre esto?: El profesor explica conceptos científicos. Se lleva a cabo un debate entre toda la clase y el profesor. El profesor explica nuestras preguntas. El profesor demuestra un concepto.	1-4
Retroalimentación percibida en clase de ciencias	¿Con qué frecuencia ocurre esto?: El profesor me informa de mi rendimiento en esta asignatura. El profesor me informa de mis puntos fuertes en esta asignatura de ciencias. El profesor me dice en qué áreas aún puedo mejorar. El profesor me dice cómo puedo mejorar mi rendimiento. El profesor me aconseja acerca de cómo lograr mis objetivos.	1-4
Adaptación de la instrucción	¿Con qué frecuencia ocurre esto?: El profesor adapta la lección a las necesidades y conocimientos de la clase. El profesor ayuda individualmente a los alumnos cuando tienen dificultades para comprender un tema o ejercicio. El profesor modifica la estructura de la clase cuando el tema resulta difícil de entender para la mayoría de los alumnos.	1-4
Motivación instrumental	¿En qué medida estás de acuerdo?: Merece la pena esforzarse en las asignaturas de ciencias porque eso me servirá para la profesión que quiero ejercer más adelante. Lo que aprendo en las asignaturas de ciencias es importante para mí porque lo necesito para lo que quiera hacer después. Merece la pena estudiar las asignaturas de ciencias porque lo que aprendo mejorará las perspectivas de mi carrera profesional. Muchas cosas que estudio en las asignaturas de ciencias de ayudarán a encontrar trabajo.	1-4
Conciencia medioambiental	¿Estás informado sobre los siguientes temas medioambientales?: El aumento de gases de efecto invernadero en la atmósfera. El empleo de organismos modificados genéticamente. Los residuos nucleares. Las consecuencias de la tala de bosques para dar otro uso al suelo.	1-4

Variable	Ítems que la componen	Escala
	La contaminación del aire. La extinción de plantas y animales. La escasez de agua.	
Optimismo medioambiental	¿Crees que los problemas medioambientales descritos a continuación mejorarán o empeorarán en los próximos 20 años?: El aumento de gases de efecto invernadero en la atmósfera. El empleo de organismos modificados genéticamente. Los residuos nucleares. Las consecuencias de la tala de bosques para dar otro uso al suelo. La contaminación del aire. La extinción de plantas y animales. La escasez de agua.	Mejorará, Quedará igual, Empeorará
Pertenencia al centro	¿En qué medida estás de acuerdo?: Me siento marginado en el centro. Hago amigos fácilmente en el centro. Me siento integrado en el centro. Me siento incómodo y fuera de lugar en mi centro. caigo bien a otros alumnos. Me siento solo en el centro.	1-4
Apoyo familiar	¿En qué medida estás de acuerdo?: Mis padres se interesan por mis actividades escolares. Mis padres apoyan mis esfuerzos y logros educativos. Mis padres me apoyan cuando me enfrento a dificultades escolares. Mis padres me animan para que tenga confianza en mí mismo.	1-4
Imparcialidad del profesorado	¿Con qué frecuencia ocurre esto?: Los profesores se relacionan conmigo menos frecuentemente que con otros compañeros. Los profesores son más duros conmigo en sus evaluaciones que con otros compañeros. Me da la impresión de que los profesores piensan que soy menos inteligente de lo que realmente soy. Los profesores son más disciplinarios conmigo que con otros compañeros. Los profesores me ridiculizan delante de los compañeros. Los profesores me dicen cosas insultantes delante de otros compañeros	1-4
Competencia en el uso de las TIC	¿En qué medida estás de acuerdo?: Me encuentro cómodo utilizando aparatos digitales con los que estoy más o menos familiarizado. Si mis amigos y familiares quieren comprar nuevos aparatos o aplicaciones digitales, puedo aconsejarlos. Me encuentro cómodo utilizando mis aparatos digitales en casa. Cuando me encuentro con problemas con los aparatos digitales, creo que puedo solucionarlos. Si mis amigos y familiares tienen un problema con los aparatos digitales, puedo ayudarlos.	1-4

