

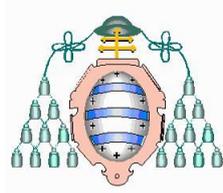
Universidad de Oviedo
Departamento de Psicología

TESIS DOCTORAL

Modelos Multinivel de los Factores de Eficacia Escolar en el Programa PISA

Elsa Peña Suárez

2011



Universidad de Oviedo
Departamento de Psicología

TESIS DOCTORAL

Modelos Multinivel de los Factores de Eficacia Escolar en el Programa PISA

Elsa Peña Suárez

Director: José Muñiz Fernández

Universidad de Oviedo 2011

A mis padres

Agradecimientos

La elaboración de la tesis conlleva un tiempo vital compartido con muchas personas relevantes para mí, bien porque suponen un apoyo emocional o bien porque son personas que han sido una influencia profesional y una suerte trabajar con ellas.

Gracias a todos, y muy especialmente a mi director, José Muñiz porque para mí es una persona que le tengo una gran admiración y un modelo profesional. El hecho de dirigirme la tesis, así como facilitarme todos los recursos necesarios y difusión de resultados supone un agradecimiento muy especial.

También quiero expresar mi gratitud a varias personas: Mis padres a los que les dedico la tesis y que han sido siempre fuente de recursos, afectos y apoyo incondicional; mi pareja Carlos porque siempre ha estado a la altura de las circunstancias, dándome todo su cariño, apoyo incondicional y haciendo posible ver las cosas de la forma más fácil y positiva posible; .mi abuela materna que aunque sólo esté en el recuerdo, sé que siempre me ha dado y me está dando todo su ánimo y a mi amiga Miriam que como ella sabe es como una buena hermana mayor .

En este apartado además quiero mencionar a mis profesores de metodología porque siempre se han interesado y me han animado en este proceso de elaboración. Especialmente a Ángel Fidalgo porque gracias a él participé en un proyecto de la Consejería de Educación; a Guillermo Vallejo por su interés; a Miguel Gallegos por su amable ayuda.

Por último gracias a mis compañeros de la Consejería de Educación y actualmente en la Facultad de Psicología porque a través de muchos de ellos he adquirido muchos conocimientos e, incluso, de algunos también tengo su amistad.

Financiación

El presente trabajo es una investigación financiada por el Vicerrectorado de Investigación de la Universidad de Oviedo. **Ref.** UNOV-08-BEDOC

ÍNDICE

PARTE TEÓRICA

Introducción.....	15
Capítulo 1. Movimiento de la Eficacia Escolar	
1. Movimiento de la Eficacia Escolar.....	19
1.1 Definición de Escuela Eficaz.....	19
2. Perspectiva histórica de la Eficacia Escolar.....	20
2.1 Antecedentes.....	21
2.1.1 Informe Coleman.....	22
2.1.2 Reanálisis de Informe Coleman.....	23
2.1.3 Informe Plowden.....	24
2.2 Primeros estudios de Eficacia Escolar.....	25
2.3 Consolidación de la Investigación en Eficacia Escolar.....	26
2.3.1 School Social Systems and Student Achievement.....	26
2.3.2 Fifteen Thousand Hours.....	28
2.3.3 Centros públicos versus privados en Norteamérica.....	30
2.3.4 The Louisiana School Effectiveness Studies (LSES).....	31
2.3.5 Junior School Project (JSP).....	33
2.4 Estudios multinivel de Eficacia Escolar.....	36
2.4.1 Eficacia diferencial: un reanálisis de los datos del School Matter.....	37
2.4.2 Eficacia Docente.....	38
2.4.3 Características de centros de Primaria y la calidad de la educación.....	39
2.4.4 The Victorian Quality of Schools Project.....	40
2.4.5 Un estudio de centros franceses.....	41
2.4.6 Study of departmental differences in academic effectiveness.....	43
3. Otras líneas de actuación de la Eficacia Escolar.....	45
4. Conclusiones.....	46
5. Índice de tablas.....	49
6. Índice de figuras.....	49

Capítulo 2: Modelos de Eficacia Escolar

1. Efectos Escolares	52
2. Modelos de efectos escolares	54
2.1. El modelo integrado de Eficacia Escolar	55
2.2. Modelo QAIT/MACRO.....	57
2.3. Modelo de eficacia docente.....	59
2.4. Modelo de Eficacia de Departamento.....	60
2.5. Modelo de Eficacia Escolar de EGB	63
2.6. Modelo iberoamericano de Eficacia Escolar	64
2.7. Modelo dinámico y multidimensional de eficacia escolar	66
2.7.1. <i>Dinamismo</i>	67
2.7.2. <i>Configuración multidimensional</i>	67
2.7.3. <i>Factores de aula</i>	68
2.8. Modelo de Mejora de la Eficacia Escolar (ESI).....	71
3. Conclusiones	72
4. Índice tablas	72
5. Índice figuras	72

Capítulo 3: Propiedades de los efectos escolares

1. Propiedades de los efectos escolares	75
1.1. Magnitud de los efectos.....	75
1.2. Consistencia de los efectos escolares	76
1.2.1. Entre medidas cognitivas	77
1.2.2. Entre medidas no cognitivas	78
1.2.3. Entre medidas cognitivas y no cognitivas.....	78
1.3. Efectos diferenciales para distintos grupos de alumnos	79
1.3.1. Rendimiento previo.....	79
1.3.2. Indicadores socioeconómicos	80
1.3.3. Características personales del estudiante.....	81
1.3.4. Diferencias en la Eficacia Docente	82
1.4. Estabilidad de los efectos.....	82
1.5. Continuidad de los efectos.....	83
2. Modelos Multinivel	83
2.1. ¿Por qué Modelos Multinivel?.....	83
2.2. Notación.....	84
2.3. Pasos a seguir en un diseño multinivel	86
3. Conclusiones	87

Capítulo 4: Eficacia escolar en España

1.	Eficacia escolar en España.....	91
1.1.	Estudios empíricos.....	91
1.1.1.	La eficacia en la educación escolar.....	92
1.1.2.	Procesos funcionales y eficacia de la escuela. Un modelo causal 92	
1.1.3.	Modelo de evaluación externa de los centros escolares de EGB 93	
1.1.4.	Rendimiento en el medio rural.....	93
1.1.5.	Factores de eficacia escolar de los centros públicos de EGB.....	94
1.1.6.	Modelo explicativo sobre calidad y gestión de recursos en los centros escolares.....	94
1.1.7.	Estudio sobre la gestión de los recursos educativos.....	95
1.1.8.	Calidad de la educación y eficacia de los centros: estudio sobre la gestión de los recursos educativos.....	95
1.1.9.	Calidad, eficacia y clima en centros educativos.....	97
1.1.10.	Mejora de Eficacia Escolar: Capacidad de cambio y adaptación de los centros.....	97
1.1.11.	Mejora de la Eficacia Escolar en Iberoamérica.....	98
1.1.12.	Variables que predicen el rendimiento académico en ESO: un análisis longitudinal multinivel.....	100
1.2.	Influencia en Políticas educativas.....	100
1.2.1.	Estudios transnacionales.....	101
1.2.2.	Estudios nacionales.....	102
1.2.3.	Definición de la Evaluación de Diagnóstico.....	103
1.2.4.	Mecanismos de la Evaluación de Diagnóstico.....	104
1.2.5.	Resultados de la evaluación general de diagnóstico.....	104
2.	Conclusiones.....	105
3.	Índice de tablas.....	106
4.	Índice de figuras.....	106

Capítulo 5: Evaluación PISA

1.	Programa Internacional de Evaluación del Alumnado	109
2.	Competencia científica	110
2.1.	Organización del área de evaluación	111
2.1.1.	Situaciones y contexto	112
2.1.2.	Capacidades científicas	113
2.1.3.	Conocimiento científico	113
2.1.4.	Conocimiento de la ciencia	113
2.1.5.	Conocimiento acerca de la ciencia	114
2.1.6.	Actitudes hacia las ciencias	115
3.	Competencia Matemática	116
3.1.	Organización del área de evaluación	117
3.2.	Situaciones y contexto	118
3.3.	Contenido matemático	118
3.4.	Procesos matemáticos	119
4.	Competencia Lectora	120
4.1.	Organización del área de evaluación	120
4.1.1.	Situaciones de la lectura	120
4.1.2.	La estructura del texto	121
4.1.3.	Características de los ejercicios	122
5.	Características metodológicas en PISA	123
5.1.	Diseño muestral	123
5.1.1.	El método BRR	124
5.2.	Instrumentos de medida	127
5.2.1.	Redacción de ítems y estudios pilotos	127
5.2.2.	Pruebas aptitudinales	128
5.2.3.	Pruebas actitudinales	131
5.2.4.	Pruebas de contexto	131
5.3.	Construcción de escalas	133
5.4.	Construcción de índices de contexto	135
6.	Resultados de España en PISA	135
7.	Planteamiento de los objetivos	136
7.1.	Objetivo general	139
7.2.	Objetivos específicos	139
7.3.	Hipótesis	140
8.	Índice de tablas	141
9.	Índice de figuras	141

PARTE EMPÍRICA

Capítulo 6: Método

1.	Muestra	144
2.	Instrumentos y variables	145
2.1.	Cuestionario de estudiante.....	145
2.1.1.	Variables de contexto/entrada.....	146
2.1.2.	Variables de proceso.....	149
2.2.	Cuestionario de competencias.....	150
2.2.1.	Variables resultado.....	151
2.3.	Cuestionario de centro.....	152
2.3.1.	Variables de contexto/entrada.....	152
2.3.2.	Variables de proceso de nivel dos.....	153
3.	Procedimiento	154
4.	Análisis de datos	156
4.1.	Objetivo 1. Estimar el tamaño del efecto de centro.....	156
4.2.	Objetivo 2. Calcular la consistencia de los efectos escolares entre las distintas variables aptitudinales y actitudinales.....	157
4.3.	Objetivo 3. Determinar la eficacia diferencial de los centros.....	157
4.4.	Objetivo 4. Determinar la aportación del contexto y del proceso de enseñanza-aprendizaje en la adquisición de la competencia científica.....	158
4.5.	Objetivo 5. Elaborar un modelo de Eficacia Escolar adecuados a la situación social y educativa científica en educación secundaria española.....	159
5.	Resumen de tablas	160
6.	Resumen de figuras	160

Capítulo 7: Resultados	
1.	Magnitud de los efectos de centro163
2.	Consistencia de los efectos escolares165
3.	Eficacia diferencial de los centros166
4.	Aportación del contexto y del proceso de enseñanza-aprendizaje en la adquisición de la competencia científica168
4.1.	Nivel de alumnado169
4.2.	Nivel de centro.....173
5.	Modelo de eficacia escolar científica176
5.1.	Comprobación de supuestos178
6.	Resumen de tablas180
7.	Resumen de figuras180
Capítulo 8: Discusión	
1.	Discusión y conclusiones183
Referencias.....189	
Anexos216	

ÍNDICE DE TABLAS

Tabla 1: <i>Factores de éxito académico (Weber, 1971)</i> . Fuente: elaboración propia.....	25
Tabla 2: <i>Modelo de cinco factores de Edmonds (1979)</i> . Fuente: elaboración propia.....	26
Tabla 3: <i>Correlaciones significativas entre rendimiento y clima escolar (Brookover et al., 1979)</i> . Fuente: Murillo (2005).....	27
Tabla 4: <i>Mapa de variables de Fifteen Thousand Hours (Rutter, 1979)</i> . Fuente: Murillo, 2005.....	29
Tabla 5: <i>Estimación de la puntuación para centros públicos y pendientes estimadas para centros privados</i> . Fuente: Murillo (2005).....	30
Tabla 6: <i>Contraste entre características de eficacia del LSES</i> . Fuente: Teddlie y Stringfield (1993, p. 132).....	33
Tabla 7: <i>Mapa de variables de Junior School Project</i> . Fuente: Elaboración propia.....	34
Tabla 8: <i>Resumen de factores de eficacia escolar</i> . Fuente: Murillo (2005 p.212).	53
Tabla 9: <i>Factores asociados con el logro académico obtenidos en los tres estudios</i> . Fuente: Murillo (2008).....	63
Tabla 10: <i>Factores de de aula y dimensiones del modelo dinámico (Creemers y Kyriakides, 2006)</i> . Fuente: elaboración propia.....	69
Tabla 11: <i>Estudios de eficacia escolar en España</i> . Fuente: Murillo (2005).....	91
Tabla 12: <i>Conceptos que componen los factores del modelo</i> . Fuente: Murillo (2011) ..	99

Tabla 13: <i>Contexto de la evaluación de ciencias</i> . Fuente: OCDE (2006a).....	112
Tabla 14: <i>Capacidades científicas en PISA 2006</i> . Fuente: elaboración propia.	113
Tabla 15: <i>Conocimientos de la competencia científica en PISA 2006</i> . Fuente: elaboración propia.	114
Tabla 16: <i>Categorías del conocimiento acerca de las ciencias en PISA 2006</i> . Fuente: elaboración propia.	115
Tabla 17: <i>Áreas de evaluación de las actitudes en PISA 2006</i> . Fuente: elaboración propia.	116
Tabla 18: <i>Replicaciones del método BRR</i> . Fuente: OCDE (2006b).....	125
Tabla 19: <i>Distribución de cuadernillos en PISA 2006</i> . Fuente: OCDE (2006b).....	129
Tabla 20: <i>Clasificación de ítems por formato y capacidad en ciencias</i> . Fuente: OCDE (2006b).	130
Tabla 21: <i>Clasificación de ítems por formato y capacidad en matemáticas</i> . Fuente: OCDE (2006b).	131
Tabla 22: <i>Clasificación de ítems por formato y capacidad lectora</i> . Fuente: OCDE (2006b).....	131
Tabla 23: <i>Clasificación de ítems por actitudes científicas</i> . Fuente: OCDE (2006b).	131
Tabla 24: <i>Representación de los valores plausibles obtenidos en PISA 2006</i> . Fuente: elaboración propia.	135
Tabla 25: <i>Evolución de los resultados globales en las tres competencias en población española</i> . Fuente: elaboración propia. Fuente: IE (2010).....	136
Tabla 26: <i>Variables contexto/entrada de alumnado</i> . Fuente: elaboración propia.	147
Tabla 27: <i>Índices de posesiones y contexto del hogar</i> . Fuente: OCDE (2009a).....	148
Tabla 28: <i>Análisis de fiabilidad y validez del ESCS</i> . Fuente: elaboración propia.	149
Tabla 29: <i>Coefficientes de fiabilidad de los índices de proceso de nivel 1</i> . Fuente: elaboración propia.	149
Tabla 30: <i>Análisis Factorial Confirmatorio de los índices de proceso de nivel 1</i> . Fuente: elaboración propia.	150
Tabla 31: <i>Variables de contexto/entrada de nivel 2</i> . Fuente: elaboración propia.....	153
Tabla 32: <i>Variables de proceso de nivel 2</i> . Fuente: elaboración propia.	154

Tabla 33: <i>Estimación de los efectos escolares brutos y netos.</i> Fuente: elaboración propia.....	163
Tabla 34: <i>Rango de valores plausibles entre los que se sitúa la media de centro.</i> Fuente: elaboración propia.....	165
Tabla 35: <i>Correlaciones entre los residuales de las distintas competencias evaluadas.</i> Fuente: elaboración propia.....	166
Tabla 36: <i>Significación de pendientes aleatorias en variables de contexto.</i> Fuente: elaboración propia.....	167
Tabla 37: <i>Modelo de pendientes aleatorias.</i> Fuente: elaboración propia.....	168
Tabla 38: <i>Porcentajes de varianza explicada por los submodelos.</i> Fuente: elaboración propia.....	169
Tabla 39: <i>Porcentajes de varianza explicada por los submodelos de centro.</i> Fuente: elaboración propia.....	169
Tabla 40: <i>Submodelos que componen el nivel 1.</i> Fuente: elaboración propia.....	170
Tabla 41: <i>Modelo final de nivel 1.</i> Fuente: elaboración propia.....	171
Tabla 42: <i>Submodelos que componen el nivel 2.</i> Fuente: elaboración propia.....	174
Tabla 43: <i>Modelo final de nivel 2.</i> Fuente: elaboración propia.....	175
Tabla 44: <i>Modelo multinivel final.</i> Fuente: elaboración propia.....	176
Tabla 45: <i>Descriptivos de los términos de error por niveles.</i> Fuente: elaboración propia.....	178

ÍNDICES FIGURAS

Figura 1: <i>Diferentes propuestas de organización de la eficacia escolar.</i> Fuente: elaboración propia.....	21
Figura 2: <i>Modelo propuesto en el Informe Plowden (1967).</i> Fuente: Murillo (2005)....	24
Figura 3: <i>Modelo conceptual de Brookover et al. (1979).</i> Fuente: Murillo (2005).....	27
Figura 4: <i>Resultados de Victorian Quality of Schools Project.</i> Fuente: tomada de Murillo (2005).....	41
Figura 5: <i>Mapa de variables del estudio de Grisay (1996).</i> Fuente: Elaboración propia	42
Figura 6: <i>Síntesis de modelos de Eficacia Escolar.</i> Fuente tomada: Reynolds et al (1999).	48

Figura 7: <i>Modelo de Eficacia Escolar de Scheerens y Creemers (1989)</i> . Fuente: Murillo (2008 p.5).....	55
Figura 8: <i>Modelo de Eficacia escolar de Scheerens (1990 p.73)</i> . Fuente: Murillo (2008 p.7).....	56
Figura 9: <i>Modelo de eficacia escolar de Stringfield y Slavin (1992, p.122)</i> . Fuente: Murillo (2008 p.8).....	58
Figura 10: <i>Modelo de Eficacia Docente (Creemers, 1994)</i> . Fuente: Murillo (2008 p.10).....	59
Figura 11: <i>Modelo de Eficacia de Departamento (Sammons, Thomas y Mortimore, 1997)</i> . Fuente: Murillo (2008 p.12).....	60
Figura 12: <i>Modelo de Eficacia Escolar en Primaria</i> . Fuente: Murillo (2008 p.22).....	61
Figura 13: <i>Resumen de relaciones entre factores de proceso</i> . Fuente: Murillo (2008 p.25).....	62
Figura 14: <i>Modelo de Eficacia Escolar en Primaria</i> . Fuente: Murillo (2008 p.22).....	65
Figura 15: <i>Modelo de mejora de la eficacia escolar (ESI)</i> . Fuente: Creemers et al., (2007 p.830).....	70
Figura 16: <i>Modelo de eficacia escolar en Primaria CIDE (1995)</i> . Fuente: Muñoz-Repiso et al. (1995).....	96
Figura 17: <i>Modelo iberoamericano de mejora de la eficacia escolar</i> . Fuente: Murillo (2011 p.54).....	99
Figura 18: <i>Orden cronológico de evaluaciones internacionales</i> . Fuente: elaboración propia.....	102
Figura 19: <i>Síntesis de evaluaciones del sistema educativo español</i> . Fuente: IE (2007).....	103
Figura 20: <i>Elementos que componen la competencia científica</i> . Fuente: OCDE (2006a).....	112
Figura 21: <i>Elementos del área de matemáticas</i> . Fuente: OCDE (2006a).....	117
Figura 22: <i>Capacidades incluidas en la competencia de matemáticas</i> . Fuente: OCDE (2006a).....	119
Figura 23: <i>Subescalas de procesos que componen la competencia lectora</i> . Fuente: OCDE (2006a).....	122
Figura 24: <i>Representación de la participación de España en PISA 2006</i> . Fuente: IE (2007).....	144
Figura 25: <i>Distribución de centros según el tipo de población</i> . Fuente: elaboración propia.....	145

Figura 26: <i>Variación entre y dentro de los centros para el modelo de efectos brutos.</i>	
Fuente: elaboración propia.....	163
Figura 27: <i>Variación entre y dentro de los centros para el modelo de efectos netos.</i>	
Fuente: elaboración propia.....	165
Figura 28: <i>Distribución de los errores de nivel de alumnado.</i> Fuente: elaboración propia.....	179
Figura 29: <i>Distribución de los errores de nivel de centro.</i> Fuente: elaboración propia.....	179

Introducción

La repercusión que tiene el sistema educativo en el proceso de aprendizaje no se replantea en la actualidad, sin embargo hubo un tiempo que se puso en duda por la publicación del informe Coleman, donde se concluía que los efectos escolares ejercen una escasa influencia sobre el rendimiento de sus alumnos. Las reacciones a dicho informe fueron muy diversas y entre ellas destaca la creación de un movimiento teórico y práctico como es la línea de Investigación de Eficacia Escolar.

La finalidad de dicha línea es elaborar una teoría comprensiva capaz de integrar los elementos que ayuden a que el centro escolar sea eficaz. Dicho movimiento es el que más ha incidido en la mejora de la educación en los últimos años y ejerce una importante influencia en políticas educativas, a través de evaluaciones nacionales e internacionales sobre rendimiento académico. Y a su vez estas evaluaciones tienen una gran repercusión mediática porque es donde se fabrican los discursos que legitiman políticas y es donde se crean nuevos valores y visiones de influencia política, social y educativa.

El presente trabajo estudia de forma contextualizada el logro académico del alumnado de quince años en población española, a través del programa PISA (*Program International Student Assessment*) desarrollado por la Organización para la Cooperación y el Desarrollo Económico (OCDE). Dicho logro académico se analiza bajo una evaluación externa que se concreta en la medición de competencias. Una competencia es entendida como la capacidad de utilizar esos conocimientos aprendidos en situaciones de la vida real que deberían poseer todos los jóvenes a los quince años, cuando finalizan la educación obligatoria.

Uno de los aportes de este trabajo es la aplicación del conocimiento que ofrece la Investigación sobre Eficacia Escolar a la evaluación PISA 2006. Se aporta conocimiento al estado de la cuestión sobre las propiedades de los efectos escolares de dos formas. Por una parte se mide la magnitud, consistencia y eficacia diferencial de los efectos de centro entre variables aptitudinales que en la literatura no se encuentran con frecuencia, porque se suele estudiar la influencia de tales efectos a través de la medición de áreas curriculares. Por otra, se aporta información sobre la casi inexistente investigación sobre dichos efectos entre variables afectivas o no cognitivas, variables que también son fundamentales ya que aquellos que colaboren en el desarrollo integral de su alumnado

serán centros eficaces. Además se elabora un modelo multinivel y adaptado a un esquema CIPP de indicadores de contexto, entrada, proceso y producto que predicen un éxito en la competencia científica.

La estructura del trabajo se compone de ocho capítulos. Los cinco primeros son teóricos. El *primer* capítulo describe el movimiento de la eficacia escolar de forma cronológica, señalando distintas etapas y estudios más relevantes en cada una de ellas. El *segundo* capítulo muestra diferentes modelos de eficacia escolar elaborados en la última etapa. El *tercer* capítulo se centra en describir las distintas propiedades de los efectos escolares y los análisis aplicados que subyacen a la estimación de las mismas. Además este capítulo termina conceptualizando y explicando el modelado multinivel. El *cuarto* capítulo se centra en la descripción del movimiento en España, se describen estudios académicos y se presentan trabajos surgidos al amparo de políticas educativas. El capítulo *quinto* cierra la parte teórica haciendo una exposición de los contenidos evaluados en PISA y la metodología que subyace a dicho programa. Toda esta parte teórica cierra con el planteamiento de cinco objetivos que se van a desarrollar en la parte empírica. Dicha parte se inicia con el *sexto* capítulo donde se describe el método del trabajo, seguido del *séptimo* donde se describen los resultados encontrados en cada uno de los objetivos propuestos y por último, el *octavo* capítulo presenta la discusión y las conclusiones obtenidas.

Además se aportan cuatro anexos que complementan contenidos tratados como ejemplos de ítems, las preguntas que componen cada uno de los índices incluidos en la elaboración del modelo multinivel para la competencia científica, la distribución de los estudiantes de los países de la OCDE en las distintas escalas evaluadas y los descriptivos de los índices.

CAPÍTULO 1

MOVIMIENTO DE LA EFICACIA ESCOLAR

Capítulo 1: Movimiento de la Eficacia Escolar

1.	Movimiento de la Eficacia Escolar	19
1.1	Definición de Escuela Eficaz	19
2.	Perspectiva histórica de la Eficacia Escolar	20
2.1	Antecedentes	21
2.1.1	Informe Coleman	22
2.1.2	Reanálisis de Informe Coleman.....	23
2.1.3	Informe Plowden	24
2.2	Primeros estudios de Eficacia Escolar.....	25
2.3	Consolidación de la Investigación en Eficacia Escolar	26
2.3.1	School Social Systems and Student Achievement.....	26
2.3.2	Fifteen Thousand Hours	28
2.3.3	Centros públicos versus privados en Norteamérica	30
2.3.4	The Louisiana School Effectiveness Studies (LSES)	31
2.3.5	Junior School Project (JSP).....	33
2.4	Estudios multinivel de Eficacia Escolar	36
2.4.1	Eficacia diferencial: un reanálisis de los datos del School Matter.....	37
2.4.2	Eficacia Docente.....	38
2.4.3	Características de centros de Primaria y la calidad de la educación	39
2.4.4	The Victorian Quality of Schools Project	40
2.4.5	Un estudio de centros franceses.....	41
2.4.6	Study of departmental differences in academic effectiveness	43
3.	Otras líneas de actuación de la Eficacia Escolar	45
4.	Conclusiones	46
5.	Índice de tablas	49
6.	Índice de figuras	49

1. Movimiento de la Eficacia Escolar

El Movimiento teórico-práctico de Eficacia Escolar es la línea de investigación educativa que más ha incidido en la mejora de la educación en los últimos años (Townsend, 2007). Dentro de este movimiento destacan sus propios medios de difusión del conocimiento como la creación en 1990 de la revista internacional “*School Effectiveness and School Improvement*”, con un factor de impacto de 0.659 según datos de *Journal Citation Report* de 2009 (<http://www.accesowok.fecyt.es>), celebraciones cada cuatro años del “*International Congress for School Effectiveness and Improvement*” (ICSEI). Además de dos ediciones del manual “*The International Handbook of School Effectiveness Research*” (Teddlie y Reynolds, 2000; Townsend, 2007) que recogen las publicaciones más relevantes del ámbito. Contenidos en español se puede encontrar en la “*Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*” (REICE) editada por la “*Red Iberoamericana de Investigación sobre Cambio y Eficacia Escolar*” (RINACE).

La finalidad de la Investigación sobre Eficacia Escolar (IEE) es elaborar una teoría comprensiva capaz de integrar los elementos que ayudan a que el centro escolar sea eficaz (Reynolds et al., 2005; Townsend, 2007). El Movimiento está compuesto por estudios empíricos que tienen por objeto la estimación de la magnitud de los efectos escolares, el análisis de sus propiedades científicas y el estudio de las características escolares, de aula y de contexto que caracterizan a una escuela eficaz (Murillo, 2005). Bajo esta línea se entiende por efecto escolar la capacidad de los centros educativos para influir en los resultados de su alumnado. Técnicamente se define efecto escolar como el porcentaje de varianza del rendimiento del alumno debido a variaciones entre centros (Murillo, 2007).

1.1 Definición de Escuela Eficaz

La definición de escuela eficaz de referencia es la que propone Murillo (2005, pp. 29-30) entendiendo que una escuela eficaz es “*aquella que consigue un desarrollo integral de todos y de cada uno de sus alumnos mayor de lo que sería esperable teniendo en cuenta el rendimiento previo y la situación social, económica y cultural de las familias*”, de tal forma una escuela eficaz presenta estas tres características definitorias:

1. *Mejora el desarrollo integral del alumno.* Se refiere no sólo a que su alumnado domine destrezas básicas de leer, escribir y contar, sino también, debe ayudar a la adquisición de habilidades creativas y críticas,. Esto es, debe enseñar contenidos de un “*currículo global*” o “*currículo humano*” compuesto por habilidades y actitudes deseables en las comunidades del futuro (Townsend y Otero, 1999).
2. *Equidad.* Una escuela será eficaz si promueve el desarrollo de cada uno de sus alumnos. Por tanto, el concepto de eficacia está íntimamente relacionado con el término de equidad (Murphy, 1992; Stoll y Fink, 1996), según el cual no pueden considerarse “*buenos resultados*” aquellos que logra una parte del alumnado con exclusión de otra.
3. *Valor añadido.* Para conocer la eficacia de una escuela, no es suficiente datos directos como resultados medios en una prueba. Es imprescindible ajustar el rendimiento de los alumnos en función de su rendimiento previo y de la situación social, económica y cultural de su familia, es lo que se conoce valor añadido educativo o valor agregado (Webster, Mendro y Almaguer, 1994; Saunders, 1999). En una escuela eficaz los alumnos progresan más allá de lo que sería esperable teniendo en cuenta sus condiciones de entrada (Mortimore, 1991). Técnicamente, una escuela es eficaz cuando sus residuales escolares – la diferencia entre el resultado del centro y el valor previsible teniendo en cuenta el nivel social, económico y cultural de los alumnos y su rendimiento previo- sean positivos y lo más elevados posibles (Peña-Suárez, Fernández-Alonso y Muñiz, 2009).

2. Perspectiva histórica de la Eficacia Escolar

Existen distintas propuestas de organización de la historia de la IEE representadas en la figura 1 (Creemers, 1996; Jansen, 1995; Murillo, 2005; Murphy, Hallinger y Mesa, 1985; y Reynolds y Teddlie, 2000) Aunque todas ellas comparten dos analogías: el inicio del movimiento en torno a 1966 con la publicación del Informe Coleman y el inicio de la última etapa en 1986 con la publicación del artículo de Aitkin y Longford (1986), a partir del cual se sigue una metodología multinivel en la estimación de los efectos escolares.

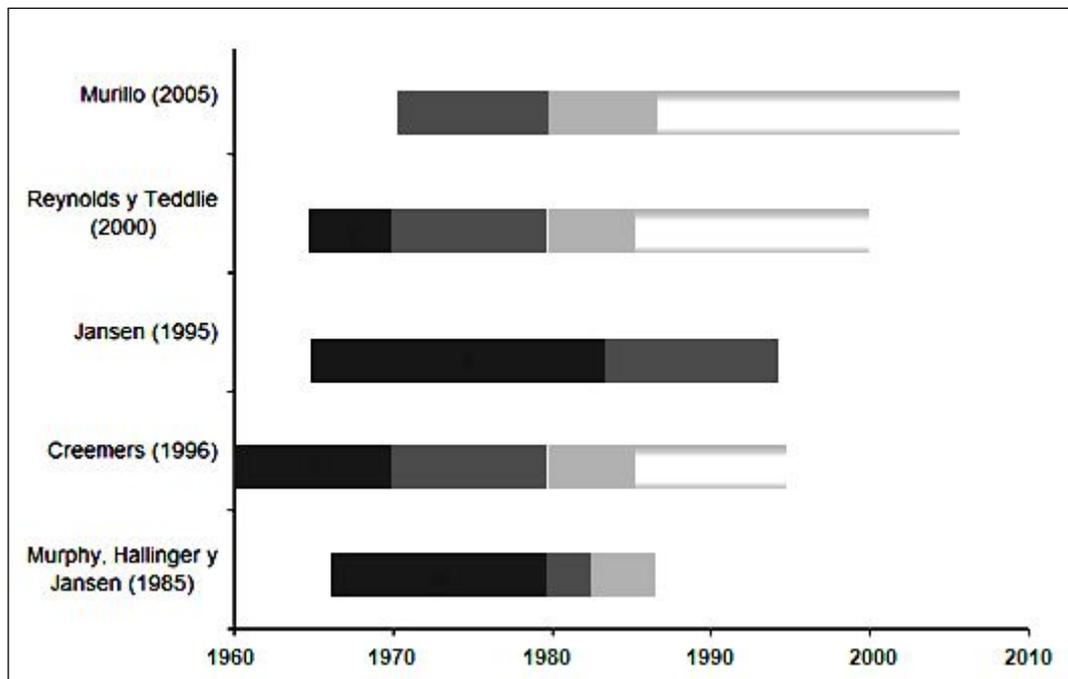


Figura 1: Diferentes propuestas de organización de la eficacia escolar. Fuente: elaboración propia

Para el desarrollo de este apartado se toma como referencia la clasificación de Murillo (2005) dado que es la más reciente. Se dividen en tres períodos: antecedentes, primeros estudios (1971-1979), consolidación de la eficacia escolar (1979-1986) y la última generación de trabajos (1987-actualidad).