Variable	Ítems que la componen	Escala
Interés en el uso de las TIC	¿En qué medida estás de acuerdo?: Me olvido del tiempo cuando utilizo aparatos digitales. Internet es un gran recurso para obtener la información que me interesa (p. ej. noticias, deportes, diccionarios). Es muy útil contar con redes sociales en internet. Me emociona mucho descubrir nuevas aplicaciones o aparatos digitales. Me molesta mucho que no haya conexión a internet. Me gusta usar aparatos digitales.	1-4
Autonomía en el uso de las TIC	¿En qué medida estás de acuerdo?: Si necesito nuevos programas informáticos, los instalo yo mismo. Leo información sobre aparatos digitales para ser autónomo. Uso aparatos digitales porque quiero usarlos. Si tengo un problema con los aparatos digitales comienzo a resolverlo por mí mismo. Si necesito una nueva aplicación, la elijo yo mismo.	1-4
TIC como tema de interacción social	¿En qué medida estás de acuerdo?: Para aprender algo nuevo sobre dispositivos digitales, me gusta hablar sobre ellos con mis amigos. Me gusta intercambiar soluciones a los problemas con dispositivos digitales con otra gente por internet. Me gusta quedar con mis amigos y jugar a videojuegos con ellos. Me gusta compartir información sobre dispositivos digitales con mis amigos. Aprendo mucho sobre dispositivos digitales a través de discusiones con mis amigos y familiares.	1-4

Tabla AII.2. Variables a nivel de centro educativo

Variable	Ítems que la componen	Escala
Liderazgo educativo	¿Con qué frecuencia ocurre esto?: Utilizo los resultados de los alumnos para elaborar los objetivos educativos del centro. Me ocupo de que las actividades para la formación permanente del personal docente serán acordes con los objetivos académicos del centro. Me aseguro de que el personal docente trabaje siguiendo los objetivos educativos del centro. Promuevo prácticas docentes basadas en investigaciones educativas recientes. Alabo a los profesores cuyos alumnos participan activamente en el aprendizaje. Cuando un profesor tiene problemas en su clase, tomo la iniciativa para tratar el asunto. Hago notar al profesorado la importancia del desarrollo de las capacidades críticas y sociales del alumnado. Presto atención a los	1-6

Variable	Ítems que la componen	Escala
	comportamientos que perturban las clases. Procuero al personal la ocasión de participar en la toma de decisiones del centro. Implico al profesorado en la construcción de una cultura escolar de mejora continua. Pido al profesorado que participe en la evaluación de las prácticas de gestión. Cuando un profesor saca a la luz un problema de clase, lo resolvemos juntos. Hablo de los objetivos académicos del centro con el profesorado en los claustros.	
Desarrollo del currículo	¿Con qué frecuencia ocurre esto?: Utilizo los resultados de los alumnos para elaborar los objetivos educativos del centro. Me ocupo de que las actividades para la formación permanente del personal docente serán acordes con los objetivos académicos del centro. Me aseguro de que el personal docente trabaje siguiendo los objetivos educativos del centro. Hablo de los objetivos académicos del centro con el profesorado en los claustros.	1-6
Liderazgo instruccional	¿Con qué frecuencia ocurre esto?: Promuevo prácticas docentes basadas en investigaciones educativas recientes. Alabo a los profesores cuyos alumnos participan activamente en el aprendizaje. Hago notar al profesorado la importancia del desarrollo de las capacidades críticas y sociales del alumnado	1-6
Desarrollo profesional	¿Con qué frecuencia ocurre esto?: Cuando un profesor tiene problemas en su clase, tomo la iniciativa para tratar el asunto. Presto atención a los comportamientos que perturban las clases. Cuando un profesor saca a la luz un problema de clase, lo resolvemos juntos.	1-6
Participación del profesorado	¿Con qué frecuencia ocurre esto?: Procuero al personal la ocasión de participar en la toma de decisiones del centro. Implico al profesorado en la construcción de una cultura escolar de mejora continua. Pido al profesorado que participe en la evaluación de las prácticas de gestión.	1-6
Responsabilidad del centro sobre los recursos	¿Quién toma las siguientes decisiones?: Contratación de docentes. Despido de docentes. Salario inicial de los docentes. Subidas de salario para los docentes. Elaborar	Director, Profesores, Equipo Directivo,

Variable	Ítems que la componen	Escala
	el presupuesto del centro. Decidir en qué se emplea el presupuesto dentro del centro.	Consejería, Ministerio
Responsabilidad del centro sobre el currículo	¿Quién toma las siguientes decisiones?: Establecer las políticas de evaluación de los estudiantes. Elegir los libros de texto a utilizar. Determinar los contenidos curriculares. Decidir qué materias se ofertan.	Director, Profesores, Equipo Directivo, Consejería, Ministerio
Actividades extraescolares de carácter creativo	¿Se oferta esta actividad en su centro?: Banda, orquesta o coro. Teatro. Anuario o periódico escolar. Voluntariado. Club de ciencias. Competiciones científicas. Club de ajedrez. TIC. Arte. Deportes. Seminarios o conferencias de agentes externos al centro.	Sí/No

ANEXO III. Documentos presentados en los centros durante la fase cualitativa

1. Acuerdo de confidencialidad entre el Equipo Directivo y el Grupo de Investigación
2. Permiso de recogida de datos de los estudiantes, para las familias
3. Carta informativa para los grupos de discusión con Equipos Directivos
4. Carta informativa para los grupos de discusión con docentes

ACUERDO DE CONFIDENCIALIDAD ENTRE EL EQUIPO INVESTIGADOR Y EL CENTRO EDUCATIVO

En Salamanca, a de de 2018.

REUNIDOS

De una parte, María José Rodríguez Conde, DNI 22728082B, Catedrática de Universidad de Métodos de Investigación y Diagnóstico de la Universidad de Salamanca, Directora del Instituto Universitario de Ciencias de la Educación, como Investigadora Principal (IP) del Proyecto I+D titulado *Evaluación de impacto del desarrollo de competencias básicas sobre el rendimiento académico en educación secundaria: propuesta de formación e innovación docente* (Ref: EDU2015-64524-P²³) y representante del Equipo Investigador de dicho Proyecto.

De otra parte, director del centro educativo, en representación del Centro.

MANIFIESTAN

I.- Que el Equipo Investigador del Proyecto I+D EDU2015-64524-P y el centro educativo “.....” van a colaborar en el marco del Proyecto I+D mencionado anteriormente.

II.- Que, en este marco, ambas partes van a intercambiar información susceptible de ser protegida mediante condiciones de confidencialidad.

Para proteger toda esta información y salvaguardar su carácter confidencial, las partes, de común acuerdo, formalizan el presente Acuerdo de Confidencialidad con sujeción a las siguientes,

CLAUSULAS

1. Ambas partes se obligan a mantener, en todo momento, estrictamente confidencial cualquier dato que pueda contribuir a la identificación del centro educativo, como su nombre, el nombre de alguno de sus integrantes, o la localidad en la que se encuentra (de aquí en adelante, se referirá a esta información como INFORMACIÓN CONFIDENCIAL). En todas las publicaciones dirigidas a la divulgación del protocolo o los resultados de la investigación a la que se refiere el presente acuerdo, el centro educativo será identificado con un código alfanumérico y situado en “una capital provincial de la Comunidad Autónoma de Castilla y León”, sin especificar la localidad.
2. El Equipo Investigador del Proyecto I+D EDU2015-64524-P se compromete a entregar al centro educativo los documentos especificados a continuación con el objetivo de contribuir a su mejora:
 - a. Informe de resultados de la investigación llevada a cabo en el centro educativo
 - b. Propuesta de formación adaptada a las necesidades detectadas en el centro