2.1 Antecedentes

En la literatura se encuentran diferentes indicios de un interés creciente por conocer los efectos de los centros educativos años anteriores al Informe Coleman. Por ejemplo Kreft (1987) en su tesis doctoral recoge una serie de comunicaciones presentadas en el 27º Anuario del *National Society for the Study of Education* celebrado en Estados Unidos en 1927, como: “*rendimiento y asistencia escolar*” y “*rendimiento y métodos escolares*” que inciden en la importancia de los efectos escolares. Incluso este interés por los efectos escolares convivía con otras líneas de investigación educativa como la Eficacia Docente con la publicación de Gage (1963) “*Handbook of Research on Teaching*”; un auge de trabajos empíricos centrados en conocer los efectos de la educación en el desarrollo psicosocial (Creemers, 1996) y una corriente sociológica dedicada al estudio de la autonomía escolar y su papel activo a la hora de generar resultados en los alumnos. Por tanto, el Informe Coleman (Coleman et al., 1966)

supone frenar más que potenciar el interés de los investigadores por temas de Eficacia Escolar (Murillo, 2005).

2.1.1 Informe Coleman

En 1964 Estados Unidos aprobó la Ley de Derechos Civiles, entre sus artículos aparecía el mandato de realizar una encuesta nacional con la finalidad de estudiar la naturaleza y extensión de las desigualdades en la escuela pública estadounidense. Este estudio recibió el nombre de “*Equality of Educational Opportunity Study*” liderado por James S. Coleman.

El resultado más sobresaliente tras eliminar la influencia del origen étnico y la situación socioeconómica del alumno, es que las características escolares medidas explicaban apenas un 10% de la varianza del rendimiento de los alumnos. Uno de los resultados más famosos fue que las diferencias entre centros escolares estadounidenses de secundaria explicaban el 4,95% de la varianza del rendimiento en matemáticas de los alumnos blancos de 14 años, y el 8,73% de los alumnos afroamericanos de la misma edad –ambas puntuaciones corregidas por la situación socio-económica del alumnado–. Llegando a la conclusión de que las “*escuelas ejercen escasa influencia sobre los alumnos, por lo que no se puede sostener que su rendimiento sea independiente de su estatus y contexto social*” (Coleman, 1966 p. 325, tomado de Murillo, 2005).

Las reacciones al Informe Coleman por parte de la comunidad científica fueron muy variadas, algunos asumieron los resultados como válidos y, simplemente, obviaron el nivel escuela en sus estudios. De esta manera se potenció una perspectiva individualista o psicológica, que buscaba conocer los factores exclusivamente personales y contextuales asociados al rendimiento.

Otros trabajos, dado que no se encontró eficacia en factores de centro, se centraron en la eficiencia o cómo optimizar los recursos para conseguir objetivos de rendimiento, omitiendo lo que acontece dentro del centro docente. A esa línea se la llamó estudios de productividad escolar (e.g. Hanushek, 1986; Fraser et al., 1987; Gómez Dacal, 1989; Hanushek, 1997) Los investigadores asumen que los centros docentes son sistemas en los cuales los *inputs* se transforman en *outputs* influidos por determinados marcos tecnológicos, estructuras organizativas y escenarios institucionales.

Finalmente otros investigadores no llegaron a creer las conclusiones de Coleman, así que realizaron diversos reanálisis de los datos, intentando descubrir posibles incorrecciones. De entre los reanálisis destacan tres trabajos por su importancia y posterior influencia: el reanálisis de Smith (1972), el estudio de Jenck y colaboradores (1972) y el coordinado por Mayeske (1972) que se describen brevemente.

2.1.2 *Reanálisis de Informe Coleman*

Los estudios de reanálisis, a pesar de sus objetivos iniciales, supusieron una confirmación de los resultados de Coleman et al. (1966). La causa se atribuye a que compartían el mismo modelo de estudio: *modelo input-output* o *modelo de caja negra* a través del cual se estudiaba la relación entre variables de entrada de tipo humano, económico y social y variables resultado, sin tener en cuenta lo que acontecía dentro del centro.

2.1.2.1 *Smith (1972)*

El Informe de 1966 asumía que los cinco grupos de variables seleccionados para explicar el rendimiento poseían una naturaleza aditiva y que las variables contextuales debían estar incluidas antes en el modelo que los factores escolares. Smith (1972), por el contrario, defendía que los factores de background compartían el mismo poder explicativo con los recursos escolares, dado que ambos elementos estaban relacionados.

2.1.2.2 *Jenck et al. (1972)*

Jencks et al. (1972) ampliaron la muestra del Informe Coleman llevando a cabo un estudio de carácter longitudinal midiendo a más de un centenar centros de Secundaria. Los resultados de este trabajo indicaron que los centros docentes contribuyen escasamente a atenuar las diferencias dadas en rendimiento por el estatus económico. Afirmaron que el rendimiento del centro estaba básicamente determinado por circunstancias familiares del alumnado, por lo que cualquier reforma educativa, como programas compensatorios, resultaban inútiles.

2.1.2.3 *Mayeske et al. (1972)*

Mayeske y cols. (1972) intentaron corregir dos principales problemas que presentaba el Informe Coleman: por una parte, la variable de rendimiento utilizada y, por otra, la metodología de análisis. De tal forma que elaboraron un índice compuesto

por cinco variables a partir de los resultados de un análisis factorial: comprensión lectora, competencia verbal y no verbal, competencia matemática e información de carácter general; y aplicaron el método de partición de la varianza. Este método permite separar la contribución única y específica que hace cada una de las variables independientes y la compartida por dos o más variables independiente en la explicación de la varianza de la variable dependiente. Los resultados dieron cuenta que un buen porcentaje de la varianza se explicaba por la contribución común de variables de contexto y de centro.

2.1.3 Informe Plowden

Al mismo tiempo se publicó el estudio “*Children and their Primary School*” o también conocido como el “*Informe Plowden*” en Reino Unido (Plowden Committee, 1967). Esta investigación intentó indagar en la relación entre el bajo rendimiento y características del contexto (e.g. el tamaño de la unidad familiar) y a su vez planteó medidas políticas que debía adoptar el gobierno. Las variables analizadas se representan en la figura 2.

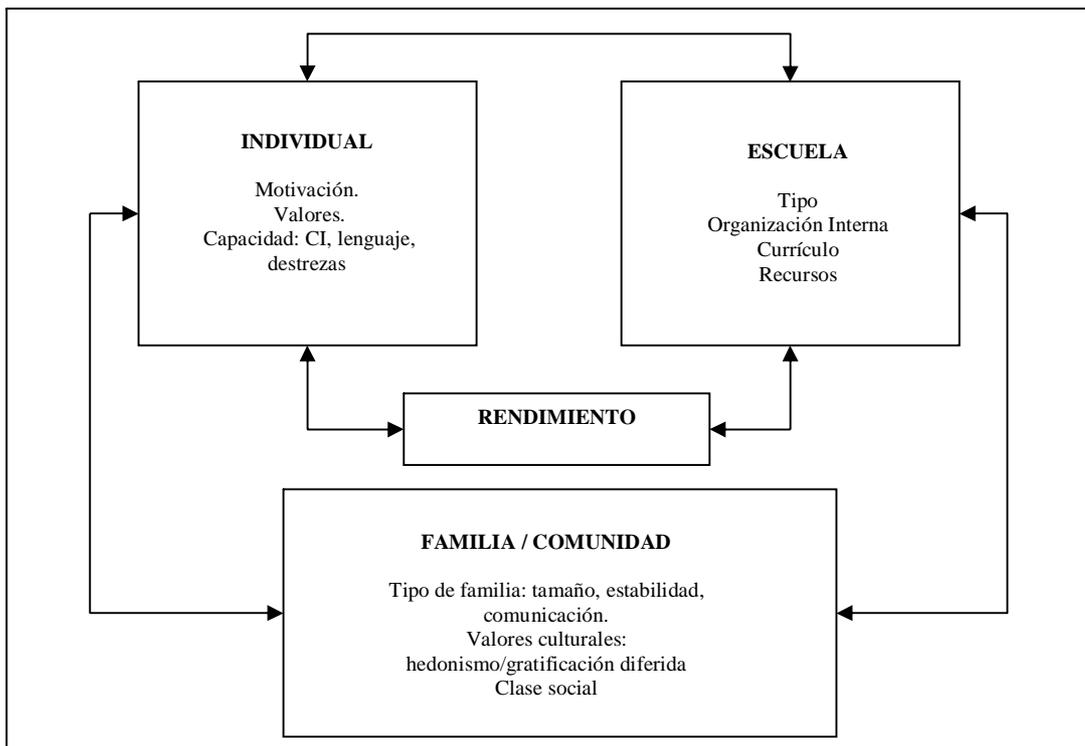


Figura 2: Modelo propuesto en el Informe Plowden (1967). Fuente: Murillo (2005 p.48)

Entre sus resultados destaca la síntesis que hace de los indicadores, agrupándolos en tres grupos independientes entre si: actitudes de los padres,

circunstancias familiares y situación de la escuela. El hallazgo más destacable fue sus conclusiones similares a los del Informe Coleman (1966): *“las diferencias entre familias explican más de la varianza de los niños que las diferencias entre escuelas”*. (Plowden Committee, 1967, p. 35).

2.2 Primeros estudios de Eficacia Escolar

Este período se caracteriza por considerar como objeto de estudio a centros con resultados peculiares o inusuales. El término de *escuelas ejemplares* se refiere a *“centros cuyos resultados son claramente superiores a los previstos en función de sus resultados previos y de la población a la que acogen”* (Murillo, 2005, p.50). Realmente el término de escuelas ejemplares o inusualmente eficaces está dentro de lo que se podría considerar *escuelas prototípicas* u *outliers*. Se entienden centros prototípicos *“aquellos centros que no se ajustan a los modelos prefijados; aquellos casos que, por su rareza, son habitualmente eliminados de los análisis que pretenden conocer los patrones habituales, pero que resultan extremadamente útiles para conocer comportamientos o resultados extraordinarios”* (Murillo, 2005, pp. 51-52).

Todos los estudios que se van hacer referencia en este apartado se encuentran dentro de una tendencia llamada *“Movimiento de Escuelas Ejemplares”* (Miller, 1985) desarrollada fundamentalmente en Estados Unidos.

Weber (1971) analizó cuatro centros que cumplían dos características fundamentales: alto rendimiento lector y población económicamente en desventaja. Encontró una serie de factores que parecían determinar sus altos rendimientos, tabla 1:

Tabla 1: Factores de éxito académico (Weber, 1971). Fuente: elaboración propia.

Factores que determinan altos rendimientos	
1.	Fuerte liderazgo instructivo
2.	Buen clima escolar (centrado en la tarea, tranquilo y ordenado)
3.	Tareas centradas en la enseñanza de la lectura
4.	Altas expectativas sobre los alumnos
5.	Enseñanza individualizada
6.	Utilización del método fónico en la enseñanza de la lectura
7.	Evaluación constante del progreso de los alumnos

Posteriormente Weber (1974) comenzó un segundo estudio encargado por la Oficina de Mejora de la Educación en Nueva York identificó dos centros urbanos con población socioeconómica baja: el primero con alto y el segundo con bajo rendimiento,

encontrando que los factores que los diferenciaba eran los del trabajo de 1971 y dos factores nuevos como:

1. El trabajo administrativo de los directivos
2. Utilización de una misma estrategia de enseñanza de la lectura por parte de todos los docentes del centro

Ambos estudios tienen una gran relevancia por considerarse una primera relación de factores de eficacia escolar (Fuentes, 1986) y, también, abren una nueva aproximación metodológica dedicada al estudio de centros docentes outliers (e.g. Austin, 1978; Brimer, 1978; Brookover y Scheneider, 1975; Ellis, 1975; Lezotte, Edmonds y Ratner, 1974; New York State Department of Education, 1976; Madden, Lawson y Sweet, 1976; Kappa, 1980; Reynolds, 1975, 1976; Reynolds y Sullivan, 1979; Spartz et al., 1977; Venezky y Winfield, 1979 y Wilder, 1977). Aunque el trabajo de esta etapa con más repercusión es el realizado por Edmonds (1979) donde elabora un “*modelo de cinco factores de eficacia*”, tabla 2.

Tabla 2: Modelo de cinco factores de Edmonds (1979). Fuente: elaboración propia.

Modelo de cinco factores	
1.	Poseen un liderazgo fuerte
2.	Hay un clima de altas expectativas en el rendimiento de los alumnos
3.	Atmósfera es ordenada sin ser rígida
4.	Objetivo prioritario del centro es la adquisición de destrezas y habilidades básicas
5.	Evaluación constante y regular del progreso de los alumnos

2.3 Consolidación de la Investigación en Eficacia Escolar

Esta nueva etapa muestra tres características distintivas: la vuelta de los estudios de muestras representativas; una mayor preocupación por recoger adecuadamente los datos sobre procesos escolares, lo que supone el paso de un *modelo input-output* a un *modelo input-process-output*; y los resultados obtenidos de esta línea comienzan a influir en el *Movimiento de la Mejora Escolar*. Los siguientes epígrafes presentan los estudios más sobresalientes de esta etapa.

2.3.1 School Social Systems and Student Achievement

Brookover et al. (1979) rompe con el modelo input-output para introducir dos constructos intermedios: la estructura social y el clima social del centro. Además presenta como novedad la ampliación de la visión escolar incluyendo como *outputs* el

rendimiento académico, el autoconcepto de la capacidad académica y la autoconfianza del alumnado. La figura 3 representa el marco conceptual de Brookover et al. (1979).

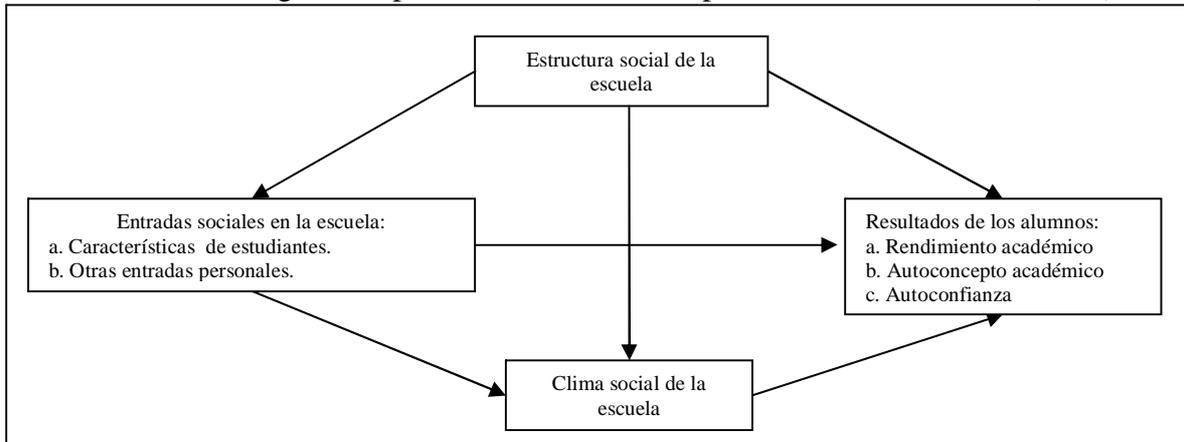


Figura 3: Modelo conceptual de Brookover et al. (1979). Fuente: Murillo (2005 p.59)

Se analizaron 68 centros de Primaria seleccionados aleatoriamente de entre la población de centros del Estado de Michigan, aplicando tres técnicas de análisis: correlación producto-momento, análisis de regresión por el método de introducción de variables independientes *paso a paso*, y la partición de la varianza explicada en función de los resultados del análisis de regresión.

Fuente de datos	Factor	Correlación con el rendimiento
Alumno	Sensación por parte del alumno de inutilidad de lo académico	0,769
	Apreciaciones y expectativas futuras	0,218
	Apreciaciones y expectativas presentes percibidas	-0,568
Profesor	Apreciaciones y expectativas para seguir estudios universitarios	0,228
	Apreciaciones y expectativas para finalizar la secundaria superior	0,664
	Implicación hacia la mejora del rendimiento de los alumnos	-0,105
	Percepción de las expectativas del directivo	0,198
	Inutilidad académica de los docentes	-0,129
Director	Interés de los padres por los estudios de los hijos	0,320
	Inclinación por mejorar el rendimiento educativo	-0,237
	Impresión de la reputación de la escuela	0,365
	Evaluaciones y expectativas de los alumnos	0,377

Tabla 3: Correlaciones significativas entre rendimiento y clima escolar (Brookover et al., 1979). Fuente: Murillo (2005 p.60)

Los resultados obtenidos mostraron varias correlaciones significativas y positivas. Destacan las relaciones entre las variables que componen el clima escolar y el rendimiento, tabla 3. Sin embargo, las correlaciones entre variables cognitivas y las no cognitivas no fueron significativas, demostrando la existencia de dos criterios diferentes de eficacia.

Del análisis de regresión *paso a paso* se encontró: (tomado de Scheerens, 1992. p. 122):

1. Variables de entrada explicaban el 75% de la varianza de las puntuaciones medias de las escuelas;
2. Variables de estructura explicaban el 41%;
3. Variables de clima explicaban el 72%;
4. Analizando las variables de clima después de haber considerado las variables de entrada, apenas explicaba el 6%; las variables de estructura en la misma situación explicaban el 4%; y
5. cuando las variables de entrada eran introducidas después de las variables de clima, aquéllas explicaban el 9% de la varianza.

Todo ello indica que la varianza entre las escuelas puede ser explicada no sólo por las variables de entrada sino también por las de clima, y, en menor medida, por indicadores de estructura. A su vez enfatiza que los resultados se distorsionan si no se tiene un control previo de las variables de entrada.

Los análisis de regresión con las variables autoconcepto y autoconfianza indicaron que los porcentajes varianza explicada por las variables de clima y de estructura eran superiores (21 y 14% respectivamente).

2.3.2 *Fifteen Thousand Hours*

A lo largo de los doce primeros años de vida escolar, los alumnos pasan quince mil horas en el centro docente por lo que el informe trata de averiguar ¿qué impacto tiene ese tiempo en su formación? y ¿qué es lo que genera ese impacto?

Se llevó a cabo un estudio longitudinal del progreso de unos dos mil alumnos, pertenecientes a 12 centros de Secundaria con diferentes características situados en

Londres. La metodología del estudio posee tres notas definitorias: distintas medidas output, un diseño de investigación longitudinal y variables de características escolares. La tabla 4 recoge el mapa de variables analizadas por Rutter et al. (1979).

Tabla 4: Mapa de variables de Fifteen Thousand Hours (Rutter, 1979). Fuente: Murillo (2005 p.63).

Tipo de variables	Especificación de variables analizadas
Variables de entrada	<ol style="list-style-type: none"> 1. Inteligencia de los alumnos 2. Rendimiento lector previo 3. Problemas de comportamiento 5. Delincuencia 6. País de origen 7. Situación familiar
Características del proceso del funcionamiento del centro	<ol style="list-style-type: none"> 1. Énfasis en lo académico 2. Comportamiento docente 3. Uso de recompensas y castigos durante la enseñanza 4. "Amigabilidad" de la escuela para los alumnos 5. Responsabilidades específicas de los alumnos 6. Estabilidad del profesorado 7. Estabilidad del círculo de amigos de los alumnos 8. Organización del personal
Factores ambientales	<ol style="list-style-type: none"> 1. Situación socioeconómica media de los alumnos del centro 2. Proporción de alumnos respecto a su situación en CI, socioeconómico o etnia
Características físicas y administrativas del centro	<ol style="list-style-type: none"> 1. Situación del edificio escolar 2. Criterios de distribución de alumnos en aulas
Variables de resultado	<ol style="list-style-type: none"> 1. Asistencia escolar 2. Buen comportamiento de los alumnos 3. Delincuencia 4. Resultados de exámenes

Del estudio se extrajeron algunas conclusiones (tomado de Murillo, 2005, p. 64):

1. Los centros de secundaria estudiados difieren considerablemente tanto en el comportamiento de los alumnos como en su rendimiento académico;
2. Las diferencias permanecen estables durante un período de cuatro a cinco años;
3. Los centros puntúan de forma parecida en todos los criterios de eficacia utilizados;
4. Las características de proceso pueden ser manipuladas por el profesorado;
5. Los resultados de los exámenes fueron mejores en centros con una importante proporción de alumnos con cocientes de inteligencia cercanos al promedio; y

6. La relación entre las puntuaciones generales de proceso y criterio fue mucho mayor que la relación con los indicadores de procesos individuales

2.3.3 Centros públicos versus privados en Norteamérica

Coleman, Hoffer y Kilgore, (1981 y 1982) realizaron a principios de los años 80 un estudio con el objetivo de comparar resultados de centros públicos y privados para conocer qué tipo de enseñanza era mejor y cuáles eran los factores generaban tales diferencias.

El estudio toma los datos del trabajo longitudinal *High School and Beyond* (HSB) formado por una muestra representativa del total de alumnos estadounidenses de Secundaria. Los directivos rellenaron un cuestionario sobre el funcionamiento del centro y los docentes aportaron información sobre sus alumnos y sobre ellos mismos. Como variables de rendimiento se midieron sus puntuaciones en lectura, vocabulario y matemáticas. Los factores de proceso medidos se agruparon en cinco grandes bloques: deberes para casa, asistencia escolar, clima disciplinado percibido por los alumnos y comportamiento de los alumnos en clase. Además de diversos estadísticos descriptivos, los resultados proceden de análisis de regresión múltiple. La tabla 5 muestra los resultados en las pruebas evaluadas tras ajustar las puntuaciones a variables de *background* familiar.

Tabla 5: Estimación de la puntuación para centros públicos y pendientes estimadas para centros privados. Fuente: Murillo (2005)

	Lectura	Vocabulario	Matemáticas
Alumnado de escuelas públicas	3,60	3,69	9,40
Incrementos para escuelas católicas	0,32	0,36	0,58
Incrementos para otros centros privados	0,14	0,33	0,56

La comparación entre las puntuaciones brutas y las ajustadas indican que la mitad de las diferencias entre escuelas públicas y privadas se debían a los factores del contexto sociocultural. Posteriormente ajustaron las puntuaciones con el rendimiento del año anterior, y corrigieron esos datos con la tasa de abandono. Esta tasa de abandono suponía el 24% para los centros públicos, el 12% para los centros católicos y el 13% para otros centros privados, por tanto, también se encontraron diferencias entre los centros públicos y los privados. Tras el ajuste los centros privados mostraban mejores resultados con la excepción en lectura.

Por lo que la política escolar afecta de manera importante en el rendimiento de los alumnos. Los factores en los que difieren significativamente los centros privados son la disciplina y el comportamiento del alumno y mayores exigencias académicas.

2.3.4 *The Louisiana School Effectiveness Studies (LSES)*

Teddlie y Stringfield comenzaron con el LSES en 1980 como respuesta a un mandato legislativo con el objetivo de llevar a cabo un estudio a largo del rendimiento en centros de Primaria.

La primera fase (LSES-I), desarrollada a lo largo del curso escolar 1981/82, fue apenas un estudio piloto para validar los instrumentos y métodos a utilizar. La segunda fase (LSES-II) se seleccionaron a 76 centros de Primaria del estado a través de un muestreo estratificado simple. La muestra estaba compuesta por 5.400 alumnos y 250 docentes y 76 directivos. Al alumnado se les aplicaron varias escalas como una prueba estandarizada de rendimiento y un cuestionario de clima escolar, de autoconcepto y de locus de control. A los docentes se les pasaron encuestas relativas al clima escolar, al autoconcepto y al locus de control y a las direcciones un cuestionario general. Los datos permitieron efectuar comparaciones entre escuelas eficaces, ineficaces y típicas.

La tercera fase (LSES- III) se diseñó un análisis más detallado de un pequeño número de centros, introduciendo en el diseño informaciones sobre eficacia docente. Concretamente se analizaron 16 centros, ocho pares de centros completamente opuestos en sus resultados pero equiparados en sus características sociales. Como instrumentos, se utilizaron cuestionarios para directivos, docentes y alumnos, sistemas de observación del aula de alta y baja inferencia etc. (Teddlie, 1984 y Teddlie, Kirby y Stringfield, 1989). Por último en la cuarta fase (LSES-IV) se desarrolló un estudio longitudinal cuya muestra la componían 16 centros de la anterior fase, convirtiéndose en un análisis de los procesos de mejora de los centros.

En general los resultados de las distintas fases se pueden agrupar en siete conclusiones (Stringfield y Teddlie, 1989):

1. La mayor parte de la varianza del rendimiento se observa entre los alumnos de la misma clase, mientras que el 12% se produce entre docentes de la misma escuela y el 13% entre escuelas;

2. Los resultados encontrados son similares a los de otros estudios norteamericanos, como por ejemplo el de Brookover et al. (1979);
3. Las medidas de producto de proceso y las relaciones entre ellas son estables en el tiempo;
4. El clima es el mejor predictor del rendimiento de una escuela;
5. Las variables de contexto también se han mostrado relevantes como el nivel socioeconómico, el hábitat donde se encuentra la escuela, su titularidad, las características del profesorado y los directivos, las instalaciones, etc.;
6. Existe una fuerte relación entre las variables de eficacia docente y eficacia escolar. Por ejemplo los centros más eficaces son aquellos que más tiempo dedican a la enseñanza directa; y
7. Todas las escuelas eficaces tienen un buen líder instructivo

La tabla 6 muestra el comportamiento de factores de eficacia en dos centros prototípicos equiparados en variables de contexto pero con resultados completamente opuestos.

Tabla 6: *Contraste entre características de eficacia del LSES*. Fuente: Teddlie y Stringfield (1993 p. 132)

Centro eficaz	Centro menos eficaz
<i>Directivo</i>	
1. Liderazgo estable y adecuado	1. Liderazgo inestable generalmente inadecuado
2. Estructura organizativa adecuada e informal	2. Estructura organizativa inadecuada e informal
3. Liderazgo académico compartido con el profesorado	3. Liderazgo académico no compartido
4. Resistente al cambio externo	4. Acepta el cambio externo
5. Relación cercana entre los administradores	5. Relaciones tensas entre los administradores
6. Buen uso del personal docente de apoyo	6. Utilización no imaginativa del personal docente de apoyo
<i>Profesorado</i>	
7. El profesorado es cálido y amistoso	7. El profesorado es frío y reservado
8. Fuerte cohesión entre el profesorado	8. Falta de cohesión entre el profesorado
9. Inexistencia de conflictos personales claros entre docentes	9. Enfrentamientos abiertos entre los docentes
10. Integración del personal de apoyo entre el profesorado	10. Integración inadecuada del personal docente de apoyo
11. Esfuerzos de cooperación para mejorar la enseñanza	11. Efecto de arriba abajo para mejorar la enseñanza
12. Alta estabilidad entre el profesorado	12. Estabilidad del profesorado moderadamente baja
13. Alto tiempo dedicado a la tarea/clima de aula positivo	13. Escaso tiempo dedicado a la tarea /mal clima evidente
14. Completa uniformidad en la enseñanza entre las diferentes aulas	14. Importantes diferencia en la enseñanza entre aulas
15. Apoyo dado a los nuevos docentes	15. Poco apoyo dado a los nuevos docentes
<i>Alumnado</i>	
16. Excelente disciplina y comprensión de las normas	16. Pobre disciplina y comprensión de las normas
17. Los alumnos están implicados en el funcionamiento del centro	17. Poca o ninguna implicación de los alumnos en el funcionamiento del centro
18. Escasa utilización del castigo físico	18. Uso excesivo del castigo físico
19. Clima orientado al alumno	19. Clima orientado al adulto
20. Consistencia en el alto rendimiento de los alumnos	20. Consistencia en el bajo rendimiento de los alumnos

2.3.5 *Junior School Project (JSP)*

Mortimore et al. (1986 y 1988) lideró una investigación con el título oficial de *Inner London Educational Authority's Junior School Project (JSP)*, más conocido con el nombre del informe final.- *School Matters*-. Su trascendencia es tal que Goldstein (1997) afirmó que es la primera investigación en la historia de la Eficacia Escolar que cumple todos los requisitos mínimos y necesarios para realizar cualquier tipo de inferencia válida. El trabajo analizó cerca de 2.000 alumnos durante cuatro años de escolarización incluyendo un cambio de etapa: de Primaria a Secundaria. Se plantearon

tres objetivos de eficacia: en primer lugar, analizar si los centros educativos y las aulas varían en eficacia al controlar la variación de entrada de los alumnos; en segundo término, conocer si algunos centros o algunas aulas son más eficaces para ciertos grupos de alumnos; y, por último, determinar los factores relacionados con la eficacia de los centros y de las aulas. Los datos recogidos y los instrumentos de medida aplicados se muestran en la tabla 7:

Tabla 7: Mapa de variables de Junior School Project. Fuente: Elaboración propia.