²³ RESOLUCIÓN DE 6 DE MAYO DE 2016, de la Secretaría de Estado de Investigación, Desarrollo e Innovación (ministerio de Economía y Competitividad, entonces) por la que se conceden ayudas correspondientes al Programa Estatal de Fomento de la Investigación Científica y Técnica de Excelencia, Subprograma Estatal de Generación de Conocimiento, convocatoria 2015, modalidad 1: Proyectos de I+D, en el marco del Plan Estatal de Investigación Científica y Técnica y de Innovación 2013-2016.

3. La INFORMACION CONFIDENCIAL será utilizada exclusivamente por el Equipo Investigador en el ámbito de la COLABORACION. Se entenderá por COLABORACIÓN la investigación realizada en el marco del Proyecto I+D EDU2015-64524-P y en aquellos proyectos posteriores que pudiesen estar relacionados con este. El Equipo Investigador se compromete a no revelar, comunicar ni de cualquier otro modo poner en conocimiento de terceros la INFORMACION CONFIDENCIAL. Del mismo modo, el centro educativo se compromete a mantener confidencial cualquier información y resultados adquiridos o derivados de la COLABORACION (informe de resultados del centro), siendo cualesquiera resultados derivados de la misma, propiedad exclusiva del Equipo Investigador a todos los efectos de uso o divulgación.
4. El Equipo Investigador se compromete a que los informes entregados al centro educativo, así como las publicaciones de divulgación científica que pudiesen derivarse de la investigación, se abstendrán de identificar personalmente a ninguno de los informantes, presentando en su lugar la información agregada por colectivos (estudiantes, docentes, equipo directivo).
5. La INFORMACION CONFIDENCIAL sólo será puesta en conocimiento de aquellos empleados o colaboradores del Equipo Investigador que necesiten conocer dicha información, y, estos empleados/colaboradores, deberán ser informados, previamente, por el responsable del Equipo Investigador del carácter confidencial de dicha INFORMACION CONFIDENCIAL.
6. Las partes convienen que, en caso de que el Equipo Investigador incumpla parcial o totalmente con las obligaciones a su cargo derivadas del presente contrato, será responsable de los daños y perjuicios que dicho incumplimiento llegase a ocasionar al centro educativo.

Y para que conste y en prueba de conformidad, lo firman ambas partes, por duplicado en el lugar y fecha señalados en el encabezamiento

Dña. MARÍA JOSÉ RODRÍGUEZ CONDE

IP del Proyecto I+D EDU2015-64524-P

D.

Director del Centro

AUTORIZACIÓN PARA LA RECOGIDA DE DATOS DE MENORES DE EDAD

DATOS DEL PADRE, MADRE O TUTOR/A LEGAL DEL MENOR

NOMBRE..... APELLIDOS.....

MANIFIESTA

I.- Que autoriza a su hijo/a (NOMBRE Y APELLIDOS DEL MENOR)..... a participar en la recogida de datos relativa al Proyecto I+D+i *Evaluación de impacto del desarrollo de competencias básicas sobre el rendimiento académico en educación secundaria: propuesta de formación e innovación docente* (Ref: EDU2015-64524-P²⁴) que el personal del Instituto Universitario de Ciencias de la Educación (Universidad de Salamanca) va a llevar a cabo en su centro educativo durante el mes de en las condiciones abajo descritas.