Variables e instrumentos de medida de Junior School Project	
<i>1. Características de entrada</i>	
1.1.	Características de <i>background</i> de los alumnos, características familiares, situación socioeconómica, lengua materna y país de origen
1.2.	Rendimiento previo en lectura, aritmética y destrezas visuales y espaciales (medidas al inicio de Primaria)
1.3.	Evaluación del comportamiento de los alumnos al comienzo de primaria por parte de los docentes
<hr/>	
<i>2. Características de la escuela y de la escolarización en el aula</i>	
2.1.	Entrevistas con los directivos sobre la organización escolar y la política de la escuela, incluyendo opiniones sobre liderazgo, filosofía de la educación, cualificación y experiencia, procedimientos de agrupamiento de los alumnos, criterios de asignación de los profesores a los grupos, implicación de los profesores en la toma de decisiones y división de responsabilidades entre el personal
2.2.	Estrategias didácticas evaluadas mediante observación directa y estructurada a través de listas de control y manejo de datos cualitativos, tales como descripciones de los observadores
<hr/>	
<i>3. Características del entorno y ciertas características de la escuela determinadas por el entorno</i>	
3.1.	Opiniones de los padres acerca del aprendizaje de sus hijos, y de su implicación en el mismo
3.2.	Características descriptivas de la escuela. Relaciones del centro con organizaciones de etapas escolares anteriores y titularidad del centro
<hr/>	
<i>4. Rendimiento en aprendizaje cognitivo (tomadas en varios momentos de la escolarización de los alumnos)</i>	
4.1.	Rendimiento obtenido en pruebas estandarizadas de lectura y aritmética
4.2.	Pruebas individuales en matemáticas
4.3.	Medidas anuales de escritura creativa
4.4.	Medidas de destrezas de expresión verbal (entre una muestra aleatoria de alumnos)
<hr/>	
<i>5. Productos no cognitivos del sistema educativo</i>	
5.1.	Evaluación del comportamiento del alumno por su docente
5.2.	Actitudes frente a la educación (medidas a través de informes que elaboraban periódicamente los alumnos)
5.3.	Medidas de cómo los alumnos creen que los perciben sus profesores
5.4.	Medidas de la autoimagen de los alumnos

Entre los resultados destaca que el 23% de la varianza en el progreso de los alumnos en lectura es debido a factores escolares y sólo el 6% es debido a factores de *background* del alumno. Por áreas las diferencias fueron de 25 puntos y comparada con la media general de 54 para lectura; dicha diferencia fue de 12 puntos en matemáticas y comparada con la media general de 28 puntos. Respecto a los factores escolares que ayudan a explicar esas diferencias, los autores seleccionaron una serie de escuelas eficaces tanto en el área social como en la académica. Estas escuelas presentaban doce características:

1. *Liderazgo con propósito*. La dirección estaba activamente implicada en el centro, y era capaz de compartir el poder con los docentes, especialmente en la toma de decisiones respecto al gasto y las orientaciones curriculares.
2. *Implicación de la subdirección*. La subdirección estaba implicada en las decisiones políticas hacía incrementar el progreso de los alumnos.
3. *Implicación de los docentes*. En las buenas escuelas, los docentes desarrollaban funciones de planificación del currículum y cuestiones políticas de centro.
4. *Consistencia entre docentes*. La coherencia entre los docentes tenía un efecto positivo, los alumnos obtuvieron mejores resultados cuando había un enfoque de enseñanza consistente.
5. *Sesiones estructuradas*. Los alumnos rendían más cuando el día escolar estaba estructurado. En las escuelas eficaces, el trabajo de los alumnos estaba organizado por el profesor, aunque también permitían alguna libertad en la estructura.
6. *“Enseñanza intelectualmente desafiante”*. La calidad de la enseñanza era un factor fundamental para promover el progreso y el desarrollo de los alumnos. Los resultados mostraron con toda claridad que los alumnos avanzaban más cuando los profesores les estimulaban y se mostraban entusiastas. La existencia de afirmaciones, preguntas de orden superior y docentes que insistían en que sus alumnos utilizaran estrategias de resolución de problemas parecía especialmente útil.
7. *Ambiente centrado en el trabajo*. Los alumnos disfrutaban con su trabajo y se mostraban *impacientes* por comenzar nuevas tareas. El ruido en las aulas era bajo y

el movimiento en la sala prácticamente se reducía a lo requerido en las tareas escolares.

8. *Atención reducida a pocos temas en cada sesión.* Los estudiantes progresaban más cuando los profesores dedicaban toda su energía a una sola tarea o, como mucho, a dos.
9. *Máxima comunicación entre profesores y alumnos.* Cuanta más comunicación había entre el profesor y el alumno sobre el contenido del trabajo más avanzaba el alumno. Los profesores que utilizan cualquier oportunidad para hablar a la clase en su conjunto, por ejemplo leyendo un relato o formulando una pregunta, fueron más eficaces.
10. *Escribir y utilizar evaluaciones.* Son actuaciones que servían no sólo para el alumnado sino también para el profesorado, permitiéndole una autorrevisión del trabajo como docente y una planificación actividades posteriores.
11. *Implicación de las familias.* Las escuelas más eficaces eran aquellas en la que se estimula a los padres a que colaboren con la formación de sus hijos y a que visiten la escuela.
12. *Clima positivo.* Los buenos centros poseían una atmósfera agradable de trabajo. En estos centros y en sus aulas había un menor énfasis en el castigo y el control y mayor en la recompensa; los docentes estaban preocupados por fomentar el autocontrol por parte de los alumnos, más que por enfatizar los aspectos negativos de su comportamiento. También era importante la actitud de los docentes hacia los alumnos: los profesores disfrutaban enseñando y dedicaban tiempo no lectivo a preocuparse de sus alumnos.

2.4 Estudios multinivel de Eficacia Escolar

El artículo de Aitkin y Longford (1986) supuso cambiar la aproximación metodológica de los estudios de Eficacia Escolar. A partir de ese momento, “*la utilización de los modelos multinivel se presenta como un requisito casi imprescindible en cualquier estudio de eficacia escolar*” (Murillo, 2005. p.75). A su vez se vuelve habitual en los trabajos la observación de escuela y aula, así como la utilización de cuestionarios para estudios a gran escala (Creemers, 1996).

Se abren nuevas líneas de interés para los investigadores, entre ellas destaca el estudio de las propiedades científicas de los efectos escolares tales como la eficacia diferencial, la consistencia de los efectos o la estabilidad de los mismos. Estas propiedades de los efectos escolares han suscitado gran interés dentro de la IEE (e.g. Mandeville y Anderson, 1987; Nuttall et al., 1989; Luyten, 1994; Sammons *et al.*, 1995; Thomas, Sammons y Mortimore, 1997) como se verá en el capítulo tres.

En esta etapa se produce una fusión entre la línea de investigación de Eficacia Docente o Eficacia Instructiva y la Eficacia Escolar. Este nuevo enfoque tiene como finalidad la búsqueda de factores de aula relacionados con el logro educativo, incluido el comportamiento del docente y sus características. El trabajo de Creemers (1994) es el máximo exponente en esa fusión.

Una característica definitoria es la transformación de modelos *input-process-output* a modelos CIPP: *contexto-entrada-proceso-producto* (Scheerens, 1992). Este planteamiento tiene su máxima expresión en los modelos comprensivos de eficacia escolar que empiezan a elaborarse en los comienzos de la década de los 90 que se desarrollarán en el capítulo dos.

El número de trabajos de Eficacia Escolar desarrollados en esta época es ingente, como ejemplo en 1997 se publicaron un total de 114 estudios de Eficacia Escolar sólo en el Reino Unido (Sheffield y Saunders, 2002). Este crecimiento, sin embargo, no ha sido parejo en todos los países, de hecho el centro de gravedad del movimiento se traslada desde Estados Unidos a Reino Unido y Países Bajos. A continuación se presentan una serie de estudios de esta etapa.

2.4.1 *Eficacia diferencial: un reanálisis de los datos del School Matter*

Entre las conclusiones del *School Matters* (Mortimore et al., 1988) destacaba que los centros eficaces en promover el progreso de un grupo de alumnos (bien sea procedentes de una clase social determinada, de un género o de un grupo étnico) fueron normalmente también eficaces para los niños de otros grupos. Análogamente, los centros que fueron ineficaces para un grupo tienden a ser también ineficaces para todos. Otra conclusión que destacaba es que el impacto de la desventaja social puede ser mitigado en los resultados, sin ser eliminado ni siquiera por los centros más eficaces.

Sammoms, Nuttall y Cuttance (1993) reanalizaron los datos del *Junior School Project* utilizando técnicas multinivel con el objetivo de analizar la *Eficacia Diferencial Escolar*. Esto es, se plantearon si todos los centros son igualmente eficaces para los distintos grupos de estudiantes atendiendo a su rendimiento previo y características del contexto (género, clase social y grupo étnico). Entre las conclusiones se encuentran:

1. La varianza del rendimiento de los alumnos explicada por las diferencias entre centros parece aumentar con el tiempo, especialmente en Matemáticas;
2. Los factores de *background* de los alumnos y el rendimiento previo explican mayor proporción de varianza en el rendimiento lector que en el rendimiento matemático;
3. Existen marcadas diferencias entre centros individuales en términos de sus efectos sobre el rendimiento del centro (residuales del nivel 2) después de controlar los factores de *background* y de rendimiento previo;
4. Se han encontrado algunas evidencias acerca de la Eficacia Diferencial (pendientes distintas de centros). Los centros eficaces mejoraron el rendimiento de todos los alumnos, sobre todo en Matemáticas; y
5. No se encontró eficacia diferencial para diferentes grupos de alumnos (divididos por género, grupo étnico o clase social); y
6. A pesar de que los efectos escolares tuvieron una correlación positiva (estimados a través de los residuales del nivel 2 para Matemáticas y para Lengua). Sólo cuatro centros tenían residuales significativamente positivos para ambas áreas y seis marcadamente negativos, por lo que el concepto general de eficaz vs ineficaz parece demasiado simple para describir la dimensionalidad de los efectos escolares

2.4.2 *Eficacia Docente*

Virgilio, Teddlie y Oescher (1991) buscaron replicar los resultados del LSE-III, antes descrito, ampliando muestra e incidiendo en el estudio del comportamiento del docente (Teddlie, Kirby y Stringfield, 1989), tanto en Primaria como en Secundaria. Aplicaron técnicas de análisis multivariado para analizar los datos. Las variables independientes incluidas en el modelo fueron dos: grado de eficacia del centro (centros eficaces; medios e ineficaces) y el nivel de centro (Primaria o Secundaria). Los resultados demostraron que existen rasgos diferenciales entre docentes atendiendo al

nivel de eficacia del centro. Los docentes de centros eficaces exhibían los siguientes comportamientos muy similares a los encontrados en otros trabajos como Brophy y Good (1986):

1. Utilizaron el tiempo de transición entre clases de manera eficaz;
2. Desarrollaban un enfoque positivo para gestionar el comportamiento del alumno;
3. Centraban a los alumnos en la tarea, cuando es necesario;
4. Aplicaban más estrategias docentes en la presentación de los temas que sus compañeros en centros menos eficaces;
5. Mostraban el trabajo de los alumnos en el aula; y
6. Establecían un entorno de aprendizaje positivo

2.4.3 Características de centros de Primaria y la calidad de la educación

Un trabajo representativo de la investigación sobre Eficacia Escolar desarrollada en los Países Bajos es el estudio realizado por Brandsma sobre una muestra de 252 centros de Primaria (Brandsma, 1993, citado por Scheerens y Bosker, 1997). Sus objetivos se centraron en determinar la existencia de diferencias en eficacia entre los centros analizados y detectar cuáles son las diferencias en las características escolares.

Los datos obtenidos fueron analizados mediante modelos multinivel. Concretamente se empleó un modelo de dos niveles (escuela/aula y alumno) con coeficientes aleatorios y con pendientes e interceptos aleatorios. Los resultados indicaron que en lengua, las variables del contexto y organización de centro explican más varianza que las características de aula, del profesorado y del líder escolar, mientras que, en aritmética son las variables de la organización de aula las que tenían un mayor impacto.

En cuanto al efecto de las variables sólo la evaluación frecuente y la cooperación entre el profesorado (medido como número de reuniones) parecían ser especialmente relevantes. Por el contrario, otros factores comúnmente encontrados en otros trabajos tales como liderazgo educativo, clima ordenado, énfasis sobre las destrezas básicas y orientación del centro hacia el logro, no fueron significativos.

2.4.4 *The Victorian Quality of Schools Project*

Hill, Rowe y Jones (1995) coordinaron un proyecto con un total de 90 centros y 903 profesores y 13.909 alumnos de Primaria y Secundaria. Los datos de los alumnos se recogieron en tres cursos diferentes: 1992, 1993 y 1994. Como variables de producto se utilizaron tanto medidas cognitivas como no cognitivas. Entre las variables *output* cognitivas se evaluó el rendimiento en lengua inglesa (lectura, escritura y lengua hablada) y en matemáticas (geometría y cálculo) y como variables de producto no cognitivas recogieron datos de la atención del alumno en las clases y su actitud frente al aprendizaje. Como variables de *background* de los alumnos se midió la habilidad del alumno (a través de pruebas estandarizadas de lectura, escritura y matemáticas), el nivel socioeducativo, su lengua materna, el hábitat donde residen y el género. Para medir comportamientos en las variables medidas aplicaron escalas de control recogidas por los propios docentes.

Los resultados mostraron que la varianza entre escuelas era del 18%, utilizando un modelo de tres niveles la varianza entre aulas estaba en torno al 50% (más baja para inglés y para Secundaria) y entre centros se reducía al 8%. Lo que indica, de nuevo, que la influencia del aula en el progreso del alumno es mucho mayor que el impacto del centro. Los resultados del modelo multinivel se recogen en la figura 4, ofrecen una imagen de los factores que explican el progreso en Inglés, interrelaciones y pesos.

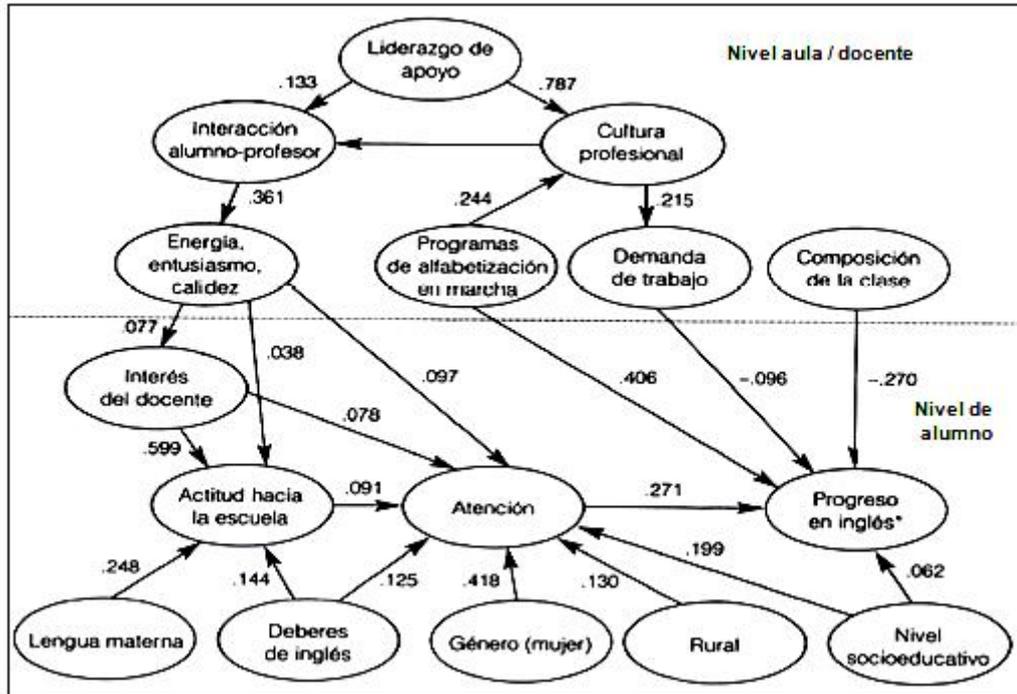


Figura 4: Resultados de Victorian Quality of Schools Project. Fuente: tomada de Murillo (2005 p.87)

2.4.5 Un estudio de centros franceses

Todas las investigaciones hasta ahora analizadas se han llevado a cabo en población escolarizada en Sistemas Educativos en los que el centro goza de una gran autonomía escolar, lo que hace plantearse si los resultados son replicables en sistemas centralizados como el sistema educativo francés.

El trabajo de Grisay (1996) pretende establecer las características del centro que inciden en el desarrollo cognitivo y socioefectivo de los alumnos, estudiando a un centenar de centros de Secundaria durante cuatro cursos escolares. Las variables analizadas se recogen en la figura 5.

Los análisis realizados fueron muy variados, pero básicamente se desarrolló un proceso de modelado multinivel de tres niveles y tres tipos de variables explicativas incorporadas independientemente: rendimiento previo, variables de *background* del alumno y características del entorno educativo.

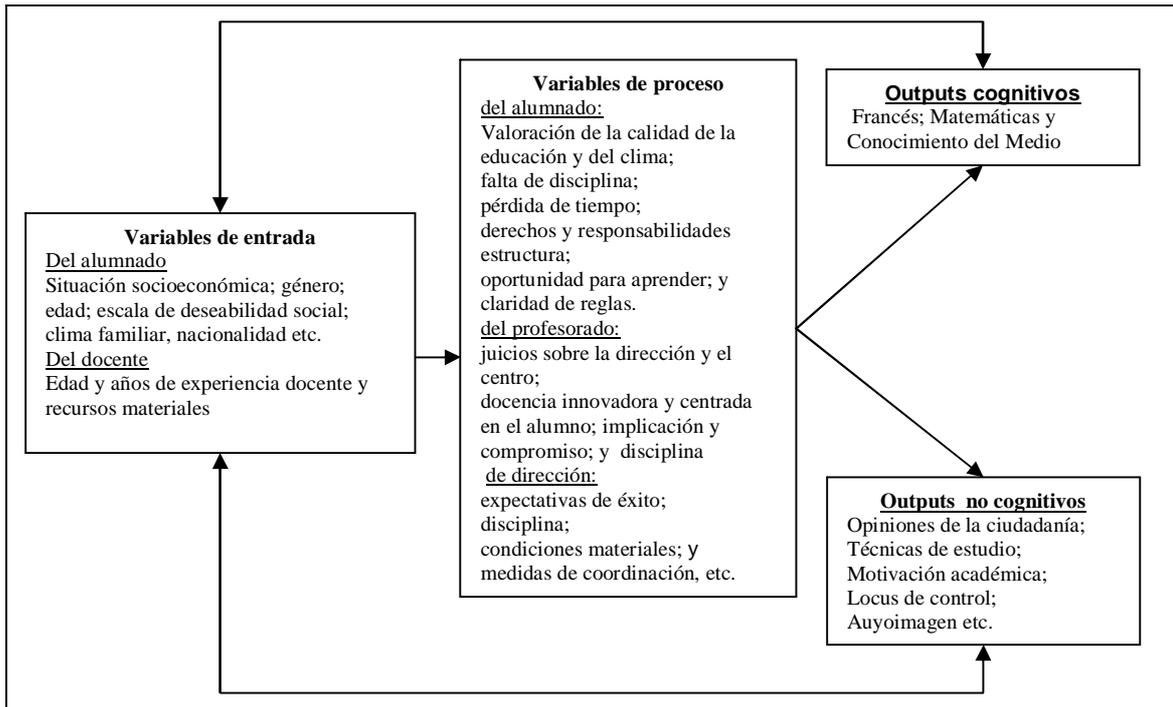


Figura 5: Mapa de variables del estudio de Grisay (1996). Fuente: Elaboración propia

Los resultados mostraron que el porcentaje de varianza explicado por el centro, después de ser ajustado por rendimiento previo y las variables de *background* fue muy bajo, de 5,4% para matemáticas y de 2,6% para francés. La varianza explicada por las aulas fue de 4,7% para matemáticas y del 1,9% para francés. Estos resultados algo bajos (10% y 4,7% sumando la varianza explicada por el centro y por el aula) pueden ser motivados por las características del sistema educativo francés, como anteriormente se señaló. No obstante encontraron una serie de variables de proceso relacionadas con el rendimiento en francés y/o matemáticas (ajustado por el rendimiento previo y las variables de *background*):

1. Juicio positivo de los alumnos sobre la calidad de la educación y el clima;
2. Disciplina y ausencia de pérdida de tiempo;
3. Derechos y responsabilidades de los alumnos;

4. Apoyo estructurado a los alumnos (tutorías, cursos de refuerzo);
5. Oportunidad para aprender;
6. Juicio positivo del centro según los profesores;
7. Implicación y compromiso de los profesores, y
8. Edad y experiencia de los docentes

Respecto a los resultados acerca de las variables de rendimiento no cognitivas se encontró que la varianza explicada de rendimiento no cognitivo por el centro y aula era mucho menor que en el rendimiento cognitivo, de hecho actitudes hacia la escuela estaban muy influidas por las características de *background* de los alumnos. No obstante las variables de proceso asociadas con el rendimiento cognitivo coincidían con las asociadas con el rendimiento no cognitivo.

La principal aportación de este estudio es el análisis del sistema educativo de carácter centralizado, por lo que en una evaluación del contexto educativo en España sería previsible encontrar más puntos en común con este trabajo que similitudes con estudios norteamericanos, británicos u holandeses, en donde los centros gozan de mayor autonomía (Murillo, 2005).

2.4.6 *Study of departmental differences in academic effectiveness*

Sammons, Thomas y Mortimore (1997) propusieron una ampliación del modelo de Eficacia Escolar de Creemers (1994) incluyendo en su estudio un nivel de departamento. Los tres objetivos de la investigación fueron:

1. Conocer el tamaño, alcance y estabilidad en el tiempo de los efectos escolares de eficacia en el rendimiento final de Secundaria (*General Certificate of Secondary Education, GCSE*).
2. Explorar el alcance de cualquier variación interna en la Eficacia Escolar en:
 - 2.1. el nivel del departamento; y
 - 2.2. los distintos grupos de alumnos
3. Investigar en detalle las razones que subyacen a cualquier diferencia en eficacia en relación con los procesos escolares y departamentales

Evaluaron a 94 centros de Secundaria y 7.000 alumnos de tres cohortes diferentes (1990, 1991 y 1992). Recogieron además de los resultados de las pruebas en seis materias (inglés, literatura, matemáticas, francés, historia y ciencias), datos de variables antecedentes de los alumnos tales como edad, género, grupo étnico, y grupo socioeconómico y rendimiento previo.

El estudio se dividió en tres partes, la primera fase se centró en estimar Eficacia Escolar teniendo en cuenta diferencias entre centros y departamentos, mediante modelos multinivel. En la segunda etapa se realizó un estudio de 6 centros y 30 departamentos, la selección de los casos se basó en la consistencia y significación estadística de los Valores Añadidos estimados a partir de los resultados obtenidos en la primera parte. La tercera fase se centró en analizar las relaciones entre las variaciones en los factores de proceso del departamento y el centro y sus relaciones con el rendimiento del alumno, mediante una metodología de encuesta aplicada a directivos y jefes de departamento.

Los resultados indicaron un efecto claro de los equipos docentes, sobre todo en literatura, francés e historia. La diferencia entre centros era mayor con respecto al rendimiento general, mientras que en inglés era especialmente sensible al efecto de la familia.

En cuanto a la estabilidad en el tiempo de los efectos escolares y departamentos docentes, los resultados apuntaban que la puntuación global era bastante estable (una correlación media de 0,88 entre cohortes), mientras que entre materias existían diferencias, siendo francés la más inestable (aproximadamente 0,4 de correlación) e historia la más estable (0,8).

Tanto para la puntuación global como para la mayoría de las materias se encontraron efectos diferenciales para las cuatro variables analizadas (rendimiento previo, género, grupo étnico y situación socioeconómica). Esto indica que para medir la eficacia de un centro no es suficiente analizarlo mediante el cálculo de la media global de las puntuaciones de los alumnos, sino que es necesario analizarlo para los distintos grupos de alumnos.

En la segunda parte se encontró que las escuelas y los departamentos eficaces compartían las siguientes características:

1. Énfasis académico;

2. Altas expectativas;
3. Una visión compartida entre docentes;
4. Un equipo directivo eficaz;
5. Enseñanza de buena calidad; y
6. Alto nivel de interés e implicación de los padres

3. Otras líneas de actuación de la Eficacia Escolar

Tradicionalmente la Eficacia Escolar y la Mejora Escolar (Fullan, 1991; Hargreaves et al., 1998; Bolívar, 1999 y Murillo y Muñoz-Repiso, 2003) han tenido un desarrollo histórico paralelo. De hecho la Mejora Escolar surgió como un enfoque distinto que buscaba un cambio educativo planificado, a través del cual los resultados de los estudiantes mejoraban y fortalecía la capacidad de la escuela para gestionar cambios (Filp et al., 1984).

Actualmente, a partir de Hopkins y Reynolds (2002), se establecieron unos estándares comunes, por lo que ambas corrientes confluyeron en un nuevo enfoque llamado *Movimiento de la Mejora en Eficacia Escolar* (Reynolds et al., 1996; Murillo, 2001; Muñoz-Repiso y Murillo, 2003). Desde esta perspectiva se entiende que los factores encontrados por la Eficacia Escolar pueden ser incorporados en líneas de actuación de la Mejora Escolar, suponiendo un beneficio mutuo (Creemers y Reezigt, 1997; Reynolds, Hopkins y Stoll, 1993). Así la Eficacia proporciona experiencias y conocimientos que se incorporan en proyectos de Mejora, y a su vez, la Mejora puede validar experiencias de eficacia y nuevas posibilidades de factores de eficacia (Creemers et al., 2007).

Entre diversos trabajos destaca *Success for all* (Slavin et al., 1996) y *Capacity from Change and Adaption in the case of Effective School Improvement (ESI) Project* desarrollado entre 1998 y 2001 en el que participaron ocho países entre los que se encontraba España (Muñoz-Repiso y Murillo, 2003). Más recientemente cabe señalar a Murillo (2011) donde se elabora un modelo de mejora de la eficacia escolar para Iberoamérica y dónde también se incluye a España. Dicho trabajo, en parte, es bibliométrico y conjuga tres estudios de diferentes aproximaciones metodológicas (comparación de factores del contexto del sistema educativo; revisión de

investigaciones de factores de eficacia de proceso y opinión de investigadores sobre cambio escolar). Estas dos últimas investigaciones se tratarán con más detalle en el capítulo de la eficacia en España.

Esta actual línea tiene una gran influencia en Inglaterra (Reynolds, 2007). De hecho es palpable en políticas nacionales, en marcos de inspección de OFSTED (Sammons, Hillman y Mortimores, 1995) y en Instituciones de Educación Superior que ofrecen desarrollo profesional a docentes en estas temáticas con la creación de la *Red de Mejora Escolar (School Improvement Network (SIN))*.

Por otra, al amparo de las corrientes comentadas, ha surgido una nueva línea más específica y que también trata de mejorar la calidad de la educación y encontrar factores de aula que contribuyen más eficazmente a que los alumnos aprendan. Este nuevo enfoque toma el nombre de *Enseñanza Eficaz (Effective Teaching)*, y sus resultados están aportando interesantes elementos para la reflexión sobre cómo tiene que desarrollarse la acción educativa en el aula para conseguir criterios de mejora de la calidad y equidad de la educación (Murillo, Martínez-Garrido y Hernández-Castilla, 2011). Algunos autores se pueden enmarcar bajo esta línea como Wittrok (1986), Teddlie, Kirby y Springfield (1989), Kaplan y Owings (2002), Roman (2008), Borich (2009), Brown (2009), Good, Wiley y Florez (2009) Orlich et al. (2010); Murillo, Martínez-Garrido y Hernández-Castilla (2011)

Por último también se encuentra la influencia de la IEE en políticas a través de encuestas internacionales procedentes de la IEA o de la OCDE, tales como “*Third International Maths and Science Study*” (TIMSS) o “*Program for International Student Achievement*” (PISA) y mediante evaluaciones de sistemas educativos que responden a mandatos legislativos (e.g. Porter, 1991; Fitz-Gibbon, 1996 y Willms, 1992) que han sabido aplicar conocimiento tanto conceptual como metodológico procedente de la IEE (Kyriakides, 2005 y Reynolds, 2007).

4. Conclusiones

Este capítulo muestra un recorrido cronológico y los principales estudios que componen la línea de la Eficacia Escolar. El objetivo de este movimiento es construir una teoría capaz de integrar los elementos que ayudan a que una escuela sea eficaz (Reynolds et al., 2005 y Townsend, 2007). La definición de escuela eficaz que se toma

como referencia es la que propone Murillo (2005) según la cual, un centro eficaz es aquel que promueve el desarrollo integral de cada uno de sus alumnos teniendo en cuenta el rendimiento previo y el contexto social.

La figura 6 representa una síntesis de los distintos modelos planteados en cada una de las etapas del Movimiento de la Eficacia Escolar (Reynolds et al., 1999). Se puede ver la evolución del modelo *Input-Output* del Informe Coleman (Coleman, 1966) hasta modelos posteriores más elaborados, incluyendo la inclusión de variables de proceso (Rutter, 1979; Teddlie y Stringfield, 1989 o Mortimore, Sammons y Stoll, 1987) y la construcción modelos contextualizados gracias a la inclusión de variables de contexto y al uso de una metodología multinivel propuesta por Aitkin y Longford (1986). De esta etapa son el estudio de Sammons, Nuttall y Cuttance (1993); Sammons, Thomas y Mortimore (1997) entre otros. Por último destaca la influencia de esta línea en políticas educativas mediante proyectos internacionales como PISA; su confluencia con proyectos de Mejora Escolar, generando el Movimiento de la Mejora en Eficacia Escolar y el actual hincapié en el estudio de factores de aula.

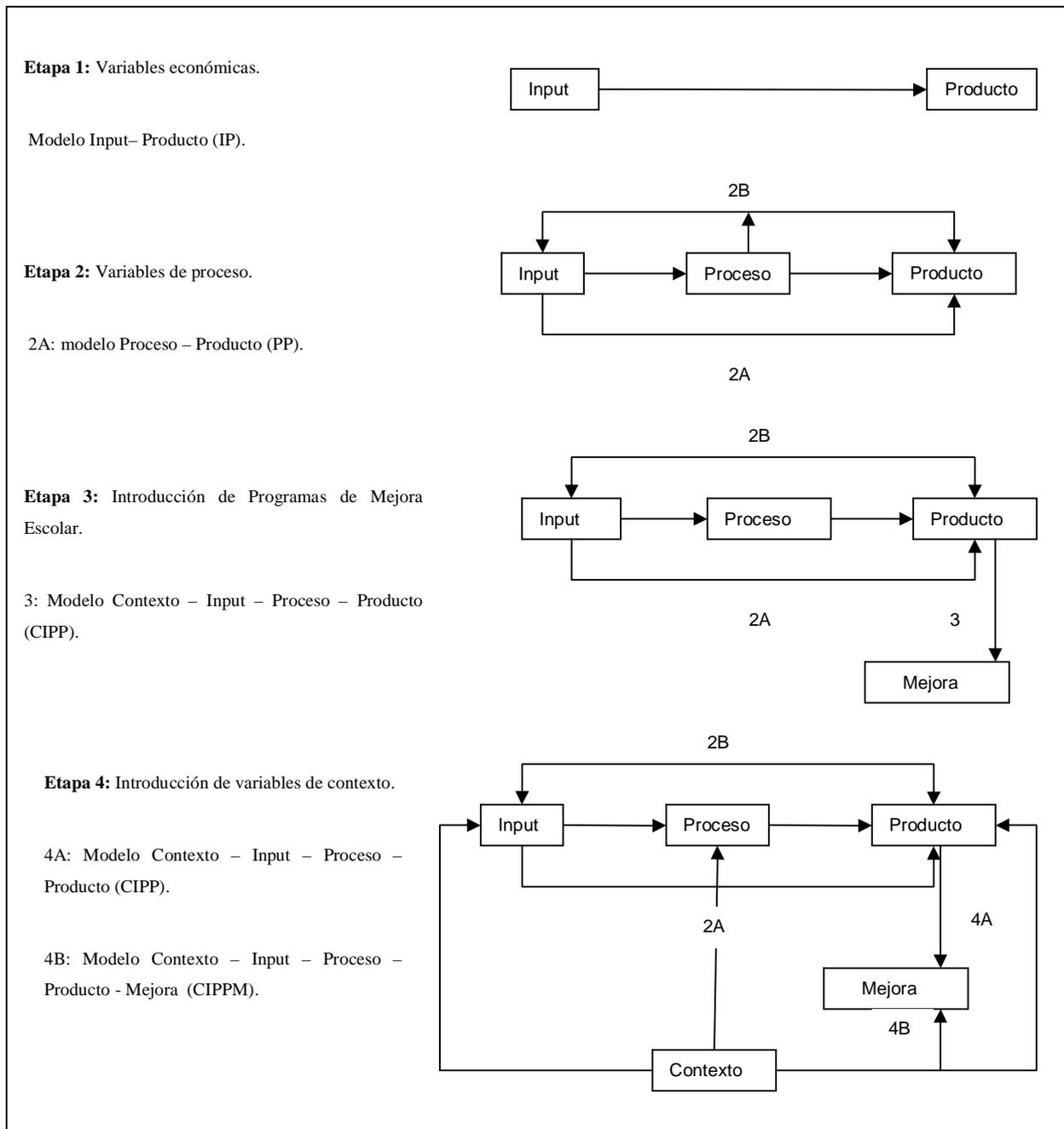


Figura 6: *Síntesis de modelos de Eficacia Escolar*. Fuente tomada: Reynolds et al (1999 p.5).

5. Índice de tablas

Tabla 1: <i>Factores de éxito académico (Weber, 1971)</i> . Fuente: elaboración propia.	25
Tabla 2: <i>Modelo de cinco factores de Edmonds (1979)</i> . Fuente: elaboración propia. ...	26
Tabla 3: <i>Correlaciones significativas entre rendimiento y clima escolar (Brookover et al., 1979)</i> . Fuente: Murillo (2005)	27
Tabla 4: <i>Mapa de variables de Fifteen Thousand Hours (Rutter, 1979)</i> . Fuente: Murillo, 2005.	29
Tabla 5: <i>Estimación de la puntuación para centros públicos y pendientes estimadas para centros privados</i> . Fuente: Murillo (2005)	30
Tabla 6: <i>Contraste entre características de eficacia del LSES</i> . Fuente: Teddlie y Stringfield (1993, p. 132)	33
Tabla 7: <i>Mapa de variables de Junior School Project</i> . Fuente: Elaboración propia.....	34

6. Índice de figuras

Figura 1: <i>Diferentes propuestas de organización de la eficacia escolar</i> . Fuente: elaboración propia.....	21
Figura 2: <i>Modelo propuesto en el Informe Plowden (1967)</i> . Fuente: Murillo (2005)....	24
Figura 3: <i>Modelo conceptual de Brookover et al. (1979)</i> . Fuente: Murillo (2005).....	27
Figura 4: Resultados de Victorian Quality of Schools Project. Fuente: tomada de Murillo (2005).....	41
Figura 5: Mapa de variables del estudio de Grisay (1996). Fuente: Elaboración propia	42
Figura 6: <i>Síntesis de modelos de Eficacia Escolar</i> . Fuente tomada: Reynolds et al (1999).	48



CAPÍTULO 2

MODELOS DE EFICACIA ESCOLAR

Capítulo 2: Modelos de Eficacia Escolar

1. Efectos Escolares	52
2. Modelos de efectos escolares	54
2.1. El modelo integrado de Eficacia Escolar	55
2.2. Modelo QAIT/MACRO	57
2.3. Modelo de eficacia docente	58
2.4. Modelo de Eficacia de Departamento	59
2.5. Modelo de Eficacia Escolar de EGB	62
2.6. Modelo iberoamericano de Eficacia Escolar	63
2.7. Modelo dinámico y multidimensional de eficacia escolar	65
2.7.1. <i>Dinamismo</i>	66
2.7.2. <i>Configuración multidimensional</i>	66
2.7.3. <i>Factores de aula</i>	67
2.8. Modelo de Mejora de la Eficacia Escolar (ESI)	69
3. Conclusiones	71
4. Índice tablas	71
5. Índice figuras	71

1. Efectos Escolares

El análisis cronológico desarrollado en el anterior capítulo no aporta una visión clara de cuales son los factores escolares que contribuyen a que un centro sea eficaz. Algo que es fundamental para la construcción de modelos, capaces de generar una teoría comprensiva de la EE. En la literatura se recogen varias revisiones de investigaciones con el ánimo de encontrar esos factores comunes, como el metaanálisis propuesto por Murillo (2005). La tabla 8 es una síntesis que hace este autor de seis revisiones realizadas en Estados Unidos, Inglaterra y Países Bajos, donde encuentra las siguientes características comunes en estos estudios:

1. Metas compartidas, consenso, trabajo en equipo, configurando el sentido de comunidad;
2. Liderazgo educativo;
3. Clima escolar y de aula;
4. Altas expectativas;
5. Calidad del currículo/estrategias de enseñanza;
6. Organización del aula;
7. Seguimiento y evaluación;
8. Aprendizaje organizativo/ desarrollo profesional;
9. Compromiso e implicación de la comunidad educativa.

Estos factores incluso se han encontrado en contextos diferentes como países en vías de desarrollo, que configuran una nueva línea de trabajo dentro del Movimiento (e.g. Riddell, 1989; Pennyquick, 1993; Fuller y Clarke, 1994; Welch 2000; Scheerens; 2001 y Murillo, 2003). Además en estos países se encuentra un décimo factor distintivo, el cual resume indicadores de calidad y adecuación de las instalaciones, mobiliario, material didáctico y recursos económicos como es:

10. Recursos educativos

En relación con el décimo indicador destaca el trabajo de Murillo y Román (2011), cuyo propósito era determinar la incidencia de los servicios y recursos educativos de los centros escolares en tercer y sexto grado de Primaria. La muestra

estaba compuesta por 180.000 estudiantes y 3.000 centros de 15 países de América Latina procedente de un estudio de la UNESCO. Los resultados mostraron que la disponibilidad de infraestructuras y servicios adecuados como agua, electricidad..., contar con medios didácticos como instalaciones deportivas, laboratorios etc., así como disponer de un buen número de libros en las bibliotecas y ordenadores en el centro mejora el rendimiento en los países evaluados.

Tabla 8: *Resumen de factores de eficacia escolar.* Fuente: Murillo (2005 p.212).

Edmonds (1979)	Purkey y Smith (1990)	Levine y Lezotte (1990)	Cotton (1995)	Sammons, Hillman y Mortimore (1995)	Scheerens y Bosker (1997)
	Metas claras sobre destrezas básicas		Metas planificadas de aprendizaje	Visión y metas compartidas	Consenso y cooperación entre el profesorado
Atmósfera ordenada y tranquila	Atmósfera cooperativa y clima ordenado	Clima y cultura productiva		Entorno de aprendizaje y refuerzo cooperativo	Clima escolar y de aula
Adquisición de destrezas y habilidades básicas	Política orientada al rendimiento	Foco en las destrezas de aprendizaje	Énfasis en el aprendizaje	Concentración en la enseñanza y el aprendizaje	Orientación hacia el rendimiento
Evaluación constante y regular del progreso del alumnado	Evaluación frecuente	Seguimiento adecuado	Evaluación	Seguimiento del progreso	Potencial evaluativo
		Desarrollo del profesorado orientado a la práctica	Desarrollo profesional Aprendizaje cooperativo	Organización del aprendizaje	
Liderazgo fuerte	Liderazgo fuerte	Liderazgo sobresaliente	Gestión y organización escolar, liderazgo y mejora, liderazgo y planificación	Liderazgo profesional	Liderazgo educativo
		Implicación de los padres	Implicación de la comunidad		Implicación de los padres
	Tiempo en tareas de refuerzo		Gestión y organización del aula	Relación hogar-escuela	Gestión del aula
		Adaptaciones instructivas eficaces	Diseño y planificación del currículo	Enseñanza con propósito	Calidad de currículo, instrucción estructurada y adaptativa
Altas expectativas	Altas expectativas	Altas expectativas	Interacciones profesor-alumno	Altas expectativas	Altas expectativas
				Derechos y responsabilidades del alumno	
			Interacción escuela-distrito		
			Equidad		
			Programas especiales		
					Feedback y refuerzo
					Aprendizaje independiente

En otro estudio Murillo (2007) analiza los sistemas educativos de nueve países (Bolivia, Chile, Colombia, Cuba, Ecuador, España, Panamá, Perú y Venezuela) encontrando el décimo factor con un efecto significativo. El logro académico en este trabajo se relaciona con el uso y gestión de los recursos e instalaciones y las condiciones económicas y laborales de los docentes, mientras que el factor evaluación obtenía un efecto residual. El autor atribuye la causa de la escasa incidencia de la evaluación a la casi nula tradición de evaluación de centros y acciones autoevaluadoras.

Además de los diez factores mencionados destacan aquellos relacionados con el docente. El proyecto ISERP (Internacional School Effectiveness Research Project) desarrollado por Reynolds, Creemers, Stringfield, Teddlie y Schaffer (2002) hace una comparación entre países de habla no inglesa y países de habla inglesa. Como aspectos significativos encuentran en países de habla no inglesa factores de eficacia por ejemplo el estilo de instrucción, currículo, experiencias previas y la influencia familiar. Mientras que en países de habla inglesa destaca el efecto de la dirección, las expectativas de los estudiantes, las metas y las relaciones entre docentes. No obstante esta distinción es un tanto forzada, porque el constructo de eficacia está medido en términos concretos en países no ingleses y en términos más abstractos en países de habla inglesa (Garber, 2003). Pero críticas a parte, destaca el valor que Reynolds et al. (2002) dan a una serie de variables de proceso que implican directamente al docente y a sus prácticas y que inciden en los resultados académicos. Los comportamientos docentes que encuentran se refieren al empleo de un feedback positivo, lecciones estructuradas, importancia de la evaluación, motivar al alumnado a probar y elaborar sus respuestas y generar altas expectativas de logro (Creemers, Stringfield y Guldmond, 2002).

2. Modelos de efectos escolares

El movimiento de investigación sobre Eficacia Escolar intentó desde el primer momento elaborar una teoría (Townsend, 2007). Sin embargo es un requisito elaborar un modelo explicativo y global capaz de especificar o visualizar de una forma simplificada o reducida un fenómeno que no es fácilmente o directamente observable (Sheerens, 1992, p.13)

La definición que se va a tomar de modelo es la que propone Murillo (2008, p.4) entendiéndolo por modelo *“una imagen simplificada y gráfica de un conjunto de*

unidades (hechos, conceptos y variables), un sistema de relaciones entre las mismas y unas interpretaciones globales y predictivas de sucesos empíricos, con el objetivo de explicar relaciones entre fenómenos”.

Scheerens y Creemers (1989) fueron los pioneros en esbozar ideas clave que componen posteriores modelos de IEE, tales como: a) partir de teorías instructivas y del aprendizaje; b) modelos que expongan lo que se sabe y que además planteen hipótesis que pueda orientar a investigaciones posteriores; y c) reflejar elementos distribuidos en tres niveles (centro, aula y alumno), modulados por características del contexto y relaciones mutuas. El esbozo propuesto se presenta en la figura 7.

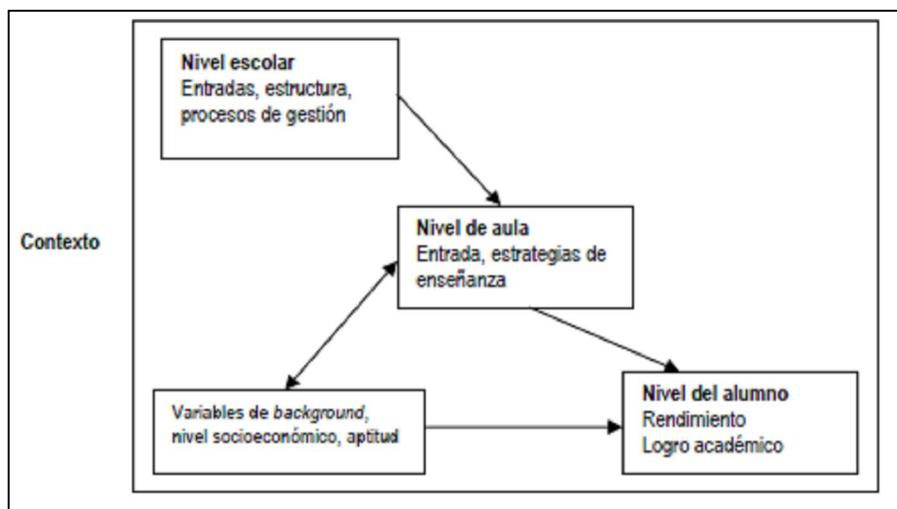


Figura 7: *Modelo de Eficacia Escolar de Scheerens y Creemers (1989)*. Fuente: Murillo (2008 p.5).

Los siguientes epígrafes tratan de describir los modelos más relevantes y con más influencia siguiendo un orden cronológico.

2.1. El modelo integrado de Eficacia Escolar

El modelo de Scheerens (1990) es conocido como el *Modelo Integrado de Eficacia Escolar* por ser una propuesta que combina las aportaciones de distintos enfoques. Con el fin de entender el funcionamiento de las organizaciones del sector público combina desde principios básicos de la Teoría de la Contingencia, los desarrollos en el campo de la Teoría de las Organizaciones hasta conocimientos de la Teoría Microeconómica. Bajo este prima se entiende, por una parte, que la eficacia de estructuras organizativas o de los procesos de gestión dependen de las características contextuales, tales como las restricciones ambientales y tecnológicas; y, por otro, que

las condiciones meso (organización de la escuela) parecen facilitar las condiciones micro (nivel alumno) (Murillo, 2008).

Entre las aportaciones de Scheerens son sus *principios ordenados*, los cuales se van a enumerar y que están presentes en el modelo, figura 8.

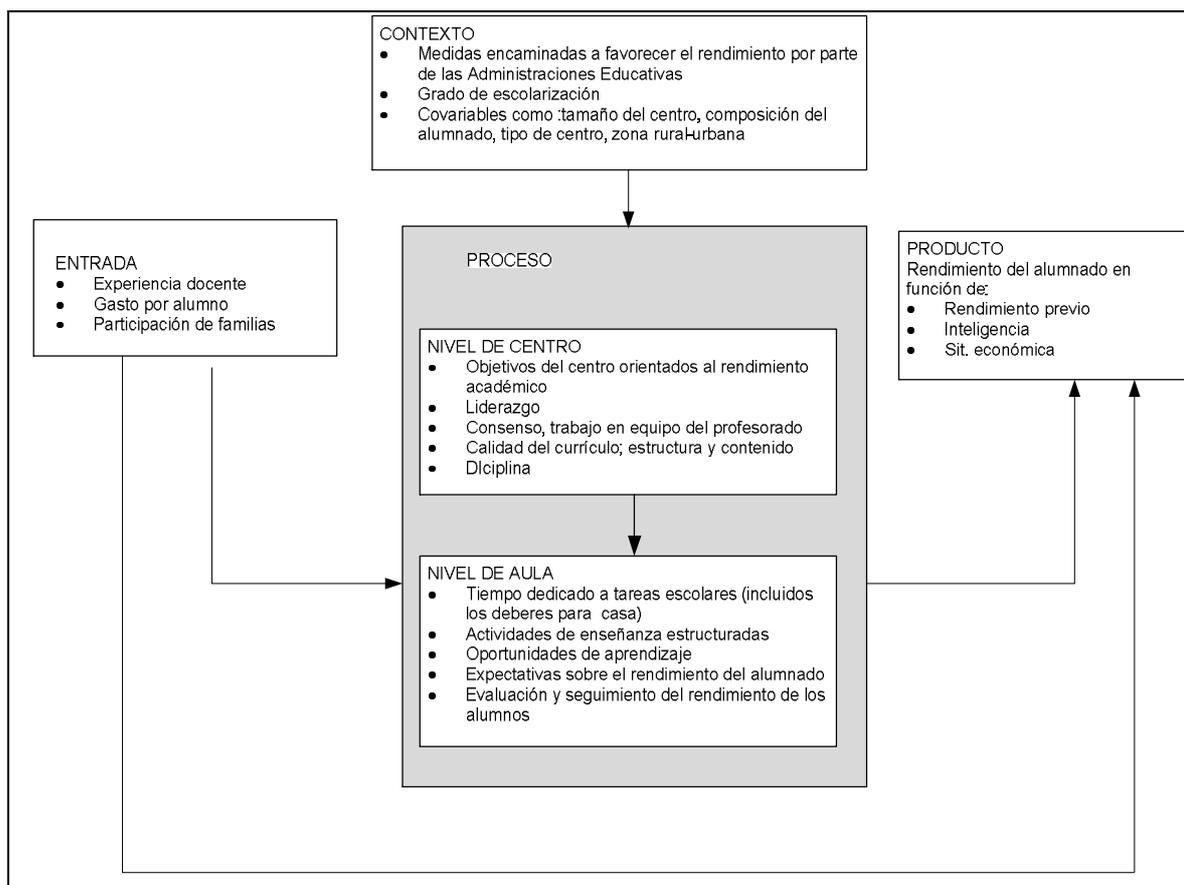


Figura 8: *Modelo de Eficacia escolar de Scheerens (1990 p.73)*. Fuente: Murillo (2008 p.7).

1. Elaboración de un marco general de referencia para determinar la posición de indicadores de proceso en un modelo analítico de sistemas, contexto, entrada, proceso y variables de resultado.
2. Una estructura multinivel, capaz de dar cuenta de indicadores de proceso definidos desde los niveles de entorno escolar, centro y aula. Además anclar los indicadores de proceso a los resultados educativos, lo cuales han de estar medidos en el nivel más bajo de agregación.
3. Disponer de una perspectiva teórica que permita señalar las interrelaciones entre las variables en diferentes niveles.

2.2. Modelo QAIT/MACRO

El modelo conocido por los acrónimos QAIT/MACRO fue propuesto por Stringfield y Slavin (1992), figura 9. QAIT indica términos de Calidad, Adecuación, Incentivo y Tiempo (*Quality, Appopriateness, Incentive y Time*); y MACRO se refiere a Metas con propósito, Atención a lo académico, Coordinación, Selección y Formación, y Organización (*Meaningful goals, Attention to acedemic focus, Coordination, Recruitment and training, y Organization*). Está compuesto por cuatro niveles:

1. El alumno influido por el modelo de Carroll (1963).
2. El nivel del aula se compone por los profesionales que están en interacción directa con los alumnos. Este nivel procede, fundamentalmente, de teoría de la organización del aula eficaz de Slavin (1987). Según la misma, este nivel estaría conformado por los cuatro elementos del modelo de Carroll potencialmente controlados por el profesor: Calidad (oportunidad para aprender, tiempo dedicado a la tarea y comportamientos eficaces del profesor), Adecuación (nivel de dificultad de la materia), Incentivo (estimulación de la motivación para aprender) y Tiempo dedicado a la tarea.
3. El nivel del centro educativo se centra en la dirección, otro personal del centro y desarrollo de programas. Aquí destacan dos elementos: los recursos humanos (establecimiento de metas y que éstas sean compartidas por todo el personal del centro, el liderazgo instructivo, la selección y formación del profesorado); y las características de la organización escolar (coordinación, organización de la jornada lectiva y uso del personal de apoyo).

4. El nivel del contexto, nivel superior al centro, que incluye la comunidad, el distrito escolar y resto de autoridades educativas. Se centra en las relaciones con las familias y la comunidad local, el distrito escolar, los programas especiales y la financiación de la escuela definida por los gobiernos federal y estatal.

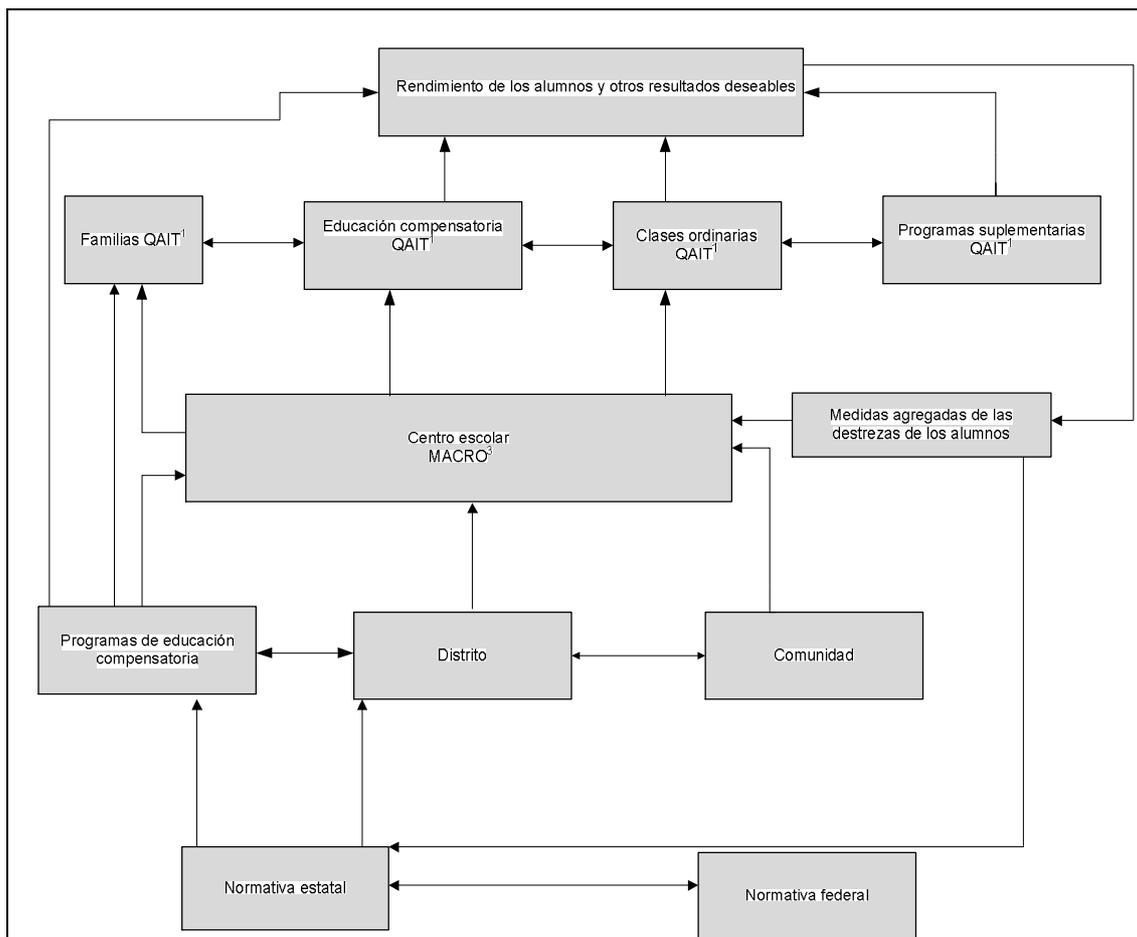


Figura 9: *Modelo de eficacia escolar de Stringfield y Slavin* (1992, p.122). Fuente: Murillo (2008 p.8).

- (1) QAIT: Calidad, adecuación, incentivo y tiempo para la instrucción.
- (2) Educación especial, educación bilingüe, etc.
- (3) MACRO: metas con propósito, atención a lo académico, coordinación, selección y formación, organización.

2.3. Modelo de eficacia docente

El modelo de eficacia docente de Creemers (1994) es el sucesor del *modelo de productos educativos* (Creemers, 1992) en el cual confluyen la investigación holandesa sobre Eficacia Docente y Eficacia Escolar. El modelo que se va a presentar es novedoso porque refleja resultados empíricos aplicando una clara distinción entre los factores

correspondientes al sistema educativo, al centro, al aula y al alumno. De entre los distintos niveles se centra en el nivel de aula. Aunque esta prioridad se verá de forma más clara con el desarrollo de un modelo propuesto por Creemers y Kyriakides (2008) que posteriormente se expondrá.

En el modelo de la figura 10 se distinguen cuatro niveles anteriormente citados y tres tipos de componentes: calidad, tiempo y oportunidad que se repiten en cada sustrato y con diferentes criterios de eficacia dependiendo del nivel que se considere.

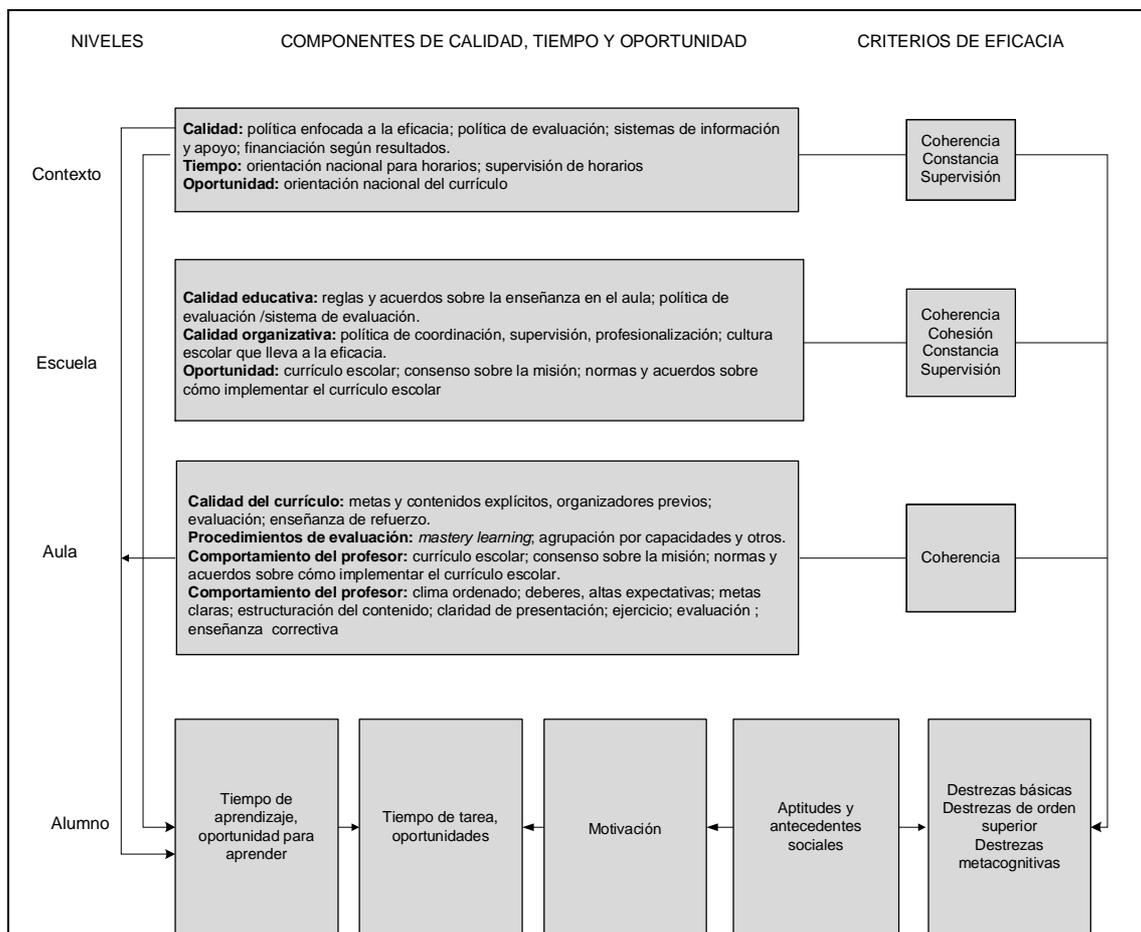


Figura 10: *Modelo de Eficacia Docente* (Creemers, 1994). Fuente: Murillo (2008 p.10)

2.4. Modelo de Eficacia de Departamento

El modelo propuesto por Sammons, Thomas y Mortimore (1997) supone una ampliación del modelo de Creemers (1994) dado que se incluye un nivel de departamento, considerado como un sustrato fundamental y responsable de la eficacia diferencial en centros de Secundaria, figura 11.

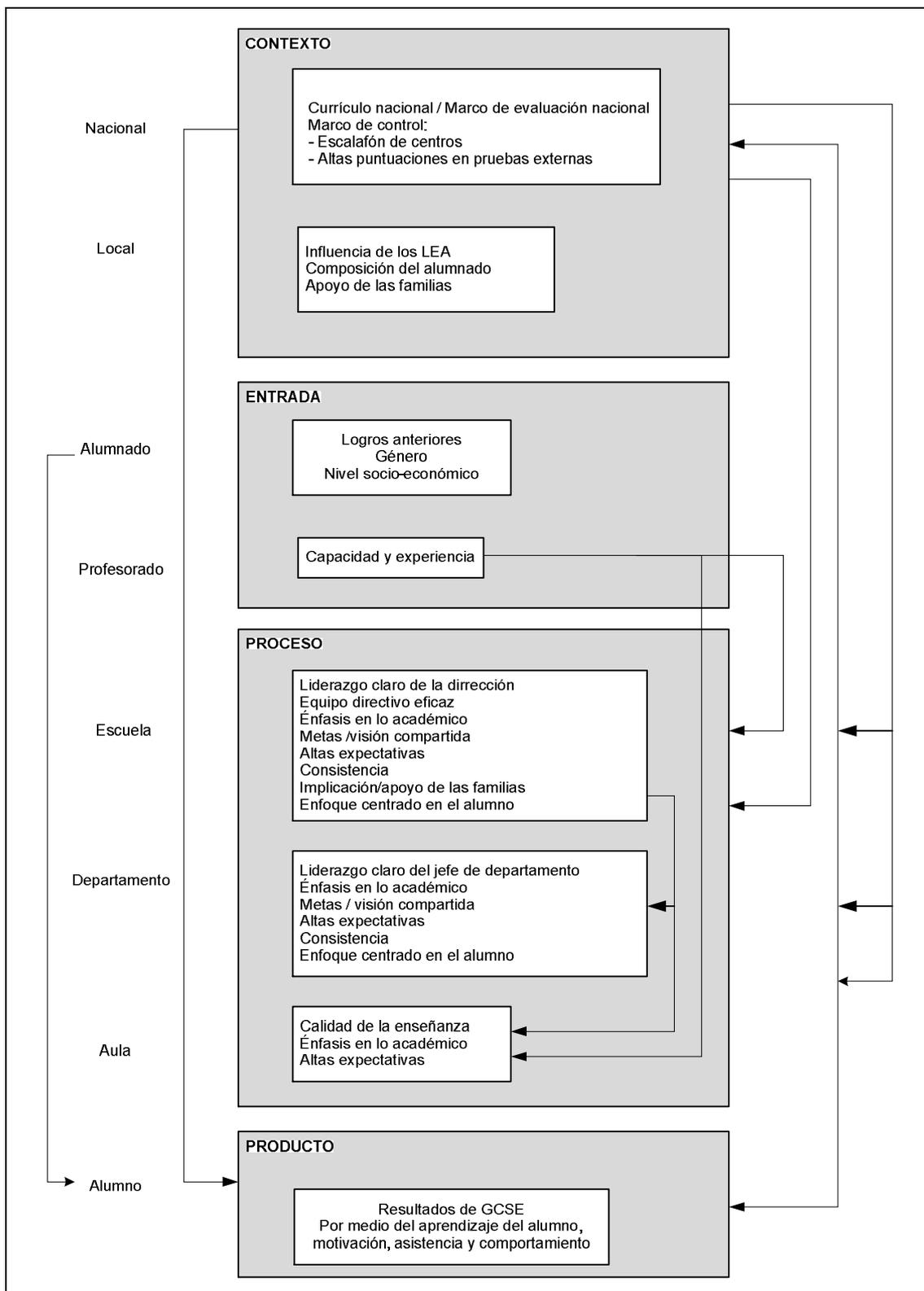


Figura 11: *Modelo de Eficacia de Departamento* (Sammons, Thomas y Mortimore, 1997).

Fuente: Murillo (2008 p.12)

Nota: GCSE: General Certification in Secondary Education.

Estos autores afirman que una de las características de las escuelas eficaces es la congruencia entre los factores que operan en diferentes niveles (escuela, departamento y aula). Por ejemplo el énfasis académico y las altas expectativas en los tres niveles, mientras que la consistencia, la visión y metas compartidas y la orientación centrada en el alumno están presentes en dos niveles (centro y departamento). A su vez los distintos niveles se relacionan a través de los distintos agentes. Así la gestión del centro incide en la eficacia académica a través del director; y el funcionamiento del equipo directivo y el liderazgo de la jefatura se relaciona con el nivel de departamento.

Los procesos de aula, especialmente la calidad de la enseñanza, ejercen un impacto directo sobre el aprendizaje y la motivación de los alumnos, los cuales afectan a los resultados académicos. El comportamiento y la asistencia a clase, sin embargo, pueden estar influenciados por los procesos escolares y de aula. La capacidad y la experiencia de los docentes son entradas del proceso educativo, y tienen una influencia directa en la calidad de la enseñanza en el nivel de aula.