CLAUSULAS

7. El Equipo Investigador a cargo del Proyecto EDU2015-64524-P no recogerá, en ningún caso, datos de carácter personal que permitan identificar al menor en el curso de la investigación (nombre, DNI, etc.), ni datos relativos a su familia (nivel socioeconómico, ocupación de sus familiares, etc.), en cumplimiento del Real Decreto 1720/2007, de 21 de diciembre, por el que se aprueba el Reglamento de desarrollo de la Ley Orgánica 15/1999, de 13 de diciembre, de protección de datos de carácter personal, y del Reglamento General de Protección de Datos.
8. La presente autorización se circunscribe a la recogida de datos relativa tres ámbitos:
 - a. Información personal de interés para la investigación (género, curso, país de origen, uso de nuevas tecnologías en casa y en el centro educativo) e historial académico del alumno/a (notas en anteriores evaluaciones, si ha repetido algún curso).
 - b. Nivel de Competencia Informacional del alumno/a.
 - c. Percepción del alumno/a sobre la Convivencia y el Clima Social del Aula en la que se encuentra.
9. Entendemos por Competencia Informacional la “habilidad para buscar, obtener, procesar y comunicar información, y para transformarla en conocimiento. Incorpora diferentes habilidades, que van desde el acceso a la información hasta su transmisión en distintos soportes una vez tratada, incluyendo la utilización de las Tecnologías de la Información y la Comunicación como elemento esencial para informarse, aprender y comunicarse”.
10. Entendemos por Clima Social del Aula aquellos aspectos relacionados con la relación entre los integrantes del aula (respeto, comunicación, cooperación, cohesión del grupo), la organización y las prácticas de disciplina que se aplican en el aula.

En Salamanca, a de de 2018.

Firmado

²⁴ RESOLUCIÓN DE 6 DE MAYO DE 2016, de la Secretaría de Estado de Investigación, Desarrollo e Innovación (ministerio de Economía y Competitividad, entonces) por la que se conceden ayudas correspondientes al Programa Estatal de Fomento de la Investigación Científica y Técnica de Excelencia, Subprograma Estatal de Generación de Conocimiento, convocatoria 2015, modalidad 1: Proyectos de I+D, en el marco del Plan Estatal de Investigación Científica y Técnica y de Innovación 2013-2016.



Proyecto I+D+i EDU2015-64524-P (“Evaluación de impacto del desarrollo de competencias básicas sobre el rendimiento académico en educación secundaria: propuesta de formación e innovación docente”)

GRUPO DE DISCUSIÓN (Equipo Directivo):

Destinado a: XXXX

Reunión:

Día: XXXXX

Lugar: XXXXX

Asistentes:

XXXXXXXXXX

OBJETIVO DE LA REUNIÓN:

Recoger la opinión del equipo directivo sobre aquellos factores que, en opinión de aquellos que ejercen el liderazgo del centro, promueven el aprendizaje de los alumnos en este centro.

ORGANIZACIÓN DE LA SESIÓN:

1º. Recordar con el equipo la finalidad del estudio (2-3 min) y las fases en las que lo estamos desarrollando. Hay que recordar que necesitaremos entregarles en junio, los informes de los resultados de los alumnos y que ellos decidan si quieren trabajar en una de las competencias en formación de profesorado que les ofrecemos (competencia informacional o competencias para la resolución de conflictos). Según los datos recogidos parece que la primera, responde más a necesidades del centro.

2º. Ronda de 2-3 minutos por persona del equipo, además de presentarse (formación previa, especialidad y curso que imparten, años de experiencia profesional y años de experiencia en el centro, etc.), expongan libremente su opinión sobre los factores que cada uno considera destacables “**en este centro concreto**”, que promueven el aprendizaje y desarrollo competencial de sus alumnos (factores de tipo organizacional, metodológico, de clima de relaciones profesorado, familias y alumnado, de recursos personales y materiales, de formación e innovación docentes, etc.). (Total:10 minutos)

3º. Cuestiones propuestas por la entrevistadora a debatir por el equipo: (25-30 min.)

3.1.- Sobre la **formación e innovación docente** en el centro.

3.2.- **Clima** del centro y del aula. Relaciones entre los alumnos, los alumnos y el profesorado, el profesorado entre sí y la relación con las familias (resolución de conflictos, convivencia...). **relación con organismos externos.**

3.3.- **Implicación, compromiso y trabajo en equipo de los docentes y del liderazgo.** Reparto de responsabilidades y roles de liderazgo, participación de los docentes en las tomas de decisiones.