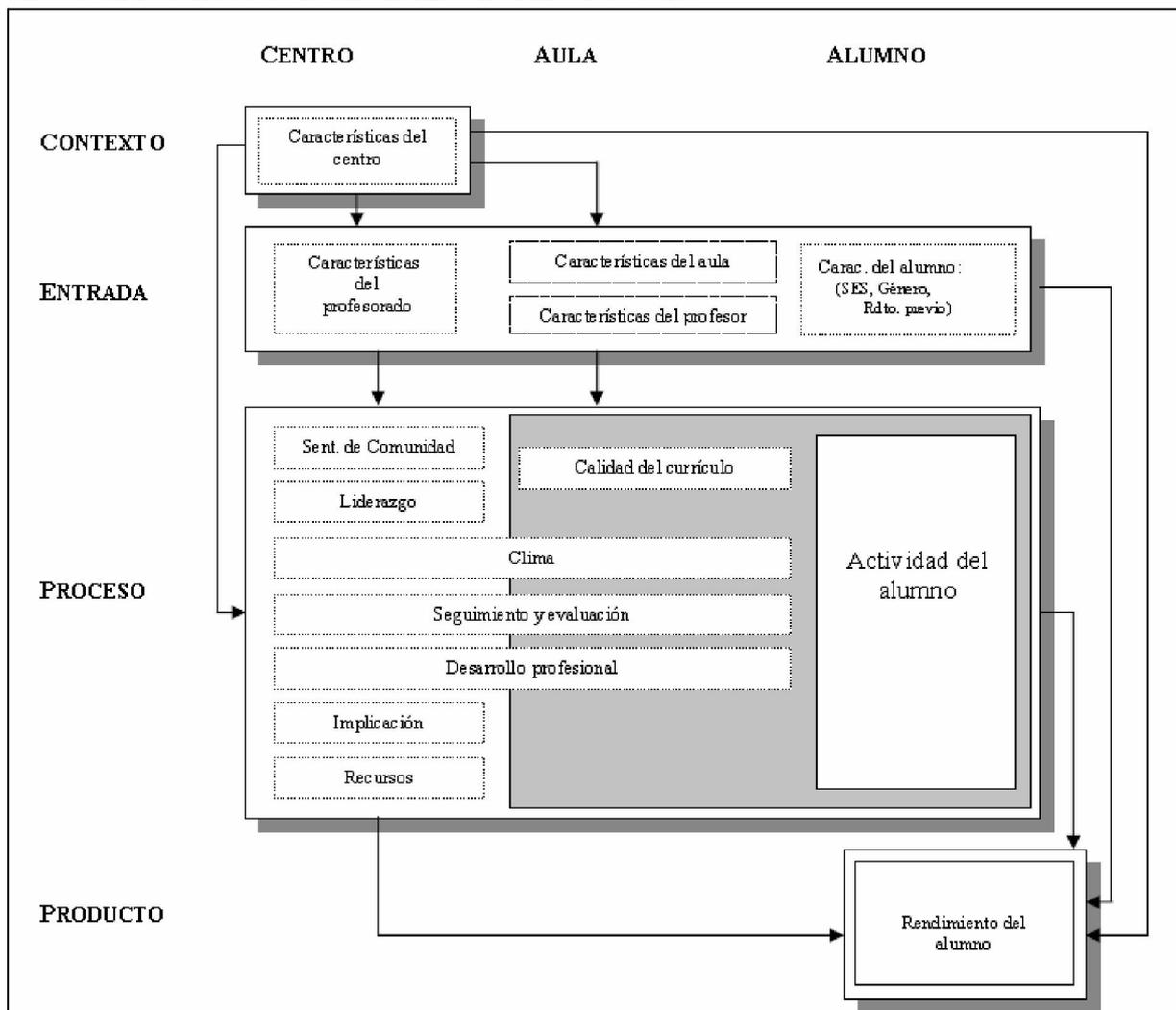


Figura 12: *Modelo de Eficacia Escolar en Primaria*. Fuente: Murillo (2008 p.22).

2.5. Modelo de Eficacia Escolar de EGB

Murillo (2008) diseñó un modelo multinivel de Eficacia Escolar para Educación Primaria, a partir de una muestra de 332 centros y 6.598 alumnos de sexto de Educación General Básica (EGB) en el curso 1994/95 en población española. Dicha evaluación medía conocimientos en lengua y matemáticas. La figura 12 presenta un modelo sistémico de factores de eficacia encontrados en el estudio, distribuidos en tres niveles (centro, aula y alumno)

Las relaciones entre los distintos indicadores de proceso se pueden ver en la figura 13. Estas relaciones se establecieron a partir de análisis de regresión tomando como nivel de significación ($p \leq 0,05$).

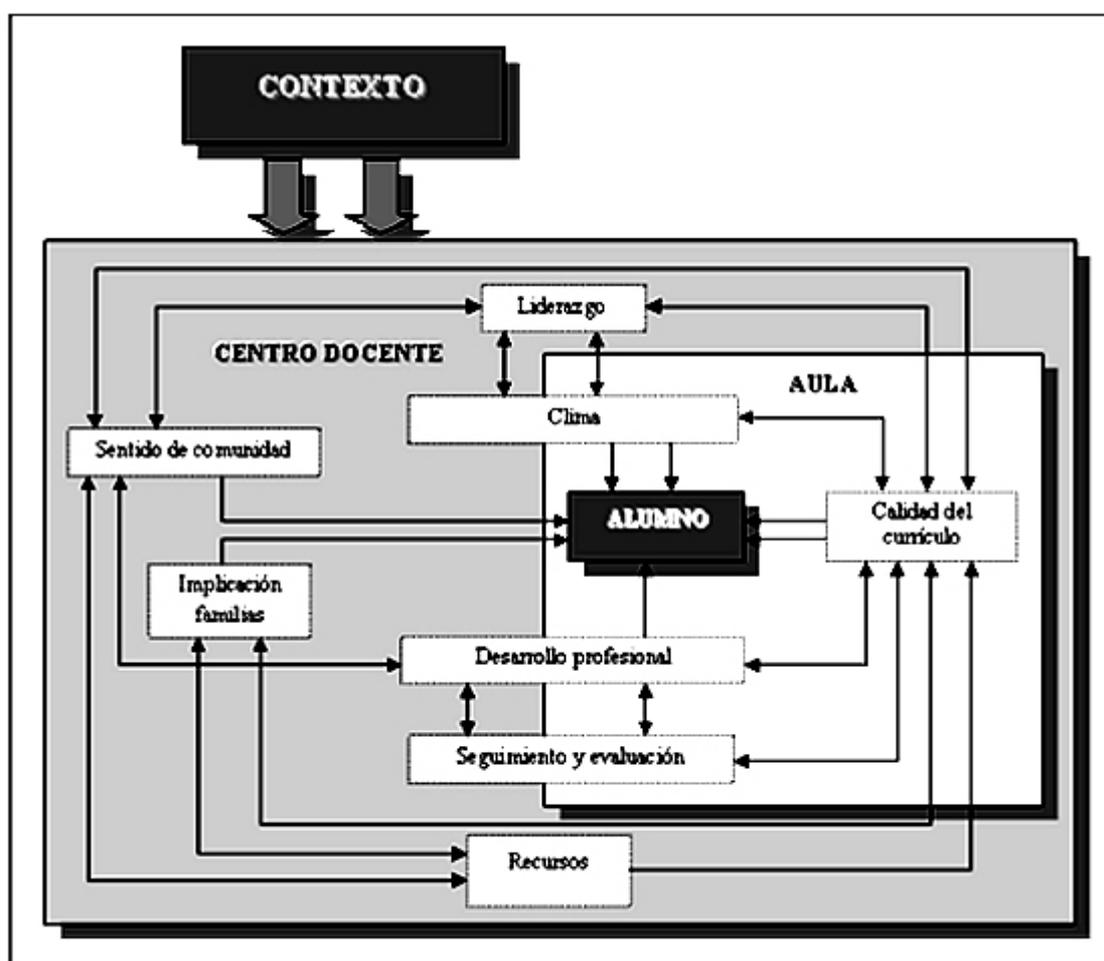


Figura 13: *Resumen de relaciones entre factores de proceso.* Fuente: Murillo (2008 p.25).

Entre los factores con más incidencia en el logro académico se encuentran la calidad del currículo y el clima de aula, pero también, aunque con menos fuerza, el desarrollo profesional de los docentes. A su vez estos tres factores se encuentran fuertemente relacionados entre sí, y también con el cuarto factor de aula -seguimiento y evaluación. De los cuatro factores de centro que aparecen en el modelo, dos están directamente relacionados con el rendimiento: el sentido de comunidad, es decir, la existencia de normas compartidas por la comunidad educativa, y la implicación de las familias, y dos más tienen una relación indirecta: el liderazgo y los recursos del centro.

2.6. Modelo iberoamericano de Eficacia Escolar

En este apartado se recoge la relación entre las variables del estudio que analizaba la situación de sistemas educativos de Iberoamérica y España comentado anteriormente (Murillo, 2007). La selección de factores del modelo procede de los resultados encontrados en dos tipos de aproximaciones metodológicas. Por una parte el estudio multinivel – de carácter cuantitativo - y por otra el estudio de eficacia percibida y el estudio de casos – de carácter cualitativo. La tabla 9 recoge los factores encontrados en las distintas aproximaciones.

Como se puede ver tanto el enfoque cuantitativo como el cualitativo han obtenido resultados coherentes y complementarios entre sí. Como señala el autor el estudio multinivel aporta información detallada, el trabajo de eficacia percibida ofrece resultados menos previstos y el estudio de casos genera una visión global y comprensiva de lo que acontece en los centros y lo qué los hace más o menos eficaces.

Tabla 9: *Factores asociados con el logro académico obtenidos en los tres estudios.*
Fuente: Murillo (2007 p. 255).

Estudio multinivel	Estudio de eficacia percibida	Estudio de escuelas prototípicas
Nivel de centro		
Características de centro	Características de centro	Características de centro
Características del profesorado	Características del profesorado	Características del profesorado
Objetivos del centro	Educación en valores	Misión del centro
Compromiso de docentes y trabajo en equipo	Implicación de docentes Trabajo en equipo	Implicación y compromiso de docentes
Clima escolar	Clima escolar	Clima escolar
Dirección escolar	Dirección escolar	Dirección escolar
Participación de familia y comunidad	Participación e implicación familiar	
Instalaciones y recursos	Instalaciones	Instalaciones

Satisfacción docente	Actitudes de los docentes		Expectativas
Nivel de aula			
Características del aula			
Características del docente			
Metodología docente	Metodología docente		Metodología docente
Gestión del tiempo	Gestión del tiempo		Gestión del tiempo
Clima de aula	Clima de aula		
Actitud de los docentes	Relación docente-alumnos		
	Satisfacción de los docentes		Organización del aula
Formación permanente del docente			
Implicación familiar	Participación e implicación familiar		
	Expectativas docentes		
Usos de instalaciones y recursos	Recursos disponibles		
Nivel de alumno			
Hábitos culturales			
Expectativas			Expectativas de la familia
Actitudes			
Relación con la familia			
Apoyo familiar y de compañeros			

Tras estos resultados se organizaron los distintos elementos bajo un modelo de cuatro niveles: Alumno, aula, centro y sistema, como se representa en la figura 14.

Según este modelo existen factores que están presentes en diferentes niveles, como el factor expectativas que se encuentra entre los elementos de proceso tanto en el nivel de centro como de aula o de alumno; o el clima que se encuentra en los niveles de centro (clima de centro) y de aula (clima de aula). En esta misma situación también aparecen el compromiso de los docentes, desarrollo profesional, implicación de la familia y uso y gestión de instalaciones y recursos.

Por otra parte incluye relaciones entre factores de distintos niveles como las metas escolares o el estilo directivo que influyen en los elementos de proceso del aula, por ejemplo en la metodología docente.

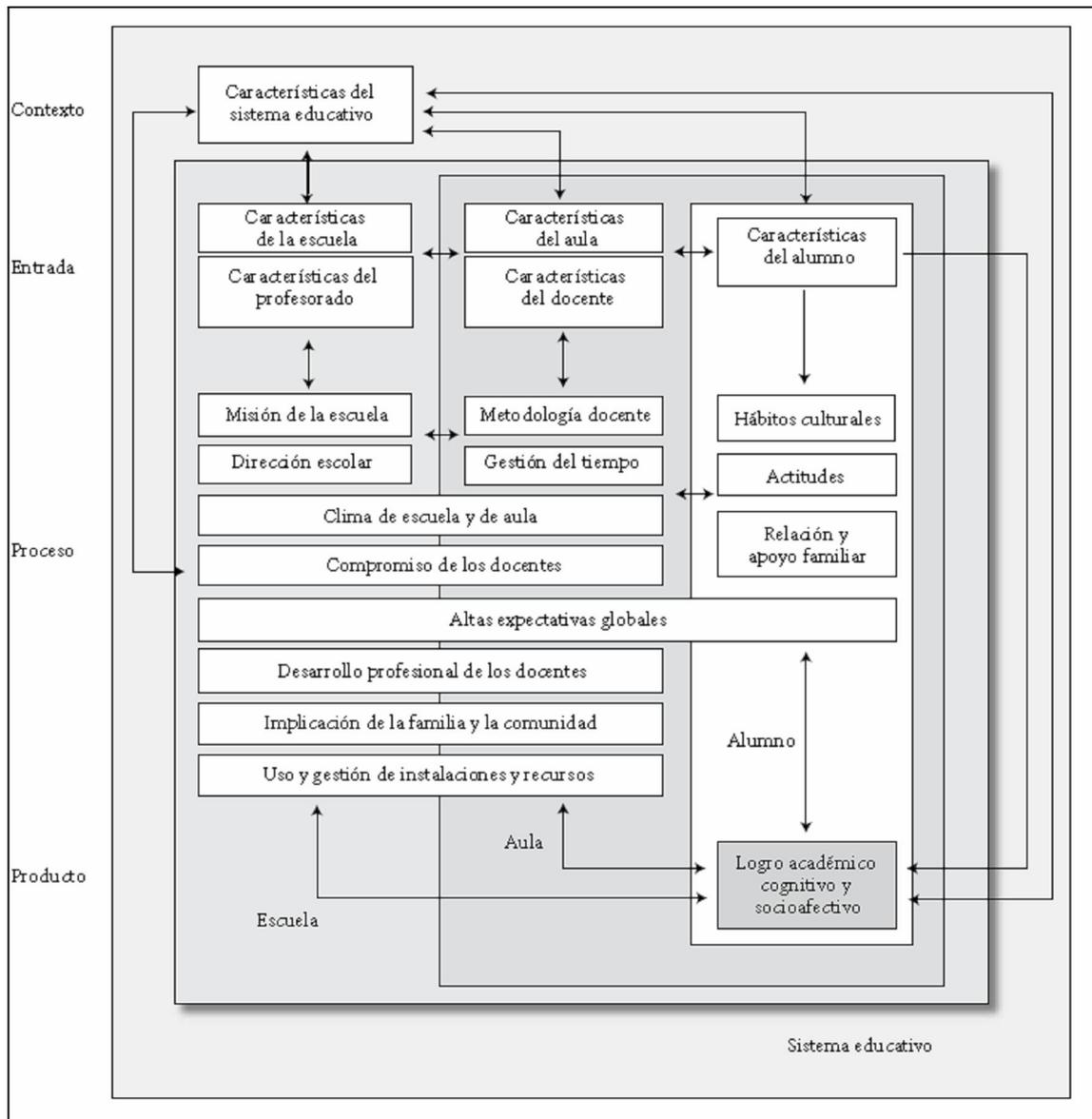


Figura 14: *Modelo de Eficacia Escolar en Primaria*. Fuente: Murillo (2008 p.22).

2.7. Modelo dinámico y multidimensional de eficacia escolar

El modelo de Creemers y Kyriakides (2006 y 2008) está configurado por factores con una naturaleza multidimensional operando dentro de una jerarquía por lo que la propuesta de estos autores tiene un carácter dinámico centrados en el aula (Kyriakides y Creemers, 2008), de hecho es un modelo que está muy relacionado con la corriente de *Enseñanza Eficaz*, como se comentó en el anterior capítulo. Los siguientes párrafos tratan de explicar estas dos peculiaridades: dinamismo y configuración multidimensional de los factores.

2.7.1. *Dinamismo*

Las relaciones entre factores es el aspecto más destacable en la visión de Creemers y Kyriakides. Advierten la necesidad de detectar relaciones curvilíneas, capaces de explicar “*evidencias problemáticas*” como el fenómeno que se da entre la frecuencia de evaluación y los resultados del alumnado. Un exceso de importancia en la evaluación produce un efecto igual de contraproducente que la ausencia de evaluación. Ya que en el primer caso reduce el tiempo efectivo de aprendizaje y la segunda situación implica no disponer de información necesaria para adaptar el tipo de enseñanza al tipo de alumnado (Creemers y Kyriakides, 2006).

Además instan a diseñar clusters de variables procedentes de distintos niveles, capaces de revelar combinaciones óptimas. El nivel de aula es donde concentran sus esfuerzos, concretamente en el comportamiento del docente. La causa de esta decisión es que el análisis del aula es un prerrequisito para definir niveles superiores (Creemers, 1994) y porque tomando como antecedentes estudios de la IEE, el aula incide especialmente en la Eficacia Escolar (e.g. Brophy y Good, 1986; Darling-Hammond, 2000; Doyle, 1990; Rosenshine y Stevens, 1986; Scheerens y Bosker, 1997; Wang, Haertel y Walberg, 1993).

2.7.2. *Configuración multidimensional*

Establece una serie de dimensiones comunes medibles en todos los factores:

1. *Frecuencia*. Es la cantidad de actividad (número de tareas) asociada con un factor de eficacia presentado.
2. *Foco*. Describe la función del factor en un rango de especificidad.
3. *Etapa*. Revela la parte de la lección o de la clase donde aparece el factor. Es esperable que los factores sean estables en el tiempo y presentes en distintas partes del proceso enseñanza-aprendizaje.
4. *Calidad*. Se refiere a las propiedades métricas y a la evaluación formativa de cada factor.
5. *Diferenciación*. Atención a las diferencias de los estudiantes en el progreso de aprendizaje. Los factores en esta dimensión se miden como ajuste a las necesidades del alumnado.

2.7.3. Factores de aula

El modelo se compone de ocho elementos que inciden en el rendimiento, estos factores derivan de estudios de Eficacia Docente (Brophy y Good, 1986; Creemers, 1994; Mujis y Reynolds, 2001; Rosenshine, 1983).

1. *Orientación.* Se refiere a la inclusión de tareas que suponen un reto y un compromiso de participación en los estudiantes. Este factor, como el modelo de enseñanza, está en la línea de promover el rendimiento hacia nuevas metas en educación tales como el desarrollo de habilidades metacognitivas.
2. *Estructuración.* Esta dimensión hace hincapié en la necesidad de mostrar una estructura clara de los contenidos, no sólo para facilitar la memorización de la información sino también porque facilita la integración y creación de relaciones entre las partes. De hecho Rosenshine y Stevens (1986) consideran que una estructura eficaz está compuesta por los siguientes pasos: a) mostrar objetivos; b) marcar las transiciones entre las partes; c) señalar principales ideas; d) revisar las principales ideas.
3. *Técnicas de preguntas.* Mujis y Reynolds (2000) indican que el docente eficaz es aquel que hace preguntas e intenta que el alumnado se involucre en discusiones de clase. Por tanto debe manejar una variedad de tipos de cuestiones, como el uso de preguntas producto (esperar una sola respuesta) o preguntas proceso que permitan a los alumnos dar explicaciones elaboradas. Incluso hay autores que señalan que el profesor eficaz se centra en este último tipo de cuestiones (Askew y William, 1995; Everston, Anderson, Anderson y Brophy, 1980).
4. *Modelo de enseñanza.* El profesor eficaz es aquel que ayuda al alumnado a desarrollar estrategias útiles y comunes ante distintos tipos de problemas (Kyriakides, 2002). Por tanto este factor se refiere a la enseñanza de estrategias que ayuden a organizar su propio aprendizaje.
5. *Práctica.* El docente eficaz es aquel que aprovecha el trabajo individual o tareas en pequeños grupos para aplicar contenidos enseñados en la clase (Rosenshine, 1983).
6. *La clase como ambiente de aprendizaje.* Este factor se centra en la contribución del profesor en crear un ambiente de aprendizaje. Para ello se toman cinco

índices que determinan el ambiente aprendizaje: interacción profesor alumno, interacción estudiante-estudiante, tratamiento profesor-alumno, rivalidad entre alumnos y conflicto en clase. Los dos primeros elementos definen el ambiente de aula (Cazden, 1986; Den Brok, Brekelmans y Wubbels, 2004; Fraser, 1991) y los tres elementos restantes dan cuenta del clima ordenado del aula (Brophy y Good, 1986; Hextall y Mahony, 1998; Scheerens y Bosker, 1997).

7. *Gestión del tiempo*. El modelo considera la participación del alumno y el tiempo de tarea como los dos factores más importantes de eficacia. Los profesores eficaces organizan y convierten el clima de clase en un ambiente de aprendizaje donde se encuentran elevadas tasas de participación (Creemers y Reezigt, 1996). Este modelo de gestión del tiempo es uno de los indicadores más importantes de la habilidad del docente.
8. *Evaluación del profesor*. La evaluación es una parte fundamental de la docencia (Stenmark, 1992) y especialmente la evaluación formativa es uno de los principales factores asociados a la eficacia particularmente en el nivel de aula (e.g. de Jong et al., 2004; Kyriakides, 2005; Shepard, 1989) La información reunida a través de la evaluación permite identificar las necesidades de sus estudiantes y conlleva evaluar la propia docencia.

El modelo teórico fue validado en una muestra de 2.503 estudiantes de 5 años, 108 aulas y 50 centros de Chipre. Para la recogida de datos aplicaron técnicas de observación de baja y alta inferencia tanto para medir conductas del alumnado como de los docentes. Mediante el empleo de técnicas de análisis multinivel, las variables de contexto controladas fueron el rendimiento previo, el sexo y el índice socioeconómico. Como indicadores de proceso de dicho nivel, se consideró cada factor medido en cada una de las dimensiones, tabla 10. Como variables dependientes se consideró el rendimiento en lengua griega, matemáticas y conocimientos de religión y, también, aspectos afectivos de religión. Los resultados revelaron que el modelo final, incluyendo todos los indicadores de contexto y de proceso, éstos explicaron más del 70% varianza en el nivel de aula.

Tabla 10: *Factores de de aula y dimensiones del modelo dinámico* (Creemers y Kyriakides, 2006). Fuente: elaboración propia.

Dimensiones						
Factores	Frecuencia	Foco	Etapas	Calidad	Diferenciación	
Orientación	Nº de tareas de orientación	Presente en toda o en parte de la lección	Fase de la lección en donde se aplica	Claridad e impacto en el proceso de aprendizaje	Tipo de orientación según el tipo de alumnado	
Estructuración	Porcentaje de tiempo dedicado a la estructura	Estructura en toda la lección o en partes del contenido	Fase de la lección en donde se aplica	Comprensión de los estudiantes. Secuencial	Tipo de orientación en función del alumnado	
Técnica de preguntas	Proporción entre distintos tipos de preguntas. Concordancia entre la pausa y dificultad de la pregunta	Coherencia de las preguntas con la lección	Manejo de preguntas en contenidos de dificultad. Preguntas que incluyen contenidos dados	Claridad de la pregunta. Pertinencia con el contenido. Tratamiento que hace el docente de la respuesta	Preguntar todo el alumnado. Adecuar el feedback a las necesidades del alumnado.	
Modelo de enseñanza	Número de tareas de modelado y tiempo empleado en ellas.	Estrategias aplicables en distintos contextos (materias, etc.)	Secuencia de las tareas	Claridad en enseñar la estrategia. Participación del alumnado	Adaptar las estrategias en función de las necesidades del alumnado	
Práctica	Tiempo aplicado en prácticas	Tareas que puedan realizar	Fase de la lección en la que se aplica	Autonomía para resolver la tarea. Práctica ajustada al contenido.	Oportunidad al alumnado con necesidades	
Ambiente de aula	<i>Ambiente de aula</i>	Nº y tipo de interacciones que promueven el aprendizaje	Propósito de las mismas	Fase de la lección que tienen lugar	Iniciativa del profesor para interactuar	Habilidad para mantener con distintos grupos las interacciones
	<i>Clima ordenado</i>	Nº de problemas en el aula	Tipo de problema y reacción del docente	Fase en la que se inicia el conflicto	Mediación del docente para resolver el problema	Solución de problemas causados por diferentes grupos
Gestión del tiempo	Duración de la clase		Tiempo dedicado a distintas partes de la lección	El rol del profesor en crear el ambiente de aprendizaje	Tiempo dedicado a los diferentes grupos	
Evaluación	Nº de tareas de evaluación	Diferentes tipos de pruebas	Momento de la evaluación. Tiempo entre la prueba y dar los resultados	Propiedades de las pruebas. Tipo de feedback constructivo	Tipos de evaluación y feedback en función de las necesidades del estudiante.	

2.8. Modelo de Mejora de la Eficacia Escolar (ESI)

Capacity from Change and Adaption in the case of Effective School Improvement (ESI) Project, es un intento de conjugar la mejora y la eficacia escolar de varios países entre los que participó España y cuyos resultados serán tratados en el capítulo de la Eficacia Escolar en España.

El concepto de mejora de la eficacia escolar utilizado en esa investigación se centro en desarrollar “*un cambio educativo planificado que incrementa los resultados del aprendizaje de los estudiantes, así como la capacidad de la escuela para gestionar el cambio*” (Creemers et al., 2007, p.826). La figura 15 ofrece una imagen sintetizada del marco comprensivo entre factores de eficacia y mejora encontrados en dicho proyecto sin distinción entre países.

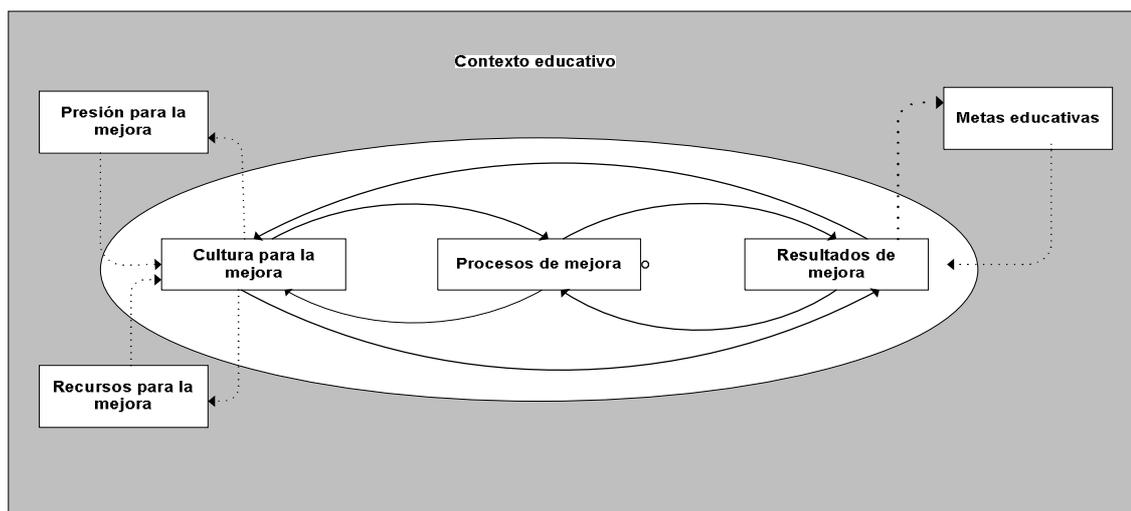


Figura 15: *Modelo de mejora de la eficacia escolar (ESI)*. Fuente: Creemers et al., (2007 p.830).

A nivel del centro educativo, los conceptos de *cultura de mejora*, *procesos de mejora* y *resultados de mejora* son esenciales. La cultura de mejora es el sustrato en el que los procesos acontecen. Los resultados de mejora son el objetivo de los centros que mejoran. Los conceptos de cultura, procesos y resultados están interrelacionados entre sí y se influyen constantemente al considerar que se trata de conceptos fundamentales, interdependientes y vinculados. Dichas interrelaciones demuestran que la mejora de la eficacia escolar es un proceso cíclico en movimiento, sin un comienzo o un final claramente marcados (Murillo, 2011).

Dentro del contexto, tres conceptos fundamentales están asociados a la mejora de la eficacia escolar: la *presión para la mejora*, los *recursos para la mejora* y las *metas educativas* existentes en el contexto educativo. Aunque en un proceso real son los centros los que han de diseñar su propio camino, sus objetivos, medios y actividades para conseguirlo, este deberá ser coherente con el contexto donde se encuentre.

3. Conclusiones

Se han presentado una serie de factores recogidos en distintas revisiones de trabajos publicados en Estados Unidos, Inglaterra y Países Bajos (Murillo, 2005), así como dos proyectos internacionales que tratan de encontrar factores comunes de eficacia en distintos contextos educativos (Murillo, 2007 y Reynolds et al., 2002). Por otra parte el capítulo incide en la presentación de distintos modelos que gozan de gran presencia en el Movimiento, con el fin de configurar una teoría de la Eficacia Escolar. Todos ellos tienen en común una visión sistémica del centro educativo; con unas relaciones e interdependencias en distintos niveles de análisis, por lo que implícitamente tienen un carácter dinámico; recogen investigación empírica y plantean hipótesis. Dentro de las peculiaridades destaca el modelo de Sammons, Thomas y Mortimore (1997) con la inclusión de un nivel de departamento; el modelo de Creemers y Kyriakides (2006 y 2008) centrado en la influencia del comportamiento docente y el modelo de Murillo (2007) por la combinación de técnicas cualitativas y cuantitativas.

4. Índice tablas

Tabla 8: <i>Resumen de factores de eficacia escolar</i> . Fuente: Murillo (2005 p.212).....	53
Tabla 9: <i>Factores asociados con el logro académico obtenidos en los tres estudios</i> . Fuente: Murillo (2007 p. 255).	63
Tabla 10: <i>Factores de de aula y dimensiones del modelo dinámico</i> (Creemers y Kyriakides, 2006). Fuente: elaboración propia.	69

5. Índice figuras

Figura 7: <i>Modelo de Eficacia Escolar de Scheerens y Creemers (1989)</i> . Fuente: Murillo (2008 p.5).	55
Figura 8: <i>Modelo de Eficacia escolar de Scheerens (1990 p.73)</i> . Fuente: Murillo (2008 p.7).	56
Figura 9: <i>Modelo de eficacia escolar de Stringfield y Slavin (1992, p.122)</i> . Fuente: Murillo (2008 p.8).....	58
Figura 10: <i>Modelo de Eficacia Docente</i> (Creemers, 1994). Fuente: Murillo (2008 p.10)	59
Figura 11: <i>Modelo de Eficacia de Departamento</i> (Sammons, Thomas y Mortimore, 1997). Fuente: Murillo (2008 p.12)	60

Figura 12: <i>Modelo de Eficacia Escolar en Primaria</i> . Fuente: Murillo (2008 p.22).....	61
Figura 13: <i>Resumen de relaciones entre factores de proceso</i> . Fuente: Murillo (2008 p.25).	62
Figura 14: <i>Modelo de Eficacia Escolar en Primaria</i> . Fuente: Murillo (2008 p.22).....	65
Figura 15: <i>Modelo de mejora de la eficacia escolar (ESI)</i> . Fuente: Creemers et al., (2007 p.830).	70

CAPÍTULO 3

PROPIEDADES DE LOS EFECTOS ESCOLARES

Capítulo 3: Propiedades de los efectos escolares

1.	Propiedades de los efectos escolares.....	75
1.1.	Magnitud de los efectos.....	75
1.2.	Consistencia de los efectos escolares.....	76
1.2.1.	Entre medidas cognitivas.....	77
1.2.2.	Entre medidas no cognitivas.....	78
1.2.3.	Entre medidas cognitivas y no cognitivas.....	78
1.3.	Efectos diferenciales para distintos grupos de alumnos.....	79
1.3.1.	Rendimiento previo.....	79
1.3.2.	Indicadores socioeconómicos.....	80
1.3.3.	Características personales del estudiante.....	81
1.3.4.	Diferencias en la Eficacia Docente.....	82
1.4.	Estabilidad de los efectos.....	82
1.5.	Continuidad de los efectos.....	83
2.	Modelos Multinivel.....	83
2.1.	¿Por qué Modelos Multinivel?.....	83
2.2.	Notación.....	84
2.3.	Pasos a seguir en un diseño multinivel.....	86
3.	Conclusiones.....	87

1. Propiedades de los efectos escolares

La IEE surgió de la necesidad de demostrar que los centros tenían alguna influencia sobre el desarrollo de los alumnos, pero estimar la magnitud de los efectos escolares no es suficiente, por lo que se ha comenzado a estudiar las propiedades que subyacen a estos factores (Murillo, 2007):

1. Consistencia de los efectos, entre diferentes áreas o medidas de resultados;
2. Estabilidad de los efectos;
3. Efectos diferenciales, entre diversos grupos de alumnos; y
4. Continuidad de efectos.

Como señalan Teddlie, Reynolds y Sammons (2000), las características científicas de los efectos de los centros y las cuestiones metodológicas están entrelazadas en el ámbito de la investigación educativa, por lo que este capítulo se centra en conocer cada propiedad y análisis aplicados que subyacen a la estimación de las mismas, donde cobran especial interés los *Modelos Multinivel* (MM), por lo que el capítulo finaliza con la conceptualización y elaboración de MM.

1.1. Magnitud de los efectos

Se refiere a la capacidad de los centros escolares para influir en los resultados de sus alumnos, se operativiza como el porcentaje de variación en el rendimiento de los alumnos debido a las características de proceso del centro en el que están escolarizados.

Para calcular la magnitud de los efectos se estima mediante MM (Aitkin y Longford, 1986; Raudenbush y Bryk, 1986), debido a que los MM están particularmente bien adaptado a esta tarea, ya que permite la partición de la variación total de las estimaciones de los parámetros (intercepto y pendientes) y en componentes de error.

No obstante “*la estimación de la magnitud de los efectos escolares es un ámbito aún abierto a controversias*” (Murillo, 2007, p.35). Los resultados obtenidos varían de manera considerable de un estudio a otro. Así, Mortimore et al. (1988) estimaron que el

tamaño del efecto de centro es del 10%, incluso cuando se emplean medidas de rendimiento de carácter más práctico; Purkey y Smith (1983), por su parte, encontraron que la diferencia en el nivel de rendimiento se encuentran entre centros más y menos eficaces se sitúa en 2/3 de desviación estándar. Mandeville y Anderson (1987) encontraron que los estudiantes adquirirían prácticamente todo el conocimiento de Matemáticas en el período de escolarización a diferencia de lo que ocurre en lengua o en arte, las cuales estaban influidas por otros factores ajenos al ambiente escolar como aquéllos procedentes del entorno familiar.

En fechas posteriores, Teddlie, Reynolds y Sammons (2000) realizaron un gran esfuerzo por recoger y sistematizar los resultados de 26 investigaciones más relevantes sobre la magnitud de los efectos escolares. Concluyeron que la magnitud de los efectos escolares se encuentra en torno al 15%, con una amplia variación entre países y estudios.

Con población española Murillo (2004) encontró en centros de Primaria un efecto de 9,26% en matemáticas, de 3,7% en lengua y de 3,47% en ciencia sociales, de 3,37% en ciencias naturales y de 6,14% en rendimiento medio. En estudios transnacionales como “*Programa para la Evaluación Internacional del Alumnado*” (PISA) de la OCDE el 13,9% de la varianza total está explicada por factores de contexto y proceso del centro para España en PISA 2006 y en PISA 2009 con la misma población se sitúa en un 19,5% (IE, 2010a). Martín et al. (2008) en un estudio longitudinal y multinivel en población española de secundaria elevó el efecto a un 20%. Recientemente el informe de la Evaluación General de Diagnóstico en Educación Primaria (IE, 2010b) indicó que la variación global que se produce en los resultados de los alumnos entre unos centros y otros es de 15,8% en la competencia en comunicación lingüística, de 14,6% en la competencia matemática, de 19,4% en la competencia en el conocimiento y la interacción con el mundo físico y de 16,3% en la competencia social y ciudadana.

1.2. Consistencia de los efectos escolares

Uno de los factores que generan la diversidad de resultados en la estimación de los efectos escolares es la utilización de varias medidas de resultado. Podría ocurrir que en un determinado centro los alumnos que tengan buenos resultados en matemáticas pero no tan buenos en comprensión lectora, o que éstos tengan alta autoestima pero con

nulos conocimientos de inglés. Es lo que se llama, técnicamente, *consistencia de los efectos escolares*. La consistencia se refiere, por tanto, a la congruencia entre las medidas de los efectos escolares realizadas a partir de una elección de distintas medidas producto (Crone et al., 1994, 1994). Esta cuestión tiene un alto interés académico y práctico: si los efectos escolares no son consistentes, complica el cálculo de la magnitud del efecto y, con ello, la búsqueda de escuelas eficaces y la determinación de factores asociados (Murillo, 2007).

La mayoría de las investigaciones miden la consistencia de los efectos escolares entre distintos productos por medio de la estimación de la correlación entre la eficacia de un centro obtenida a partir de dos variables de producto. Concretamente se elaboran los residuales de cada variable de producto - la diferencia entre el logro esperado y el obtenido teniendo en cuenta características contextuales del alumnado y del centro. Posteriormente se calcula el índice de correlación de Pearson o el coeficiente de Kappa, dicho valores un indicador del grado de discrepancia entre índices (e.g. Mandeville y Anderson, 1987; Crone et al., 1994 y 1994).

En el estudio de la consistencia de los efectos escolares es posible comparar medidas de distinta naturaleza como aquellas referidas a rendimiento cognoscitivo (lengua, matemáticas etc.) o entre medidas no cognoscitivas (resultados sociales, afectivos etc.) o, incluso, correlacionando residuales de variables de ambos tipos.

1.2.1. *Entre medidas cognitivas*

En general se puede decir que la consistencia entre áreas académicas en Primaria es moderada y positiva. Por ejemplo Sammons, Nuttall y Cuttance (1993), haciendo un reanálisis multinivel del *Junior School Project* (Mortimore et al., 1988), encontraron una correlación entre lectura y matemáticas de 0,61. Mandeville (1988) en su trabajo obtuvo correlaciones entre 0,59 y 0,74 en lectura y matemáticas en diferentes cursos de Primaria. En España Murillo (2004) encuentra correlaciones moderadamente altas, 0,82 entre ciencias sociales y naturales y en torno a 0,71 entre matemáticas con ciencias naturales y sociales. En lengua y matemáticas, Murillo (2007) encontró una correlación moderada de 0,5.

Para Secundaria destaca el estudio de Thomas, Sammons y Mortimore (1997) donde analiza la consistencia de los efectos para siete áreas a lo largo del tiempo, encontrando una correlación baja y moderada, entre 0,24 y 0,72. En España Castejón (1994) obtuvo

similares resultados, de hecho la relación entre matemáticas e idioma extranjero fue de 0,63 y entre historia e idioma extranjero fue de 0,63.

Esa disparidad de resultados en los estudios comentados referidos a población española se deben no sólo a causas de que el alumnado está en distintas etapa educativa, sino también, a que los *outputs* de Murillo (2004) proceden de la aplicación de pruebas estandarizadas de rendimiento mientras que Castejón (1994) se basa en calificaciones escolares otorgadas por docentes (Murillo, 2007).

1.2.2. *Entre medidas no cognitivas*

La evidencia sobre la consistencia de los efectos escolares entre medidas de rendimiento socio-afectivo es más escasa. Mortimore y cols. (1988) encontraron una relación de 0,33 entre actitudes hacia lengua y matemáticas; Fitz-Gibbon (1991) halló que la consistencia entre la actitud hacia las matemáticas y la actitud hacia la escuela era de 0,48. Murillo (2007) encontró una correlación baja pero significativa (en torno a 0,32) entre convivencia social y satisfacción con el centro.

1.2.3. *Entre medidas cognitivas y no cognitivas*

La discrepancia mayor se encuentra cuando se trata de medir la relación entre variables cognitivas y no cognitivas. Reynolds (1976) y Rutter et al. (1979) encontraron una alta correlación entre eficacia escolar y eficacia social (asistencia y delincuencia) y además concluyeron que en su conjunto los centros con una alta asistencia y buen comportamiento tienden a tener mejores resultados en los exámenes. Así mismo se encontró una relación fuerte entre el ámbito cognoscitivo y afectivo cuando son áreas académicas que requieren cierta *competencia afectiva* como música, arte etc. (Kochan et al., 1996). Sin embargo Gray et al. (1983), Mortimore et al. (1988), Fitz-Gibbon (1991), Knuver y Brandsma (1993) y Murillo (2004 y 2007) concluyen que los efectos escolares medidos como resultados académicos son independientes de los estimados en variables afectivas y sociales.

En general existe una falta de estudios en población española sobre el tema que se está abordando en general y particularmente en Educación Secundaria. Sobre todo se detecta una cierta ausencia de estudios aborden la magnitud de los efectos desde una aproximación multinivel. No existe una evidencia concluyente de la consistencia de tales efectos en pruebas estandarizadas de rendimiento y, menos aún, en pruebas que

evalúen competencias afectivas o conductas escolares (Peña-Suárez, Fernández-Alonso y Muñiz, 2009).

1.3. Efectos diferenciales para distintos grupos de alumnos

Reynolds et al. (1996) describieron dos tipos de efectos diferenciales en los centros: (1) debido a estudiantes con distintas características (2) debido a diferencias de unidades académicas dentro del centro. En particular, la Eficacia Diferencial se refiere a la “*variación de los efectos escolares para distintos grupos de alumnos*” (Murillo, 2007, p.41). Esto es, se trata de saber si la acción del centro educativo es similar para distintos tipos de estudiantes. Como por ejemplo, ¿los centros benefician más a aquellos alumnos con un rendimiento previo superior? o ¿son los centros igualmente eficaces para alumnos de bajo o alto nivel socio-económico?

La estrategia para estimar los efectos diferenciales se centra en comparar las diferentes pendientes de las rectas de regresión - una para cada centro - procedentes de los análisis de modelos multnível (Jeson y Gray, 1991).

“*Los Efectos diferenciales están muy relacionados con la característica de equidad que debe obtener un centro eficaz*” (Murillo, 2005 pp. 29-30). Se produce un efecto de equidad cuando el centro es capaz de compensar diferencias de entrada que inciden en el rendimiento diferencial del alumnado (Levine y Lezotte, 1990). Las características de entrada, que se han encontrado en la literatura de forma unánime, capaces de incidir en el logro académico son el rendimiento previo y las características socioeconómicas. No obstante de acuerdo con el último manual de eficacia escolar (Townsend, 2007) hace referencia a la importancia de prestar atención a otras características diferenciales de los estudiantes y del docente y sus prácticas. La labor del docente es fundamental para lograr un efecto de equidad. Los siguientes apartados hacen referencia al impacto de cada variable de contexto o de alumnado.

1.3.1. Rendimiento previo

En Primaria destaca el estudio de Rutter et al. (1979) donde concluye que los centros tienen un comportamiento similar para tres grupos creados en función de su rendimiento previo. En caso contrario, se encuentra el trabajo de Smith y Tomilson (1989) y Murillo (2007). Ambos estudios recogen diferencias en lengua como en

matemáticas, de tal forma que los alumnos con un rendimiento previo superior obtuvieron significativamente mejores incrementos en sus resultados.

En Secundaria, Nuttall et al. (1989) encontraron evidencias de una eficacia diferencial, ya que algunos centros parecen disminuir la distancia entre los alumnos con bajo y alto rendimiento previo. Willms (1992), Jesson y Gray (1991) y Sammons, Nuttall y Cuttance (1993) aportan conclusiones similares, parece que los centros que fueron eficaces son para todos los grupos.

Además en relación con el rendimiento se encuentra la condición de que un alumnado haya repetido un curso. Ser repetidor es un indicador de haber sufrido un fracaso en una escolaridad anterior, incluso, es un estigma que marca el futuro recorrido escolar (Gomolla y Radtke, 2002). En este sentido es un predictor de fracaso escolar y abandono de los estudios (Jimerson, Anderson y Whipple, 2002) y conlleva unas consecuencias emocionales y sociales relativas a la disminución de autoestima, menor atención y confianza hacia el rendimiento por parte del profesor (Bonvin et al., 2008). Por ejemplo en PISA 2000 con población española (Pajares-Box, 2005), la puntuación media de los repetidores en lectura desciende a 100 puntos menos con respecto a sus compañeros. Recientemente en PISA 2009 con la reevaluación de la competencia lectora en población española obtienen 518 puntos aquellos estudiantes que está en su curso modal, mientras que 435 puntos obtienen aquellos alumnos que han repetido un curso y 371 puntos aquellos estudiantes que han repetido dos cursos (IE, 2010a).

1.3.2. Indicadores socioeconómicos

El trabajo de Thomas, Sammons y Mortimore (1997) analizó la eficacia diferencial de los departamentos didácticos de centros de Educación Secundaria para diferentes grupos de alumnos en función de distintas variables, entre ellas su situación socio-económica. Los datos indican que los estudiantes más favorecidos socio-económicamente obtienen resultados especialmente buenos. Mientras que los centros y departamentos menos eficaces, a pesar de que todos los alumnos obtienen malas puntuaciones, los alumnos de clase más alta salen más perjudicados.

Es un indicador complejo de medir por su interacción con otras variables como la etnia o la condición de inmigrante y la titularidad de los centros. Nuttall et al. (1989) encuentra que el tipo de etnia no es una variable que se deba tener en cuenta de forma aislada sino que suele estar confundida con el índice socioeconómico de las familias.

Análisis posteriores de PISA 2000 (Pajares-Box, 2005) y de PISA 2003 (Marchesi y Martínez-Arias, 2006) con datos de España, encuentran que el mayor éxito en las pruebas por parte de centros privados y concertados en comparación con centros públicos, no es tal cuando se detrae el índice socioeconómico de las familias.

1.3.3. *Características personales del estudiante*

Inteligencia y personalidad explican gran parte de las diferencias individuales que influyen en el rendimiento: la inteligencia a través de habilidades específicas que facilitan el aprendizaje; personalidad a través de rasgos los cuales facilitan o dificultan el uso de las habilidades (Ackerman, 1996). Atendiendo a factores de personalidad es necesario considerar el modelo de “*Big Five*” (Costa y McCrae, 1997) compuesto por los siguientes factores de personalidad como son: Neuroticismo, Extraversión, Apertura a la Experiencia, Amabilidad y Responsabilidad. La mayoría de los factores han demostrado una incidencia en el rendimiento académico. Por ejemplo, la apertura a la experiencia es un factor de éxito académico (Shuerger y Kuma, 1987). Extraversión se relaciona con un bajo rendimiento asociado con conductas de distracción, sociabilidad e impulsividad (Demetriou et al., 2003) y también se ha encontrado una relación negativa entre desempeño y Neuroticismo asociado a ansiedad y estrés en situaciones de examen (Blickle, 1996; Chamorro-Premuzic y Furnham, 2003).

Dentro de las características personales también destacan los *estilos cognitivos*. *Estilo* refleja la forma en que las personas perciben, recuerdan, piensan y resuelven los problemas, fue introducido inicialmente por Allport (1937). A partir de ahí se han desarrollado teorías como la Teoría del Autogobierno Mental (Stengberg, 1988) que ha intentado encontrar factores relacionados con rendimiento del alumnado.

Dentro de las características personales también destaca la *metacognición* o el *conocimiento de las estrategias de aprendizaje habitual o más apropiadas para cada persona* (Flavell, 1979; Artelt, Schiefele y Schneider, 2001) es un factor considerado en la última edición de PISA 2009 que incide sobre la competencia lectora (IE, 2010a). La metacognición no sólo es útil para que los alumnos aprendan de manera eficiente y, además, permite a los docentes que puedan utilizarla en la creación de estrategias eficaces para distintos grupos de alumnos.

1.3.4. *Diferencias en la Eficacia Docente*

Actualmente se reconoce que responsables en política y, especialmente, las prácticas docentes son las que pueden promover la equidad del centro, dando igualdad de oportunidades de aprender a alumnado con distintas características sociales y personales distintas (Kyriakides, 2007). Por ejemplo, se ha demostrado que el alumnado menos capaz necesita un tipo de enseñanza más estructurada (Snow, y Lohman, 1984) y necesitan un feedback contingente (Kyriakides, 2007), estrategias también aplicables al el alumnado con nivel económico y cultural limitado (Brophy 1986 y Connell 1996).

1.4. *Estabilidad de los efectos*

Conocer si los efectos escolares son fiables de un año a otro, es lo que se conoce como estabilidad de los efectos en el tiempo. En la literatura se han encontrado resultados contradictorios para tratar este efecto, por una parte destacan los resultados de Mandeville y Anderson (1987) los cuales detectaron una cierta inestabilidad entre diferentes cohortes, variaban entre 0,36 y 0,66, y por otra, los resultados de Goldstein (1987) dan cuenta de correlaciones altas, entre 0,6 y 0,8.

El punto de inflexión en la medición de esta propiedad son los modelos longitudinales multinivel (Raudenbush, 1989, Willms y Raudenbush (1989) analizaron dos cohortes de estudiantes que completaron Educación Secundaria en Escocia, encontrando que la estabilidad de los efectos escolares (diferencia entre el rendimiento del alumno y el rendimiento del centro) mediante una regresión con el método de mínimos cuadrados ordinarios, se estimaba en 0,59. Aplicando una estrategia de análisis multinivel la estabilidad alcanza un 0,87 y si al diseño multinivel incluía efectos escolares ajustados a características de contexto y rendimiento previo, la estabilidad alcanzaba un valor de 0,96. Por tanto como señala Willms (1992) es necesario estimar el Valor Añadido y emplear estudios longitudinales multinivel para lograr una medida de estabilidad fiable y válida:

... para separar los efectos escolares atribuibles al centro por si mismo, descartando fluctuaciones aleatorias, es necesario describir los efectos contextuales, de proceso y resultado para sucesivas cohortes... mediante el empleo de estudios longitudinales los cuales nos dan información más completa y afinada de los efectos escolares.

Willms (1992, p.36)

1.5. *Continuidad de los efectos*

Se refiere a la pervivencia de los efectos escolares en los alumnos años después de que hayan abandonado el centro escolar. Goldstein y Sammons (1997) mediante el empleo de modelos multinivel reanalizaron los datos del *Junior School Project*. Entre sus resultados destaca que el efecto escolar en Primaria se mantenía en Secundaria e, incluso, se veía reflejado en las pruebas de certificación al final de etapa. A su vez el éxito en destrezas básicas (lectura y matemáticas) a los siete años de edad eran un buen predictor del logro académico a los dieciséis años.

2. Modelos Multinivel

2.1. *¿Por qué Modelos Multinivel?*

Los Modelos Multinivel (MM) son conocidos bajo una variedad de nombres: modelos jerárquicos lineales, modelos de coeficientes aleatorios, modelos de componentes de varianza y modelos lineales mixtos (Hox y Maas, 2005).

En los inicios de la investigación sobre Eficacia Escolar se utilizó el modelo de regresión múltiple con datos de alumnos sin agregar y mediante un *procedimiento de mínimos cuadrados* (e.g. Coleman et al, 1966 y Jencks et al, 1972). Esta estrategia olvida el contexto en el que la conducta se desarrolla, lo que se conoce como *falacia atomista* (Alker, 1969) o *paradoja de Simpson* (Kim, 2009). La atomización implica que la matriz de varianza – covarianza de los resultados no refleje la homogeneidad intragrupos (aulas, centros...), dicho de otra forma obvia la correlación intragrupos. Se hace necesario tener en cuenta la estructura jerárquica de los centros escolares lo que implica una dependencia de las observaciones de nivel micro (alumnos) dentro del nivel macro (e.g. aulas o centros). Esta dependencia se refiere a que los estudiantes del mismo grupo comparten el mismo ambiente, mismos profesores, normas y comunicación etc. (Snijders y Bosker, 1999).

Otra alternativa a la hora de trabajar con regresión múltiple es considerar el rendimiento medio del alumnado de los centros (e.g. Marks, Cox y Pomian-Szerdnicki, 1983, Prais, 1983 y Gray et al. 1983). Esta elección puede llevar a la “*falacia ecológica*” de transferir los resultados del grupo a los individuos (Robison, 1950) generando *sesgos de agregación* (Roberts y Burstein, 1980) que olvidan reconocer que la conducta es un acto esencialmente individual.

En ambas estrategias se da el fenómeno conocido como “*efectos de diseño*” (Kish, 1965; 1987) que, fundamentalmente, consiste en no reconocer las unidades de análisis adecuadamente. Los modelos multinivel resuelven el dilema entre agregación/disgregación al combinar simultáneamente varios niveles (Vallejo, Arnau y Bono, 2008). A su vez incluyen la posibilidad de establecer relaciones entre variables de distintos niveles, creando marcos teóricos más complejos y explicativos sobre las variables que influyen en los resultados (Heck y Scott, 2000). De ahí que los MM hayan suscitado gran interés en el ámbito organizacional en general y en particular a la IEE. La estructura escolar es jerárquica: alumnos están agrupados en aulas, las aulas están en centros escolares, los centros están incluidos en distritos etc., por tanto la estrategia multinivel hace recoger de forma precisa el medio natural de los datos (Heck y Scott, 2000).

Otra gran virtud de este tipo de análisis es la posibilidad de detectar diferentes fuentes de variación a través de los distintos niveles y a través de las distintas variables predictoras que se van incluyendo progresivamente en forma de efectos fijos o efectos aleatorios. Precisamente la detección de esas variables o factores ha sido una de las finalidades del Movimiento.

2.2. Notación

Los análisis de regresión multinivel tienen en cuenta que las unidades muestrales están anidadas dentro de unidades más amplias. En lugar de calcular una ecuación de regresión sobre el conjunto de datos entero, el análisis de regresión multinivel calcula una ecuación de regresión por cada unidad más amplia. Por tanto, un análisis de regresión multinivel calculará una ecuación de regresión por cada grupo.

Una vez definidos los niveles es necesario distinguir entre efectos fijos y efectos aleatorios de las variables predictoras, definiéndose matemáticamente, en los siguientes términos:

$$\text{Para un efecto fijo: } Y_{ij} = \alpha_j + \beta X_{ij} + \varepsilon_{ij} \quad [3.1]$$

$$\alpha_j = \gamma_{00} + U_{0j} \quad [3.2]$$

$$\text{Para un efecto aleatorio: } Y_{ij} = \alpha_j + \beta_j X_{ij} + \varepsilon_{ij} \quad [3.3]$$

$$\alpha_j = \gamma_{00} + U_{0j} \quad [3.4]$$

$$\beta_j = \gamma_{10} + U_{1j} \quad [3.5]$$

El subíndice i de las ecuaciones se refiere al alumno (nivel 1) y el subíndice j se refiere al grupo (nivel 2). En una ecuación, la presencia del subíndice j para un coeficiente de regresión significa que puede variar de un grupo a otro, que es lo que se considera efecto aleatorio.

El término ε_{ij} designa el residuo de la ecuación, es decir, la diferencia entre la puntuación observada Y_{ij} y la puntuación predicha \hat{Y}_{ij} . Este residuo se distribuye normalmente con una media de 0 y una varianza constante en el nivel 1, expresada como σ^2 .

Como muestran estas dos ecuaciones, el intercepto α_j se considera siempre como un efecto aleatorio. Considerar el intercepto como un parámetro fijo reduciría el modelo multinivel a la expresión de regresión lineal. El intercepto, α_j , puede dividirse en una parte fija γ_{00} -expresa el intercepto general y es igual a la media de los interceptos de grupos α_j y en una parte aleatoria.- U_{0j} - expresa la distancia desde el intercepto general al grupo. Se asume que esta distancia del coeficiente de grupo U_{0j} tiene una media de 0 y una varianza de τ_0^2 .

El coeficiente β de la primera ecuación no tiene subíndice j , lo que significa que el efecto de X no puede variar de un grupo a otro. Por lo tanto, las líneas de regresión son paralelas y el efecto de X se considera fijo. En cambio, el coeficiente β de la segunda ecuación tiene un subíndice j , lo que significa que puede variar de un grupo a otro. Las líneas de regresión ya no son paralelas y el efecto de X se considera de tipo aleatorio. Igual que antes, este coeficiente de regresión β_j puede dividirse en una parte fija y otra parte aleatoria. La parte fija γ_{10} se llama *coeficiente de regresión global* y corresponde a la media de los coeficientes de regresión β_j . La parte aleatoria U_{0j} es la distancia al grupo desde el coeficiente de regresión global. Tiene una media de 0 y una varianza llamada τ_1^2 .

Los efectos aleatorios y los efectos fijos pueden combinarse en un único análisis de regresión multinivel.

2.3. *Pasos a seguir en un diseño multinivel*

El modelado multinivel suele componerse de una serie de pasos comunes recogidos entre distintos autores (Heck y Scoot, 2000; Kim, 2009; Murillo, 2007).

Paso 1. Estimación del modelo nulo. El modelo nulo o modelo vacío es el punto de partida del proceso; contiene solo una variable respuesta y la constante o intercepto. De esta forma, el modelo posee efectos aleatorios en todos los niveles y no incluye variables explicativas en ninguno de ellos. El modelo nulo se establece como línea de base para la estimación de la varianza explicada a partir de la cual se van evaluando las aportaciones de modelos más elaborados.

Paso 2. Estimación del modelo con variables de ajuste. A partir del modelo nulo, se añaden las variables consideradas de ajuste en la parte fija y, tras comprobar que son estadísticamente significativas, se les incluye en la parte aleatoria y se analiza su comportamiento. En coherencia con la definición de Escuela Eficaz, estas variables estarán relacionadas con el nivel socioeconómico y cultural de los alumnos y de centro, con sus características personales y el rendimiento previo o el nivel sociocultural del centro.

Este modelo con las variables de ajuste es la base del planteamiento del Valor Añadido (VA), el VA se puede definir técnicamente, siguiendo a Goldstein (2001), como la diferencia entre el logro académico esperado y el obtenido teniendo en cuenta características su rendimiento inicial y la situación social, cultural y económica de las familias del alumnado. Incluso en este paso se proporcionan medios para separar los efectos derivados de las prácticas de las escuelas de variables de los estudiantes y contextuales (Linn 2004; Martínez-Arias, Gaviria-Soto y Castro-Morera, 2008; Saunders, 2000). Si se estiman los residuales de los alumnos a partir de este modelo, se tendrá el rendimiento de los alumnos al descontar esas variables de ajuste, lo que se denomina *logro académico*. Por último este modelo sirve de base para estimaciones posteriores, con los modelos ajustados. En esta fase se extraen resultados que dan respuesta a:

1. Estimación del efecto escolar, de aula y de sistema educativo para cada medida de rendimiento, a partir del cálculo de las correlaciones intraclase (ICC)¹. Dicho

¹ La correlación intra-clase es el porcentaje de varianza total del que es responsable el centro y queda expresada en la siguiente ecuación:

estadístico es una medida que indica el grado de dependencia de las observaciones de un grupo (Snijders y Bosker, 1999);

2. Análisis de la consistencia de los efectos escolares entre diferentes medidas de producto, mediante el cálculo de los residuales para cada variable y la estimación de su correlación mediante el índice de correlación de Pearson; y

Estudio de la eficacia diferencial de los efectos escolares para cada variable de rendimiento y cada variable de ajuste, a partir del análisis de la significación de la aportación de las variables de ajuste en la parte aleatoria.

Paso 3. Aportación de cada una de las variables proceso. Se elabora un modelo para cada variable de proceso o conjunto de variables incluyendo además las variables de contexto. De esta forma, se identifican los factores de aula, escolares y del sistema educativo asociados con el logro académico. El procedimiento para su estimación es análogo a la fase anterior. En primer lugar, se introduce cada una de las variables en la parte fija del modelo y se analiza si su aportación es significativa; si es así, se analiza si su aportación es significativa en la parte aleatoria.

Paso 4. Estimación del modelo final. Es una elaboración de un modelo capaz de incorporar todas las variables seleccionadas en el modelo teórico y que resultaron significativas. De esta forma, se obtiene un modelo multinivel para cada una de las variables producto. Esta fase y la anterior se repiten por cada variable producto.

Paso 5. Verificación de los supuestos del modelo: los errores presentan una distribución normal y son homocedásticos (media cero y varianza constante)

3. Conclusiones

Este capítulo muestra nuevas líneas de investigación dentro de la IEE como las propiedades científicas de los efectos escolares. Partiendo de la idea que la *escuela sí importa*, es necesario además conocer si sus efectos son consistentes en distintas áreas académicas y afectivas; si los efectos de centro se mantienen a través de distintas cohortes de estudiantes y si continúan después del período de escolarización; y averiguar si los factores escolares benefician a todos los estudiantes, independientemente de sus características personales. Para estudiar todos estos aspectos

$$\rho = \frac{\sigma_{entre_centros}^2}{\sigma_{entre_centros}^2 + \sigma_{dentro_del_centro}^2} = \frac{\tau_0^2}{\tau_0^2 + \sigma^2}$$

es necesario aplicar los avances en metodología multinivel, para ser capaces de realizar estimaciones de los efectos y factores escolares con rigor metodológico.

CAPÍTULO 4

EFICACIA ESCOLAR EN ESPAÑA

Capítulo 4: Eficacia escolar en España

1.	Eficacia escolar en España	91
1.1.	Estudios empíricos	91
1.1.1.	La eficacia en la educación escolar	92
1.1.2.	Procesos funcionales y eficacia de la escuela. Un modelo causal 92	
1.1.3.	Modelo de evaluación externa de los centros escolares de EGB 93	
1.1.4.	Rendimiento en el medio rural	93
1.1.5.	Factores de eficacia escolar de los centros públicos de EGB.....	94
1.1.6.	Modelo explicativo sobre calidad y gestión de recursos en los centros escolares	94
1.1.7.	Estudio sobre la gestión de los recursos educativos	95
1.1.8.	Calidad de la educación y eficacia de los centros: estudio sobre la gestión de los recursos educativos	95
1.1.9.	Calidad, eficacia y clima en centros educativos	97
1.1.10.	Mejora de Eficacia Escolar: Capacidad de cambio y adaptación de los centros	97
1.1.11.	Mejora de la Eficacia Escolar en Iberoamérica	98
1.1.12.	Variables que predicen el rendimiento académico en ESO: un análisis longitudinal multinivel.....	100
1.2.	Influencia en Políticas educativas	100
1.2.1.	Estudios transnacionales.....	101
1.2.2.	Estudios nacionales	102
1.2.3.	Definición de la Evaluación de Diagnóstico	103
1.2.4.	Mecanismos de la Evaluación de Diagnóstico	104
1.2.5.	Resultados de la evaluación general de diagnóstico	104
2.	Conclusiones	105
3.	Índice de tablas	106
4.	Índice de figuras	106

1. Eficacia escolar en España

De acuerdo con Pedró (2010) política, investigación e innovación son piezas que funcionan como un “*círculo virtuoso*” con una influencia recíproca en áreas como la salud. Sin embargo en la educación dicha relación es confusa, de hecho la influencia de la comunidad de investigadores que ejerce en la comunidad de usuarios potenciales no está tan clara. Teniendo en cuenta informes de la OCDE (2004a y 2007a) sobre la gestión del conocimiento, la I+D en el campo de la educación y la investigación política ponen de relieve las dificultades que surgen cuando intentan conjugar estos tres elementos.