3.4.- **Evaluación de programas, proyectos y profesorado.**

3.5.- Sobre la **metodología didáctica** y **atención a la diversidad.** Gestión del **tiempo** de aula. Recursos e instalaciones.

3.6.- Competencias digitales y resultados de aprendizaje. En concreto, **competencia informacional.**

4º. Debate abierto a quien quiera intervenir, si quiere incidir en algo importante que para él caracteriza este centro (Total: 5 min.)

TOTAL TIEMPO DE REUNIÓN: aprox. 45 minutos. Se grabará en audio, para registro objetivo de la información. La aceptación de participación en la reunión va asociada al permiso para recoger la información y poder tratarla de manera confidencial por el grupo de investigación.

- Muchas gracias por tu participación -



Proyecto I+D+i EDU2015-64524-P (“Evaluación de impacto del desarrollo de competencias básicas sobre el rendimiento académico en educación secundaria: propuesta de formación e innovación docente”)

GRUPO DE DISCUSIÓN (Profesorado):

Destinado a: XXXX

Reunión:

Día: XXXXX

Lugar: XXXXX

Asistentes:

XXXXXXXXXX

OBJETIVO DE LA REUNIÓN:

Recoger la opinión del profesorado sobre aquellos factores que, en opinión de aquellos que ejercen el liderazgo del centro, promueven el aprendizaje de los alumnos en este centro.

ORGANIZACIÓN DE LA SESIÓN:

1º. Recordar con el equipo la finalidad del estudio (2-3 min) y las fases en las que lo estamos desarrollando. Hay que recordar que necesitaremos entregarles en junio, los informes de los resultados de los alumnos y que ellos decidan si quieren trabajar en una de las competencias en formación de profesorado que les ofrecemos (competencia informacional o competencias para la resolución de conflictos). Según los datos recogidos parece que la primera, responde más a necesidades del centro.

2º. Ronda de 2-3 minutos por persona del equipo, además de presentarse (formación previa, especialidad y curso que imparten, años de experiencia profesional y años de experiencia en el centro, etc.), expongan libremente su opinión sobre los factores que cada uno considera destacables “**en este centro concreto**”, que promueven el aprendizaje y desarrollo competencial de sus alumnos (factores de tipo organizacional, metodológico, de clima de relaciones profesorado, familias y alumnado, de recursos personales y materiales, de formación e innovación docentes, etc.). (Total:10 minutos)

3º. Cuestiones propuestas por la entrevistadora a debatir por el equipo: (25-30 min.)

3.1.- Sobre la **formación e innovación docente** en el centro.

3.2.- **Clima** del centro y del aula. Relaciones entre los alumnos, los alumnos y el profesorado, el profesorado entre sí y la relación con las familias (resolución de conflictos, convivencia...). **relación con organismos externos.**

3.3.- **Implicación, compromiso y trabajo en equipo de los docentes y del liderazgo.** Reparto de responsabilidades y roles de liderazgo, participación de los docentes en las tomas de decisiones.

3.4.- **Expectativas del profesor** sobre el progreso de sus estudiantes.

3.5.- Sobre la **metodología didáctica** y **atención a la diversidad.** Gestión del **tiempo** de aula. Recursos e instalaciones.

3.6.- Competencias digitales y resultados de aprendizaje. En concreto, **competencia informacional.**

4º. Debate abierto a quien quiera intervenir, si quiere incidir en algo importante que para él caracteriza este centro (Total: 5 min.)

TOTAL TIEMPO DE REUNIÓN: aprox. 45 minutos. Se grabará en audio, para registro objetivo de la información. La aceptación de participación en la reunión va asociada al permiso para recoger la información y poder tratarla de manera confidencial por el grupo de investigación.

- Muchas gracias por tu participación -