El caso de la línea teórica que se está desarrollando, la IEE, ha tenido y tiene una importante repercusión en la educación en España a través de distintos estudios empíricos y su incidencia en el diseño de políticas educativas. Los siguientes epígrafes recogen dicho impacto tanto en la vertiente académica como política, advirtiendo la necesidad de establecer una mayor sintonía entre el ámbito académico y político en educación, tal y como se señaló al inicio.

1.1. Estudios empíricos

Los trabajos que se van a presentar están enmarcados en una tradicional línea de IEE, que tratan de conocer los factores escolares y de aula que influyen en el rendimiento académico de los alumnos. Su población de estudio es el alumnado que cursa educación obligatoria en ambos niveles. Desde el punto de vista metodológico se caracterizan por, prácticamente, una ausencia de MM. La mayoría son tesis doctorales y dos son proyectos internacionales, como se puede ver en la tabla 11. Aunque recientemente son destacables el estudio de Martín et al. (2008), ya que esta autora aplica un modelo multinivel y longitudinal y el trabajo de Murillo (2011) que perfila un modelo de Mejora de Eficacia escolar como posteriormente se verá.

Tabla 11: *Estudios de eficacia escolar en España*. Fuente: Murillo (2005 p. 113).

Autor/a y Año	Título	Tipo
Millán (1978)	<i>“La eficacia en la educación escolar”</i>	Tesis (Univ. de Valencia)
Fuentes (1986)	<i>“Procesos funcionales y eficacia de la escuela. Un modelo causal”</i>	Tesis (Univ. Complutense)
Rodríguez-Díez (1990)	<i>“Modelo para la evaluación externa de los centros escolares de EGB”</i>	Tesis (Univ. de Oviedo)

García-Durán (1991)	<i>“Calidad y eficacia en los centros educativos: análisis de las aportaciones de un modelo explicativo”</i>	Tesis (UNED)
Rodríguez-Gómez (1991)	<i>“Investigación evaluativa en torno a los factores de eficacia escolar de los centros públicos de EGB”</i>	Tesis (UNED)
Barrueco (1992)	<i>“Calidad y eficacia en los centros educativos: análisis de las aportaciones de un modelo explicativo”</i>	Tesis (Univ. de Salamanca)
Muñoz-Repiso et al. (1995)	<i>“Calidad de la educación y eficacia de la escuela. Estudio sobre la gestión de los recursos educativos”</i>	Internacional OCDE
Castejón et al. (1996)	<i>“La eficacia de centros de enseñanza secundaria: un modelo de identificación y funcionamiento”</i>	Financiada por CICy T
González-Galán (2000)	<i>Calidad, eficacia y clima en centros educativos. Modelos de evaluación y relaciones causales”</i>	Tesis (Univ. Complutense)
Muñoz-Repiso y Murillo (2003)	<i>“Capacidad para el cambio y adaptación de las escuelas en el caso de la Mejora de la Eficacia Escolar”</i>	Internacional (UE)

1.1.1. La eficacia en la educación escolar

Millán (1978) en su tesis doctoral hace una revisión narrativa de la situación de la Eficacia Escolar en aquellos momentos, concluyendo que en el mundo se estaba abordando esa temática desde cinco perspectivas diferentes: input-output, psicológica, organizativa, pedagógica y escuelas alternativas.

1.1.2. Procesos funcionales y eficacia de la escuela. Un modelo causal

Fuentes (1986) presentaba como objetivo localizar factores de Eficacia Escolar mediante el empleo de modelos causales. Entre sus conclusiones destaca el hecho de que la eficacia de los centros escolares es resultado del “*funcionamiento global*” de los mismos, pero sobre todo de la conjunción de tres influencias: las derivadas del origen social del alumnado; las ligadas a la actuación del profesorado en el aula; y procedentes del propio centro escolar mediante su funcionamiento y el clima general que lo caracteriza.

En relación con este último aspecto, encontró que las escuelas con mayor rendimiento se caracterizaban por una dedicación de todos (dirección y profesorado) a las tareas y responsabilidades colectivas: planificación y programación de la actividad docente, trabajo en equipo y participación del profesorado, liderazgo creativo y ambiente de disciplina que conjuga autoridad y libertad. En definitiva los centros eficaces son aquellos que tienen una *cultura de centro* facilitadora del rendimiento individual y de la implicación de todos en una tarea común.

Como crítica destaca que no mide el rendimiento de los alumnos de una forma directa, sino que lo hace a través de la opinión de los docentes o mediante las calificaciones en el graduado escolar.

1.1.3. Modelo de evaluación externa de los centros escolares de EGB

Rodríguez-Díez (1990) se centró en elaborar una serie de instrumentos para la evaluación de la eficacia de los centros docentes. La investigación formula y valida un modelo causal entre cuyos resultados se encuentran los siguientes:

1. Los factores con mayor influencia en la eficacia de los centros son la dotación del centro y el funcionamiento del Equipo Directivo y del Consejo Escolar;
2. El grado de satisfacción de los alumnos con la institución es menor en los estatus socioeconómico más elevado;
3. La estabilidad docente en el centro influye positivamente en el funcionamiento del Equipo Directivo, del Claustro y del aula, y negativamente en el nivel de satisfacción del alumnado; y
4. El clima del centro influye positivamente en el trabajo de los docentes en el aula y, a través de éste, en el rendimiento de los alumnos

1.1.4. Rendimiento en el medio rural

El trabajo de García-Durán (1991) se encuadró en un enfoque de casos prototípicos, bajo el que analiza las variables escolares que discriminan entre centros rurales de EGB de alto y bajo rendimiento, adaptando variables de IEE. Ordenó los centros rurales de la provincia de Cádiz de menor a mayor puntuación en el criterio (rendimiento) seleccionando siete de ellos y controlando el origen social de los alumnos. Entre los resultados del trabajo se destacan las características de los centros eficaces: revisión del plan de centro y consejo escolar, funcionamiento del claustro, orientación escolar, producción de recursos didácticos, evaluación contrastada, cooperación con las familias, los servicios de apoyo externos y altas expectativas sobre el centro.

1.1.5. Factores de eficacia escolar de los centros públicos de EGB

Rodríguez-Gómez (1991) analizó la eficacia y relación existente entre las variables personales, contextuales y funcionales de los centros públicos de EGB. La primera conclusión del trabajo fue la existencia de efectos escolares significativos, es decir, que las escuelas diferían en nivel de eficacia independientemente del origen social de sus alumnos. La segunda conclusión es la existencia de una relación entre las características de los centros y su nivel de eficacia, destacando las siguientes características de las escuelas eficaces:

1. Profesorado satisfecho de sus relaciones en el centro, expresa altas expectativas, trabaja en equipo introduciendo innovaciones y adaptando la enseñanza a las características de los alumnos;
2. Contexto escolar definido por la participación de padres, profesores y alumnos configurando un clima escolar positivo; y
3. Funcionamiento adecuado de Claustro y Consejo Escolar, mediante el trabajo en equipo y liderazgo educativo

1.1.6. Modelo explicativo sobre calidad y gestión de recursos en los centros escolares

Barrueco (1992) desarrolló un modelo causal de tres productos educativos: calificaciones escolares, formación general del alumno y niveles de satisfacción de los colectivos con mayor implicación en la escuela (padres, profesores y alumnos) para estudiar la relación entre calidad y eficacia, encontró cinco factores de eficacia:

1. Naturaleza de la comunidad educativa;
2. Autocontrol participativo de la misma;
3. Gestión técnica de la escuela;
4. Cualificación y preparación docente del profesorado; y

5. Planificación didáctica

1.1.7. Estudio sobre la gestión de los recursos educativos

Castejón (1994 y 1996) analizó factores procesuales del centro educativo que contribuyen a las diferencias de rendimiento entre centros y midió la consistencia entre áreas evaluadas. La población de estudio fue Educación Secundaria y la técnica de análisis la regresión múltiple. Entre los resultados más destacables se consideran:

1. Factores de proceso de tipo instruccional y regulativos tienen una influencia sobre el rendimiento tras controlar variables de contexto;
2. Los años de experiencia del docente inciden de forma significativa y en sentido negativo sobre el rendimiento medio del alumnado; y
3. En cuanto a factores de centro destaca el grado de satisfacción medio de profesores; percepción que tienen éstos del director; y apoyo que reciben de la dirección como elementos que ejercen una influencia en los resultados de los centros docentes.

Como críticas destaca la ausencia de MM y la ausencia de trabajar con medidas producto estandarizadas, aunque es considerado como uno de los mejores trabajos realizados (Murillo, 2005).

1.1.8. Calidad de la educación y eficacia de los centros: estudio sobre la gestión de los recursos educativos

Este trabajo corresponde a un estudio internacional titulado “*the effectiveness of schooling and of educational resource management*” (Muñoz-Repiso et al, 1995; Murillo, 1996). Los objetivos se concretaron en conocer la respuesta de los centros docentes a las reformas tendentes a mejorar la gestión de los recursos educativos; evaluar el clima escolar y la satisfacción de la comunidad educativa. El estudio recogió 542 variables que configuraban un modelo CIPP. Los instrumentos aplicados fueron entrevistas y cuestionarios a directivos, profesores, familias, alumnos e inspectores.

Un primer resultado de la investigación encontró que la relación entre las diferentes variables era completamente diferente entre los centros de Primaria y

Secundaria, con lo que se verificó la imposibilidad de elaborar un modelo de funcionamiento único.

En el modelo, figura 16, se observa que el clima de centro educativo es el factor clave y determinante del resto de los factores en Primaria. Representa el mejor predictor de las tres variables consideradas como producto y condiciona dos factores clave de proceso: la valoración del trabajo en equipo y la valoración de la metodología utilizada en el centro. Otro elemento destacable es la interrelación entre las tres variables de producto, mostrando la relación negativa entre la satisfacción general hacia el centro y el porcentaje de repetidores.

En Secundaria no es posible detectar una única dimensión que represente el grado de eficacia de un centro. Así se encuentran variables de rendimiento relacionadas fundamentalmente con el contexto del centro y la satisfacción de los miembros de la comunidad. Satisfacción que está asociada a la calidad de las relaciones entre los miembros del centro, la valoración del equipo directivo y la calidad del trabajo en equipo.

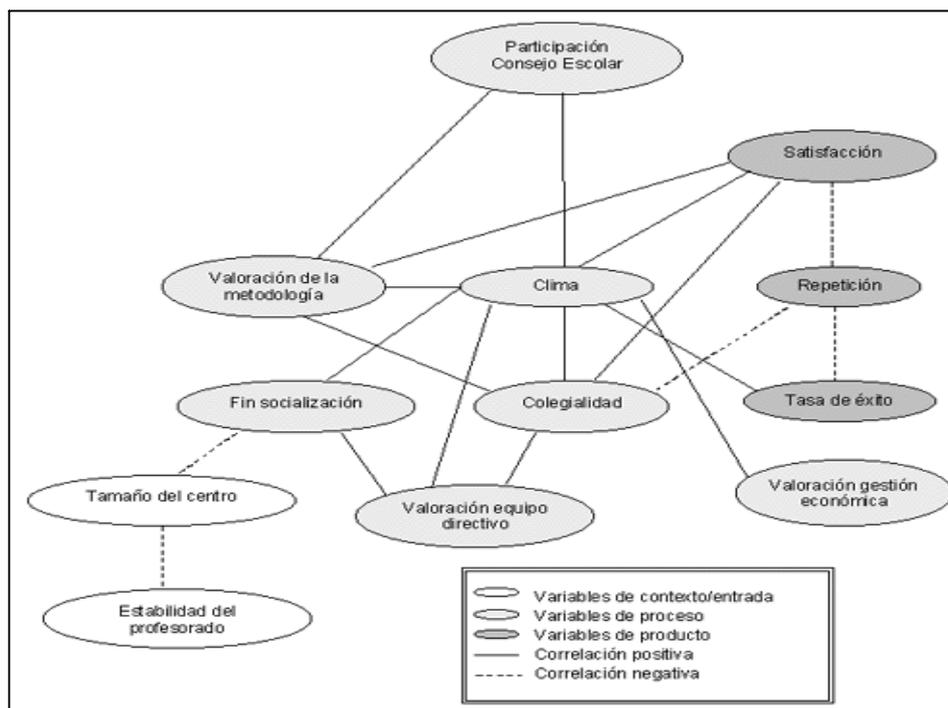


Figura 16: *Modelo de eficacia escolar en Primaria CIDE (1995)*. Fuente: Muñoz-Repiso et al. (1995 p.118)

1.1.9. Calidad, eficacia y clima en centros educativos

González-Galán (2000) validó un modelo causal de evaluación de la calidad centrado en el clima de trabajo y otras variables relacionadas con la eficacia (cultura, dirección, etc.) cuyos productos son la eficacia percibida por profesores, directivos y satisfacción de la plantilla. La novedad es la consideración de la dimensión de clima como un producto. La medida de las variables se realizó a través de escalas Likert aplicadas a 1.024 profesores y directivos agrupados en 78 centros de todas las etapas educativas de enseñanza obligatoria. Entre las conclusiones destacan:

1. El clima de trabajo y sus dimensiones explicativas se conforman como los factores esenciales en los modelos de evaluación;
2. La eficacia percibida por los profesores y directivos se explican a partir de la cultura institucional y en menor medida a partir del clima de trabajo; y
3. La eficacia presenta una relación sobre la satisfacción

1.1.10. Mejora de Eficacia Escolar: Capacidad de cambio y adaptación de los centros

El *Instituto Nacional de Calidad y Evaluación* (CIDE) participó en un proyecto junto en siete países, coordinados por Creeemers, para analizar las diferentes tradiciones teóricas que confluyen en la Mejora de la Eficacia Escolar (Muñoz-Repiso y Murillo, 2003) y analizar las experiencias de programas de Mejora de la Eficacia Escolar. La finalidad era crear un marco teórico comprensivo capaz de aportar una explicación teórica sobre factores que favorecen o impiden estos programas. El modelo surgido es el descrito en el capítulo 2 bajo el proyecto *Capacity from Change and Adaption in the case of Effective School Improvement (ESI)*. En el estudio español, en particular, se encontraron siete factores facilitadores y comunes entre los diversos sistemas educativos:

1. Cultura escolar en donde está implicada toda la Comunidad educativa y profesorado bien coordinado;
2. Papel de la dirección capaz de movilizar y lograr un compromiso de todos los docentes hacia la mejora;

3. Organización escolar que implique cambios de sistema;
4. Planificación del proceso de mejora orientado en objetivos;
5. Procesos de enseñanza y aprendizaje centrado en procesos de aula;
6. Estímulo o presión externa capaz de poner en marcha el proceso de mejora; y
7. Experiencias previas relativas a innovación y actitud favorable

A su vez destacaron una serie de factores que no incidieron en los resultados de España: la implicación de las familias y alumnado ni el desarrollo profesional del profesorado ni la influencia significativa de agentes externos de cambio.

1.1.11. Mejora de la Eficacia Escolar en Iberoamérica

Más recientemente cabe señalar a Murillo (2011) donde se elabora un modelo de mejora de la eficacia escolar para Iberoamérica y para España. Dicho trabajo, en parte, es bibliométrico y conjuga tres estudios de diferentes aproximaciones metodológicas. La primera aproximación fue un estudio comparado de la legislación de los diversos sistemas educativos en tres factores: presión del contexto y recursos del contexto para la mejora y metas del sistema educativo. El segundo estudio consta de una revisión de la literatura sobre eficacia escolar de cada país. La tercera aproximación se basa en una validación por expertos de dos factores de análisis asociados a la eficacia escolar: cultura y procesos de mejora. El instrumento de medida para esta última técnica fue una entrevista común para todos los países y colectivos. Para España el panel de expertos estaba formado por cuatro investigadores; dos inspectores y 4 directores.

Entre los resultados destaca el *modelo iberoamericano de la eficacia escolar*, figura 17, que es muy similar al propuesto al proyecto ESI, visto en el capítulo dos (figura 15). Sólo se añade la importancia del contexto socioeconómico, como viene siendo común en estudios centrados en estas poblaciones (e.g. Murillo, 2007). La composición de cada uno de los factores se especifica en la tabla 12.

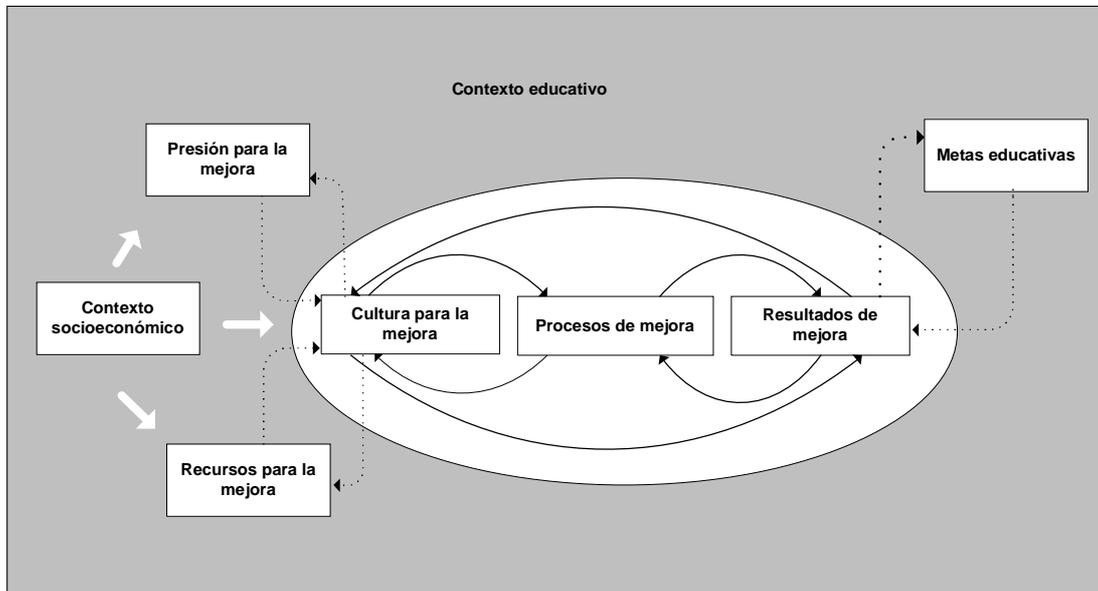


Figura 17: *Modelo iberoamericano de mejora de la eficacia escolar.* Fuente: Murillo (2011 p. 54)

Tabla 12: *Conceptos que componen los factores del modelo.* Fuente: Murillo (2011)

Factores de contexto	1. Niveles socioeconómicos del país, de la región y de la comunidad (incluido de la familia).
Presión para la mejora	1. Programas institucionales de cambio 2. Presión de incentivos para el cambio desde la Administración 3. Resultados de evaluaciones externas 4. Presión para el cambio desde inspección u otros agentes externos 5. Cambios educativos y sociales
Recursos para la mejora	1. Autonomía de los centros educativos 2. Recursos económicos y materiales 3. Apoyo de los agentes externos de cambio 4. Condiciones de trabajo favorables de las escuelas y de los docentes 5. Apoyo de la comunidad 6. Existencia de redes de escuelas y de docentes
Objetivos educativos	1. Objetivos educativos formales desde el punto de vista del rendimiento de los estudiantes
Cultura para la mejora	1. Presión interna para la mejora 2. Autonomía de las escuelas 3. Visión compartida 4. Historia de la mejora 5. Propiedad de la mejora, compromiso y motivación 6. Dirección escolar 7. Organización flexible 8. Buena disposición para convertirse en una organización de aprendizaje 9. Estabilidad del profesorado 10. Tiempo y recursos empleados para la mejora

Procesos de mejora	<ol style="list-style-type: none"> 1. Diagnóstico de las necesidades de mejora 2. Establecimientos de áreas prioritarias de mejora 3. Elaboración de objetivos de mejora 4. Planificación de las actividades de mejora 5. Aplicación de los planes de mejora 6. Seguimiento y evaluación de los procesos de cambio 7. Institucionalización
Resultados de mejora	<ol style="list-style-type: none"> 1. Cambios en la calidad del centro educativo 2. Cambios en la calidad del profesorado 3. Cambios en la calidad del rendimiento de los estudiantes (conocimientos, habilidades y actitudes)

1.1.12. Variables que predicen el rendimiento académico en ESO: un análisis longitudinal multinivel

Martín, Martínez-Arias, Marchesi y Pérez (2008) identificaron variables personales y de centro asociadas con el rendimiento académico de 965 estudiantes y 27 centros de secundaria, además, analizaron su influencia en el progreso de los alumnos a largo de la etapa. Para ello diseñaron un modelo multinivel longitudinal conformado por un nivel 1 que recoge tres momentos de la evaluación (inicio, mitad y final de etapa); nivel 2 compuesto por variables de estudiantes y un nivel 3 configurado por variables de centro. Las puntuaciones criterios proceden de escalas estandarizadas que evaluaban contenidos de lengua castellana, matemáticas y ciencias sociales.

Los resultados confirmaron la importancia de variables en el nivel personal como el sexo, el nivel sociocultural, habilidades metacognitivas y estrategias de aprendizaje. En el nivel de centro resultaron significativos el clima escolar y las expectativas del profesorado hacia los estudiantes. Mientras que el tamaño de centro, condición de repetidor y liderazgo de equipo directivo fueron significativas en algunas áreas.

1.2. Influencia en Políticas educativas

Actualmente se están desarrollando evaluaciones de sistemas educativos promovidos por organismos políticos, con el fin de convertirse en instrumentos políticos capaces de reorientar políticas educativas, basados en diagnósticos rigurosos y datos fiables. Este apartado se centra en desarrollar explotaciones de evaluaciones transnacionales, referidas a población española y evaluaciones nacionales con clara influencia de la IEE.

1.2.1. Estudios transnacionales

Como indica Martínez-Arias (2006) la idea latente de los estudios internacionales de evaluación de rendimiento educativo es considerar el mundo como una especie de *laboratorio educativo global* donde se fabrican los discursos que legitiman las políticas y se crean nuevos valores y visiones de influencia en escenarios de poder político y social.

Entre las similitudes de los diversos programas destaca su uso común de tests de desempeño (*performance assessment*). Los tests de desempeño son “*procedimientos estandarizados de evaluación en los que se demanda a los sujetos que lleven a cabo tareas o procesos en los que se demuestre su capacidad para aplicar conocimientos y destrezas a acciones en situaciones simuladas o de la vida real*” (Martínez-Arias, 2010 p.85).

Entre las diferencias de las distintas evaluaciones destaca una divergencia en sus objetivos a alcanzar (Linn, 2002). Los estudios de la *International Association for the Evaluation of Educational Achievement* (*Asociación Internacional para la Evaluación del Rendimiento Educativo* [IEA]), pretenden estimar el grado de consecución de objetivos y contenidos instructivos claramente señalizados. Por ejemplo, TIMSS (*Trends in International Mathematics and Science Study*) y PIRLS (*Progress in International Reading Literacy Study*) compara los currículos de los países participantes con el fin de encontrar un común denominador que pueda ser evaluado. Mientras que la Organización para la *Cooperación y el Desarrollo Económico* (OCDE), presentan como finalidad evaluar la competencia ante situaciones cotidianas, lo que fundamenta la originalidad de sus marcos teóricos de contenidos en PISA (*Programme for International Student Assessment*).

En la actualidad ambos organismos internacionales han irrumpido con dos nuevos programas de evaluación aplicados a población española, cuyos resultados los publica el *Instituto de Evaluación* (IE). El primero de ellos es el *Estudio Internacional de Civismo y Ciudadanía* (ICCS) elaborado por la IEA (IE, 2010c), el cual analiza de forma comparada los valores, actitudes, percepciones y comportamientos cívicos y ciudadanos de los alumnos españoles, según las respuestas de los alumnos a una prueba de comprensión y conocimientos conceptuales relacionados con la educación cívica y ciudadana de 38 países (24 países europeos).

El segundo es el *Estudio Internacional sobre la Enseñanza y el aprendizaje* (TALIS). TALIS examina aspectos importantes de la enseñanza y el aprendizaje a través de encuestas a los profesores y a los directores de Educación Secundaria Obligatoria de los 24 países participantes en el estudio. Los aspectos que evalúa son la formación y desarrollo profesional de los profesores, las opiniones y actitudes del profesorado sobre la práctica de la enseñanza, la evaluación del profesorado y del centro y su repercusión en el trabajo docente, así como el liderazgo escolar (IE, 2010d). Una visión general de los estudios comentados se muestra en la figura 18, donde se ordenan atendiendo a la fecha de publicación de informes de los distintos estudios.

Desde los años 60:	
IEA:	1964 Primer estudio internacional, matemáticas (FIMS).
	1982 Segundo estudio internacional, matemáticas (SIMS).
	1991 Comprensión lectora (RLS).
	Serie TIMSS: 1995, 1999, 2003, 2007 y 2010
	Serie PIRLS: 2001, 2006, 2011 mide competencia lectora
	1999, 2009 Estudio internacional de civismo y ciudadanía (ICCS)
	2008 Formación del profesorado de matemáticas para primaria y primer ciclo de ESO (TEDS-M)
	2013 Estudio internacional de competencia digital (ICILS)
Desde 1994:	
	OCDE: irrupción en la evaluación internacional.
	Series 2000, 2003, 2006, 2009 Programa internacional de evaluación de los estudiantes: Lectura, Matemáticas y Ciencias (PISA)
	2010 Estudio internacional sobre la enseñanza y el aprendizaje (TALIS)

Figura 18: *Orden cronológico de evaluaciones internacionales*. Fuente: elaboración propia.

1.2.2. *Estudios nacionales*

La Ley Orgánica para la Participación, la Evaluación y la Gestión de centros docentes (LOPEG), de 1995, Ley Orgánica de Calidad de la Educación (LOCE), de 2002, y, en menor medida, la Ley Orgánica de Ordenación General de Sistema Educativo (LOGSE), de 1990, tienen claras influencias de este Movimiento. Sobre todo las evaluaciones surgidas al amparo de estos marcos legales, figura 19.

Todas ellas comparten características comunes como áreas de evaluación (lengua, matemáticas, lengua inglesa etc.) y establecen diferencias entre sexo, titularidad y contexto socioeconómico.

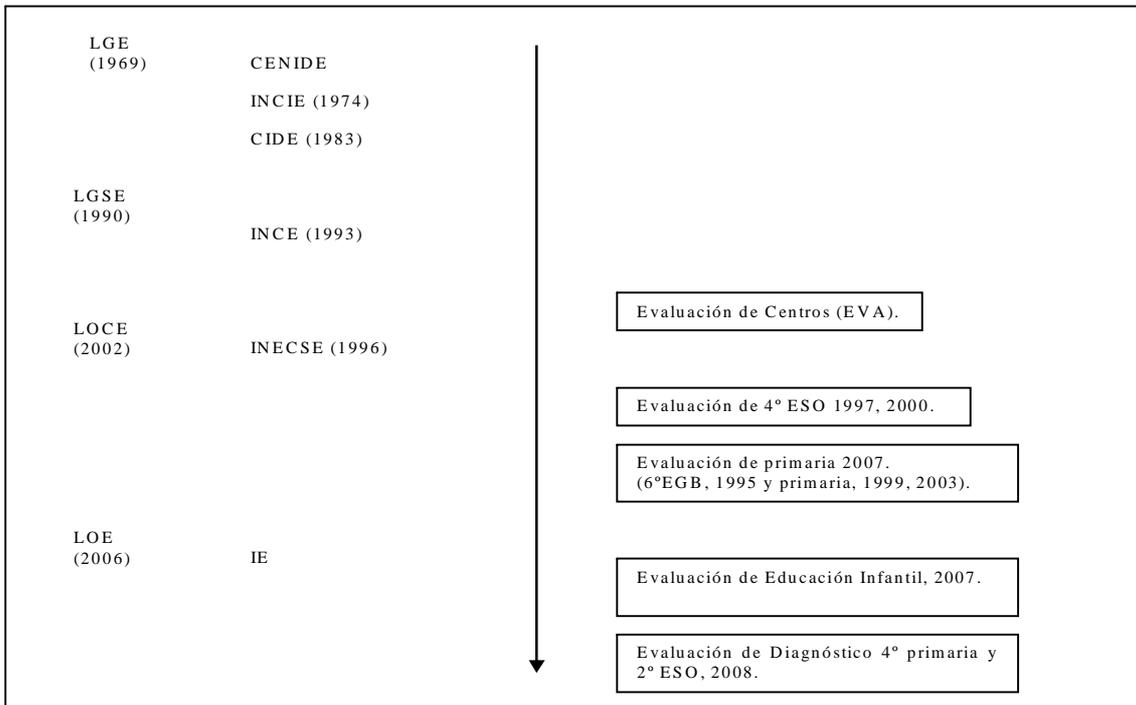


Figura 19: *Síntesis de evaluaciones del sistema educativo español*. Fuente: IE (2007)

Actualmente se está desarrollando la Evaluación de Diagnóstico, que es la propuesta de medición de vigente Ley Orgánica de Educación (LOE, 2/2006) y que se tratará a continuación.

1.2.3. *Definición de la Evaluación de Diagnóstico*

La evaluación diagnóstica se considera una medición del rendimiento del alumnado atendiendo a factores de contexto y proceso que inciden en el aprendizaje y en la orientación de políticas educativas. En consecuencia como establece la LOE en el capítulo I de su Título Preliminar la Evaluación de Diagnóstico es un seguimiento de dos objetivos educativos:

1. *La calidad de la educación* para todo el alumnado, independientemente de sus condiciones y circunstancias; y
2. *La equidad de la educación* que garantice la igualdad de oportunidades, la inclusión educativa y la no discriminación y actúe como elemento compensador de las desigualdades personales, culturales, económicas y sociales, con especial atención a las que se deriven de discapacidad

1.2.4. Mecanismos de la Evaluación de Diagnóstico

Este tipo de medición se centra en la evaluación de competencias del alumnado, realizado a partir de pruebas objetivas y cuestionarios estandarizados aplicados de manera anual en todo el ámbito territorial a través de:

1. *Las evaluaciones generales de diagnóstico*, permiten obtener datos representativos, tanto del alumnado y de los centros de las Comunidades Autónomas, como del conjunto del Estado. Estas deben surgir de la colaboración entre el Instituto de Evaluación y los organismos correspondientes de las Comunidades Autónomas, versan también sobre las competencias básicas del currículo de Educación Primaria (4º de Educación Primaria) y a la Secundaria Obligatoria (2º de Educación Secundaria Obligatoria); y
2. *Las evaluaciones de diagnóstico* al finalizar el segundo ciclo de Educación Primaria (4º) y Educación Secundaria Obligatoria (2º), realizan todos los centros de las competencias básicas alcanzadas por sus alumnos. Esta evaluación es competencia de las Administraciones Educativas, tiene carácter formativo y orientador para los centros e informativo para las familias y para el conjunto de la comunidad educativa

1.2.5. Resultados de la evaluación general de diagnóstico

Recientemente se han publicado los resultados de este tipo de evaluación en Educación Primaria (IE, 2010a). Los resultados indican que la variación global que se produce en los resultados de los alumnos entre unos centros y otros es de 15,8% en la competencia en comunicación lingüística, de 14,6% en la competencia matemática, de 19,4% en la competencia en el conocimiento y la interacción con el mundo físico y de 16,3% en la competencia social y ciudadana.

Factores de alumnado asociados al rendimiento destacan las diferencias significativas entre sexos en comunicación lingüística (14 puntos a favor de las chicas); en competencia social y ciudadana a favor de las chicas (22 puntos a favor de las chicas) y en competencia matemática (11 puntos a favor de los chicos). Otro factor de influencia es la influencia de la edad, por ejemplo se encuentran diferencias estadísticamente significativas a favor del alumnado que cursa cuarto de Educación Primaria en la edad prevista (nacidos en 1999) en comparación con aquellos alumnos cuya edad es superior (nacidos antes de 1999). Incluso las diferencias se hacen mayores

según aumenta el Índice Social, Económico y Cultural (ISEC). El rendimiento también está afectado por el lugar de nacimiento incluso cuando se controla el ISEC, los resultados indican que los estudiantes nacidos en España puntúan significativamente más alto en todas las competencias con la excepción de competencia social y ciudadana, donde no hay diferencias. El informe también destaca que los alumnos que quieren llegar a realizar estudios universitarios tienen los promedios más altos en las cuatro competencias evaluadas, con diferencias de hasta 76 puntos con respecto de aquellos alumnos que solo quieren realizar estudios obligatorios. Se obtiene una relación similar cuando se pregunta a las familias sobre el futuro académico de sus hijos.

Además encuentran que el clima escolar es una de las variables que más se relacionan con la calidad educativa de un centro escolar. Los alumnos cuyas opiniones sobre el clima y las normas de clase se sitúan en valores altos son aquellos cuyas puntuaciones medias son superiores (en torno a 515 puntos en las cuatro competencias). Al mismo tiempo, los alumnos que valoran en menor medida normas, tareas asignadas y justicia en su aplicación y corrección por parte del tutor tienen, a su vez, puntuaciones medias más modestas, en torno al *Promedio España* (500 puntos) en las cuatro competencias evaluadas.

2. Conclusiones

Para describir la presencia de la IEE en España es necesario contar con estudios procedentes del ámbito académico y prestar atención a trabajos desarrollados al amparo de políticas educativas. Entre los primeros, de carácter más “*ortodoxo*”, destaca el estudio de factores de centro que inciden en el rendimiento académico como el clima escolar, la valoración positiva de la dirección, el trabajo en equipo y la adaptación a las características del alumnado del trabajo de los docentes, expectativas de los docentes y la participación e implicación de las familias. Mientras que los trabajos enmarcados en el desarrollo de políticas educativas, en general, destacan más por sus características técnicas como tamaños de muestra representativos, pruebas estandarizadas e implementación de análisis multinivel que por el número de factores escolares de proceso encontrados y las propiedades científicas de tales efectos.

3. Índice de tablas

Tabla 11: *Estudios de eficacia escolar en España*. Fuente: Murillo (2005). 91
 Tabla 12: *Conceptos que componen los factores del modelo*. Fuente: Murillo (2011) .. 99

4. Índice de figuras

Figura 16: *Modelo de eficacia escolar en Primaria CIDE (1995)*. Fuente: Muñoz-Repiso et al. (1995) 96
 Figura 17: *Modelo iberoamericano de mejora de la eficacia escolar*. Fuente: Murillo (2011 p.54)..... 99
 Figura 18: *Orden cronológico de evaluaciones internacionales*. Fuente: elaboración propia. 102
 Figura 19: *Síntesis de evaluaciones del sistema educativo español*. Fuente: IE (2007) 103

CAPÍTULO 5

EVALUACIÓN PISA

Capítulo 5: Evaluación PISA

1.	Programa Internacional de Evaluación del Alumnado.....	109
2.	Competencia científica.....	110
2.1.	Organización del área de evaluación.....	111
2.1.1.	Situaciones y contexto.....	112
2.1.2.	Capacidades científicas	113
2.1.3.	Conocimiento científico	113
2.1.4.	Conocimiento de la ciencia	113
2.1.5.	Conocimiento acerca de la ciencia.....	114
2.1.6.	Actitudes hacia las ciencias	115
3.	Competencia Matemática.....	116
3.1.	Organización del área de evaluación.....	117
3.2.	Situaciones y contexto.....	118
3.3.	Contenido matemático.....	118
3.4.	Procesos matemáticos.....	119
4.	Competencia Lectora.....	120
4.1.	Organización del área de evaluación.....	120
4.1.1.	Situaciones de la lectura.....	120
4.1.2.	La estructura del texto	121
4.1.3.	Características de los ejercicios	122
5.	Características metodológicas en PISA	123
5.1.	Diseño muestral.....	123
5.1.1.	El método BRR.....	124
5.2.	Instrumentos de medida.....	127
5.2.1.	Redacción de ítems y estudios pilotos	127
5.2.2.	Pruebas aptitudinales.....	128
5.2.3.	Pruebas actitudinales	131
5.2.4.	Pruebas de contexto	131
5.3.	Construcción de escalas.....	133
5.4.	Construcción de índices de contexto.....	135
6.	Resultados de España en PISA.....	135
7.	Planteamiento de los objetivos	136
7.1.	Objetivo general.....	139
7.2.	Objetivos específicos.....	139
7.3.	Hipótesis	140
8.	Índice de tablas	141
9.	Índice de figuras	142

1. Programa Internacional de Evaluación del Alumnado

La primera evaluación trienal PISA tuvo lugar en el año 2000, representa la primera incursión de la Organización para la Cooperación y el Desarrollo Económico (OCDE) en el ámbito de las evaluaciones internacionales, hasta esa fecha liderada por la *International Association for the Evaluation of Educational Achievement (IEA)*.

PISA 2006 es el tercer ciclo de una estrategia de recogida de datos que fue definida en 1997 por los países participantes. Las publicaciones *Measuring Student Knowledge and Skills – A New Framework for Assessment* (OCDE, 1999) [*Medición de los conocimientos y habilidades de los Estudiantes: Un nuevo marco de evaluación*] y *The PISA 2003 Assessment Framework – Mathematics, Reading, Science and Problem Solving Knowledge and Skills* (OCDE, 2003) [*Marco de la evaluación de PISA 2003: Conocimientos y habilidades en matemáticas, lectura, ciencias y solución de problemas*], constituye el marco conceptual que sirvió de fundamento a los dos primeros ciclos de PISA. Recientemente, a finales del 2010, se ha comenzado a publicar los resultados del cuarto ciclo que cierra la evaluación de la competencia lectora como competencia principal. Todas estas publicaciones se pueden consultar a través de la dirección web: *www.pisa.oecd.org*.

En PISA el objetivo de la evaluación son las *competencias* entendidas como la capacidad de utilizar conocimientos aprendidos a situaciones de la vida real que deberían poseer todos los jóvenes a los quince años, cuando finalizan o están a punto de finalizar la educación obligatoria. El dominio de tal competencia es definido por expertos internacionales y no por un currículo.

El enfoque que adopta PISA es diferente en varios aspectos:

1. Su *origen*: una iniciativa adoptada por los gobiernos, cuyos resultados pretenden dar satisfacción a sus propias necesidades en materia de orientación política;
2. Su *regularidad*: el compromiso de cubrir múltiples áreas de evaluación con actualizaciones trienales, haciendo posible que los países realicen un seguimiento regular de los progresos obtenidos en el cumplimiento de sus principales objetivos educativos;

3. El *grupo de edad cubierto*: la evaluación de unos jóvenes que se encuentran al final del período de escolarización obligatoria permite obtener un indicador muy útil del rendimiento de los sistemas educativos; y
4. Los *conocimientos* y las *habilidades evaluadas*: sus contenidos no se definen atendiendo al común denominador que representan los currículos escolares nacionales, sino que se basan en las habilidades esenciales para que los estudiantes se desenvuelvan con éxito en su vida futura. No significa que excluya los conocimientos y la comprensión basados en el currículo, pero los evalúa en función de la adquisición de unos conceptos y habilidades de carácter más amplio que posibilitan la aplicación de los contenidos adquiridos

En definitiva evaluar directamente los conocimientos y las habilidades que se poseen cuando ya está próxima la finalización del período de escolarización básica permite evaluar los logros en función de los objetivos subyacentes a los sistemas educativos (según los define la sociedad), y no en función de la docencia y aprendizaje de un corpus de conocimientos. Esta forma de enfocar el rendimiento educativo se hace necesaria si se desea fomentar que los centros de enseñanza y los sistemas educativos se centren en los retos que plantea la sociedad contemporánea (OCDE, 2006b).

La definición genérica de competencia se concreta como:

Los conocimientos, habilidades, capacidades y otros atributos pertenecientes a los individuos que influyen de forma significativa en el bienestar personal, social y económico.

(OCDE, 2006b p. 12).

Las descripciones teóricas de la competencia científica, matemática y lectora que siguen corresponden con la aplicación del programa en el 2006, ya que la explotación de datos de la parte empírica pertenece a la aplicación del citado año.

2. Competencia científica

Las teorías actuales sobre la educación en ciencias hacen especial hincapié en la asimilación del conocimiento científico (que incluye el conocimiento del enfoque científico de la investigación) y la valoración del aporte que hacen las ciencias a la sociedad. Para alcanzar estos objetivos es necesario no sólo comprender los principales conceptos y explicaciones científicas, sino también, ser capaz de reconocer sus virtudes

y sus limitaciones. Por tanto otro objetivo consiste en desarrollar una actitud crítica y un enfoque reflexivo ante la Ciencia (Millar y Osborne, 1998).

El término que mejor engloba los objetivos generales de la evaluación en ciencias de PISA 2006 es *Competencia Científica* (Bybee, 1997; Fensham, 2000; Gräber y Bolte, 1997; Mayer, 2002; Roberts, 2001; UNESCO, 1993). Esta competencia se define como:

Un conjunto de conocimientos científicos de un individuo y uso de ese conocimiento para identificar problemas, adquirir nuevos conocimientos, explicar fenómenos científicos y extraer conclusiones basadas en pruebas sobre cuestiones relacionadas con la ciencia. Asimismo, comporta la comprensión de los rasgos característicos de la ciencia, entendida como un método del conocimiento y de la investigación humana, la percepción del modo en que la ciencia y la tecnología conforman nuestro entorno material, intelectual y cultural, y la disposición a implicarse en asuntos relacionados con la ciencia y con las ideas de la ciencia como un ciudadano reflexivo.

(OCDE, 2006b p. 12).

En PISA 2006 la definición de competencia científica es mejorada o completada en relación a PISA 2000 y 2003 por varias razones: Añade las relaciones existentes entre ciencia y tecnología y adjunta una serie de preguntas de actitud dentro de la evaluación de los conocimientos. Llevar a cabo una investigación que permita determinar en qué medida las cuestiones que se plantean en el curso de la prueba de evaluación despiertan el interés de los alumnos, contribuye a fortalecer la evaluación de una serie de elementos relativos a la actitud y a la motivación que serán importantes en el futuro compromiso con la ciencia.

2.1. Organización del área de evaluación

La definición de *competencia científica* de PISA 2006 se caracteriza por cuatro aspectos interrelacionados, figura 20:

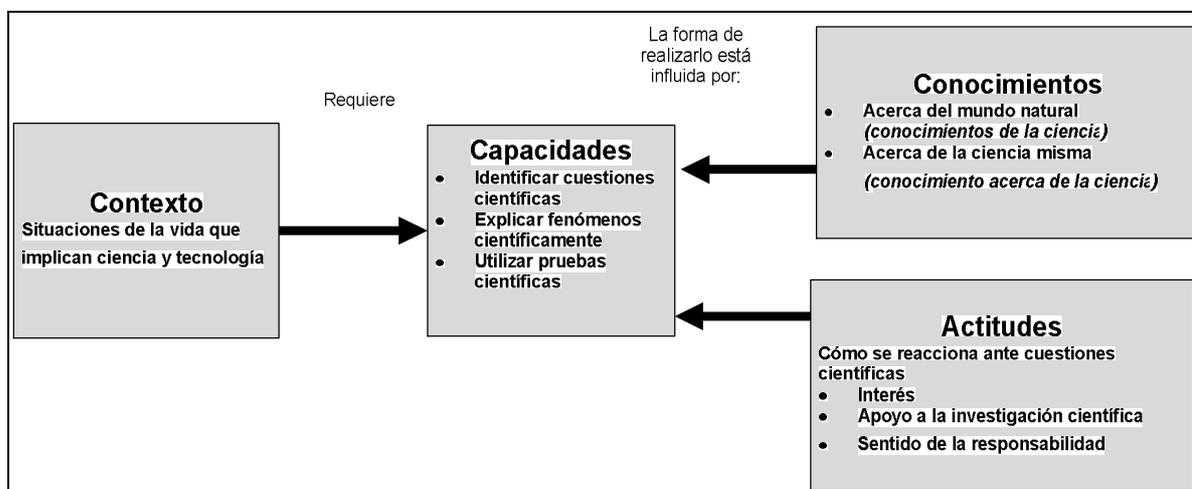


Figura 20: Elementos que componen la competencia científica. Fuente: OCDE (2006b p.27).

2.1.1. Situaciones y contexto

La situación es la parte del universo del estudiante donde se sitúan las tareas que se han de realizar. A este respecto, los ejercicios de evaluación no se limitan a las situaciones propias del entorno escolar, sino que se presentan enmarcados en una serie de situaciones relacionadas con el yo, la familia y los grupos de compañeros (*personal*), la comunidad (*social*) y la vida a escala mundial (*global*). La tabla 13 proporciona un listado de las aplicaciones de la ciencia dentro de unas situaciones personales, sociales y globales cuya principal función es servir de contextos para los ejercicios de evaluación.

Tabla 13: Contexto de la evaluación de ciencias. Fuente: OCDE (2006b p.28).

	Personal (yo, familia y compañeros)	Social (la comunidad)	Global (la vida en todo el mundo)
Salud	Conservación de la salud, accidentes, nutrición	Control de enfermedades, transmisión social, elección de alimentos, salud comunitaria	Epidemias, propagación de enfermedades infecciosas
Recursos naturales	Consumo personal de materiales y energía	Manutención de poblaciones humanas, calidad de vida, seguridad, producción y distribución de alimentos, abastecimiento energético	Renovables y no renovables, sistemas naturales, crecimiento demográfico, uso sostenible de las especies
Medio ambiente	Comportamientos respetuosos con el medio ambiente, uso y desecho de materiales	Distribución de la población, eliminación de residuos, impacto medioambiental, climas locales	Biodiversidad, sostenibilidad ecológica, control demográfico, generación y pérdida de suelos
Riesgos	Naturales y provocados por el hombre, decisiones sobre la vivienda	Cambios rápidos (terremotos, rigores climáticos) cambios lentos y progresivos (erosión costera, sedimentación),	Cambio climático, impacto de modernas técnicas bélicas

 evaluación de riesgos

Fronteras de la ciencia y la tecnología	Interés por las explicaciones científicas de los fenómenos naturales, aficiones de carácter científico, deporte y ocio, música y tecnología personal	Nuevos materiales, aparatos y procesos, manipulación genética, tecnología armamentística, transportes	Extinción de especies, explotación del espacio, origen y estructura del universo
--	--	---	--

2.1.2. Capacidades científicas

La evaluación en ciencias de PISA 2006 da prioridad a las capacidades que aparecen en la tabla 14. En todas estas capacidades se halla implícita la noción de conocimiento científico y actitudes científicas.

Tabla 14: *Capacidades científicas en PISA 2006*. Fuente: elaboración propia.

Capacidades científicas en PISA 2006	
Identificar cuestiones científicas	<ol style="list-style-type: none"> 1. Reconocer cuestiones susceptibles de ser investigadas científicamente 2. Identificar términos clave para la búsqueda de información científica 3. Reconocer rasgos clave de la investigación científica
Explicar fenómenos científicos	<ol style="list-style-type: none"> 1. Aplicar el conocimiento de la ciencia a una situación determinada 2. Describir o interpretar fenómenos científicamente y predecir cambios 3. Identificar descripciones, explicaciones y predicciones apropiadas
Utilizar pruebas científicas	<ol style="list-style-type: none"> 1. Interpretar pruebas científicas, elaborar y comunicar conclusiones 2. Identificar los supuestos, las pruebas y los razonamientos que subyacen a las conclusiones 3. Reflexionar sobre las implicaciones sociales de los avances científicos y tecnológicos

2.1.3. Conocimiento científico

El *conocimiento científico* hace referencia tanto al conocimiento sobre el mundo natural (conocimiento de la ciencia) como al *conocimiento acerca de la ciencia* en sí misma (la importancia del método científico).

2.1.4. Conocimiento de la ciencia

La evaluación en ciencias de PISA 2006 evaluó una muestra del conocimiento de la ciencia necesario para comprender el mundo natural y para dotar de sentido las experiencias que tienen lugar en distintos contextos tanto *personales, sociales y globales*. Los criterios a la hora de seleccionar los contenidos fueron los siguientes:

1. *Relevancia*: el conocimiento científico se distingue por el grado de utilidad que tiene para la vida de los individuos;

2. *Representatividad*: conceptos científicos importantes y, por tanto, de una utilidad duradera; y
3. *Adaptados*: ser acordes al nivel de desarrollo de los alumnos de 15 años

La tabla 15 muestra las categorías del *conocimiento de la ciencia* y algunos ejemplos de los contenidos seleccionados.

Tabla 15: *Conocimientos de la competencia científica en PISA 2006*. Fuente: elaboración propia.

Categorías del conocimiento de las ciencias en PISA 2006	
Sistemas físicos	1. Estructura de la materia (e.g. modelos de partículas , enlaces)
	2. Propiedades de la materia (e.g. cambios de estado, conductividad térmica y eléctrica)
	3. Cambios químicos de la materia (e.g. reacciones, transmisión de energía, ácidos/bases)
	4. Movimientos y fuerza (e.g. velocidad, fricción)
	5. Energía y su transformación (e.g. conservación, desperdicio, reacciones químicas)
	6. Interacciones de la energía y la materia (e.g. ondas de luz y de radio, ondas sonoras y sísmicas)
Sistemas vivos	1. Células (e.g. estructura y función, ADN, plantas y animales)
	2. Seres humanos (e.g. salud, nutrición, subsistemas [es decir, digestión, respiración, circulación, excreción, y sus relaciones], enfermedades y reproducción)
	3. Poblaciones (e.g. especies, evolución, biodiversidad, variación genética)
	4. Ecosistemas (e.g. cadenas tróficas, flujo de materia y energía)
	5. Biosfera (e.g. servicios del ecosistema, sostenibilidad)
Sistemas de la Tierra y el espacio	1. Estructuras de los sistemas de la Tierra (e.g. litosfera, atmósfera, hidrosfera)
	2. Energías en los sistemas terrestres (e.g. tectónica de placas, ciclos geoquímicos, fuerzas constructivas y destructivas)
	3. Historia de la Tierra (e.g. fósiles, orígenes y evolución)
	4. La Tierra en el espacio (e.g. gravedad, sistemas solares)
Sistemas tecnológicos	1. Papel de la tecnología de base científica (e.g. solución problemas, contribuye a satisfacer las necesidades y deseos de los seres humanos, diseña y desarrolla investigaciones)
	2. Relaciones entre ciencia y tecnología (e.g. las tecnologías que contribuyen al progreso científico)
	3. Conceptos (e.g. optimización, compensaciones, costes, riesgos, beneficios)
	4. Principios importantes (e.g. criterios, limitaciones, innovación, invención, solución de problemas)

2.1.5. *Conocimiento acerca de la ciencia*

Estos contenidos son similares a los propios del método de la ciencia. Entendiendo la *investigación científica* como la forma en que los científicos obtienen

datos y las *explicaciones científicas* como los objetivos de la ciencia (la forma en que los científicos usan los datos obtenidos). En la tabla 16 se muestran ambas categorías con ejemplos.

Tabla 16: *Categorías del conocimiento acerca de las ciencias en PISA 2006*. Fuente: elaboración propia.

Categorías del conocimiento <i>acerca de las ciencias en PISA 2006</i>	
Investigación científica	1. Origen (e.g. curiosidad, interrogantes científicos)
	2. Propósito (e.g. obtener pruebas que ayuden a dar respuesta a los interrogantes científicos, las ideas/modelos/teorías vigentes que orientan a la investigación)
	3. Experimentos (e.g. diversos interrogantes diversas investigaciones científicas, diseños de experimentos)
	4. Tipos de datos (e.g. cuantitativos[mediciones], cualitativos [observaciones])
	5. Medición (e.g. incertidumbre inherente, reproducibilidad, variación, exactitud/ precisión de los equipos y procedimientos)
	6. Características de los resultados (e.g. empíricos, provisionales, verificables, falsables, susceptibles a la autocorrección)
Explicaciones científicas	1. Tipos (e.g. hipótesis, teorías, modelos, leyes)
	2. Formación (e.g. representación de datos, papel del conocimiento existente y nuevas pruebas, creatividad e imaginación, lógica)
	3. Reglas (e.g. han de poseer consistencia lógica y estar basadas en pruebas, así como en el conocimiento histórico y actual)
	4. Resultados (e.g. producción de nuevos conocimientos , nuevos métodos, nuevas tecnologías; conducen a la vez a nuevos interrogantes e investigaciones)

2.1.6. *Actitudes hacia las ciencias*

La atención que presta PISA al estudio de las actitudes se basa en el convencimiento de que la competencia científica de una persona comporta toda una serie de actitudes, creencias, orientaciones motivadoras, criterios de autoeficacia, valores y, en último término, acciones. La inclusión de las actitudes y de las áreas seleccionadas para su medición se sustenta en la estructuración que hace Klopfer (1976) del dominio afectivo en el campo de la educación en Ciencias y en diversas publicaciones especializadas en la investigación del campo de las actitudes (Gardner, 1975, 1984; Gauld y Hukins, 1980; Blosser, 1984; LaForgia, 1988; Schibeci, 1984).

La evaluación en Ciencias de PISA 2006 valoró las actitudes de los alumnos en tres áreas: Interés por la ciencia, apoyo a la investigación científica y sentido de la responsabilidad sobre los recursos y los ambientes, tabla 17.

Tabla 17: *Áreas de evaluación de las actitudes en PISA 2006*. Fuente: elaboración propia.

Áreas de evaluación de actitudes de PISA 2006	
Interés por la ciencia	1. Mostrar curiosidad por la ciencia y los temas y comportamientos relacionados con la ciencia
	2. Demostrar disposición para adquirir conocimientos y habilidades científicas adicionales, utilizando diversos recursos y métodos
	3. Demostrar disposición para buscar información sobre materias científicas y poseer un interés continuado por la ciencia, incluyendo la posibilidad de considerar una opción profesional relacionada con las Ciencias
Apoyo a la investigación científica	1. Reconocer la importancia de tomar en consideración diversas perspectivas y argumentos científicos
	2. Apoyar la utilización de información factual y explicaciones racionales
	3. Expresar la necesidad de que los procesos que conducen a extraer conclusiones se realicen de una forma cuidadosa y lógica
Sentido de la responsabilidad sobre los recursos y los entornos	1. Dar muestras de que se posee un sentido de la responsabilidad personal sobre la conservación de un medio ambiente sostenible
	2. Demostrar que es consciente de la repercusión de las acciones individuales en el medio ambiente
	3. Demostrar disposición para tomar medidas a favor de la conservación de los recursos naturales.

3. Competencia Matemática

La competencia matemática se define como:

Una capacidad del individuo para identificar y entender la función que desempeñan las matemáticas en el mundo, emitir juicios fundados y utilizar y relacionarse con las matemáticas de forma que se puedan satisfacer las necesidades de la vida de los individuos como ciudadanos constructivos, comprometidos y reflexivos.

(OCDE, 2006b p. 74)

Una serie de comentarios explicativos adicionales contribuyen a clarificar la definición del área de evaluación:

1. El término competencia matemática supone una combinación de un volumen de conocimientos y habilidades matemáticas, así como la posesión de un cierto nivel lingüístico como un vocabulario rico y un conocimiento sustancial de reglas gramaticales, fonética ortografía etc.;
2. El término *mundo* hace referencia al marco natural, social y cultural en que vive el individuo: los conceptos, las estructuras y las nociones matemáticas;

3. La expresión “*utilizar y relacionarse con*” comprende tanto el uso de las matemáticas como la solución de problemas matemáticos, además de un grado de implicación personal más amplio que englobaría nociones como la comunicación, la sintonía, la valoración e incluso la apreciación y el disfrute de las matemáticas; y
4. La expresión “*la vida de los individuos*” incluye la vida privada de las personas, pero también su vida profesional, social (grupos de compañeros y familiares) y su vida como ciudadanos de una determinada comunidad

Por último PISA también reconoce el correlato que tiene en la competencia matemática cuestiones actitudinales poseer un cierto grado de seguridad en sí mismo, curiosidad, sentimientos de interés y relevancia o el deseo de realizar y comprender temas de contenido matemático.

3.1. Organización del área de evaluación

Para describir con mayor claridad el área objeto de la evaluación se distingue entre tres elementos: situaciones o contexto; el contenido matemático y los procesos matemáticos interrelacionados, representados en la figura 21.

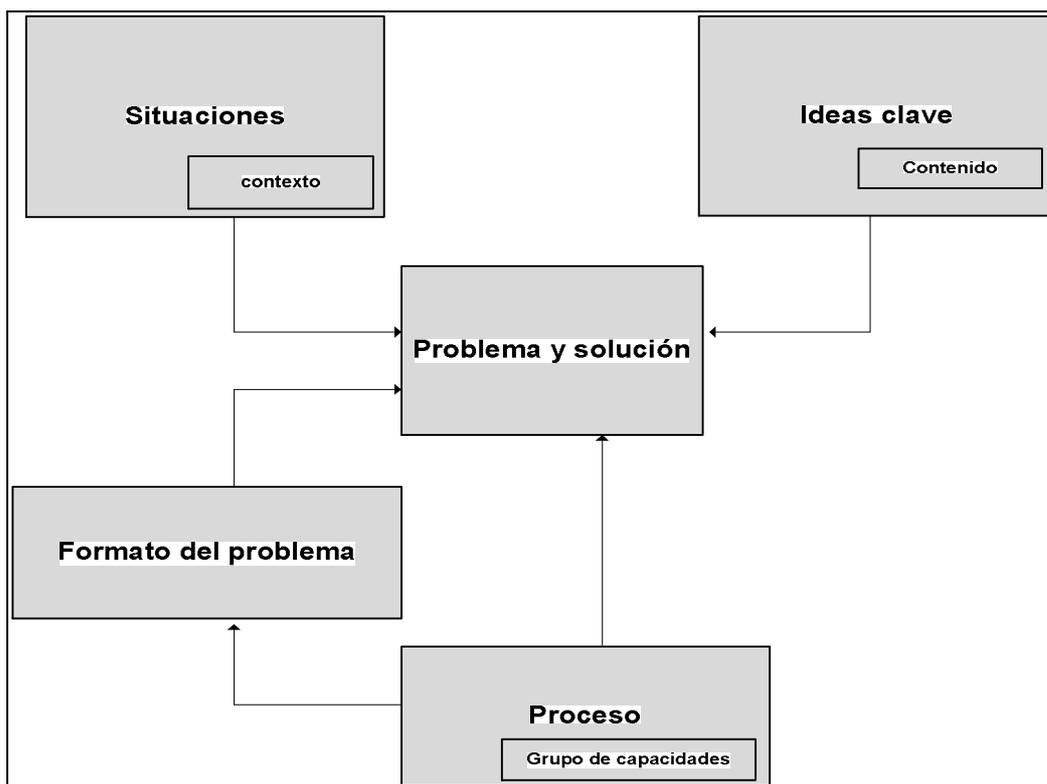


Figura 21: *Elementos del área de matemáticas*. Fuente: OCDE (2006b p. 82).

3.2. *Situaciones y contexto*

El nivel de competencia matemática de una persona se aprecia en la manera en que emplea sus conocimientos y habilidades matemáticas para resolver problemas. Dentro de la experiencia personal de cada individuo, los problemas (y su solución) pueden plantearse en una gran variedad de situaciones o contextos. La situación más cercana al alumno es su propia vida personal en PISA, luego vendría la vida escolar, laboral y el ámbito del ocio, y, posteriormente, la comunidad local y la sociedad. Las situaciones científicas, por el contrario, se encontrarían bastante más alejadas y fuera de su vida cotidiana. Los problemas que se planteen se sitúan en cuatro tipos de escenarios: *personal, educacional/profesional, pública y científica*. Mientras que el contexto de un ejercicio se concreta en el marco específico que se halla presente en una determinada situación, y a él pertenecen todos los elementos pormenorizados que se utilizan para formular el problema.

3.3. *Contenido matemático*

El contenido matemático evaluado se compone de cuatro categorías que reciben el nombre de *ideas clave*, son las siguientes: *espacio y forma, cambio y relaciones, cantidad e incertidumbre*. Estas ideas clave engloban, en toda su amplitud, los temas matemáticos que se espera que aprendan los estudiantes. En la figura 21, las ideas clave aparecen representadas en la parte superior derecha del diagrama. A partir de ellas, se extraen los contenidos que se utilizan para resolver un problema. Por esta razón el contenido está representado en el recuadro más pequeño que se encuentra inserto en el recuadro de las ideas clave.

Conceptos esenciales del espacio y forma, el cambio y las relaciones y la cantidad estarían presentes en cualquier descripción de las matemáticas y formarían parte del núcleo de cualquier currículo en todos los niveles educativos. Pero ser competente en matemáticas significa algo más, como saber abordar el campo de la incertidumbre desde una perspectiva matemática y científica. De ese modo, los elementos integrantes de la teoría de la probabilidad y la estadística dan lugar a la cuarta idea clave: la *incertidumbre*.

Los contenidos evaluados por PISA pueden no resultar demasiado familiares si se contemplan desde la perspectiva de la enseñanza de las Matemáticas y de las tendencias curriculares que suelen seguir los centros educativos.

3.4. Procesos matemáticos

Los procesos matemáticos engloban a las capacidades matemáticas para resolver problemas. Tres grupos de capacidades condensan los distintos procesos cognitivos necesarios para resolver los diversos tipos de problema.

El ámbito de los procesos aparece representado primero en el cuadrado grande, donde figuran las capacidades matemáticas generales, y luego en el cuadrado más pequeño, que representa los tres grupos de capacidades, figura 21. Las capacidades específicas que requieren la solución de un problema dependerán de la naturaleza del problema y se verán reflejadas en la solución que se dé al mismo. Esta interacción se representa mediante una flecha que va desde los grupos de capacidad al problema y su solución.

La flecha restante va desde los grupos de capacidad al formato del problema, indicando que las capacidades empleadas en resolver el problema están relacionadas con la forma que adopta el problema y con las exigencias concretas que plantea.

Las capacidades, de hecho, constituyen el núcleo mismo del concepto de competencia matemática. Solo los estudiantes que dispongan de ciertas capacidades se encontrarán en condiciones de resolver con éxito los problemas que se les planteen. Evaluar la competencia matemática supone, por tanto, determinar hasta qué punto poseen los alumnos una serie de capacidades matemáticas específicas que puedan aplicar de forma productiva a aquellas situaciones que llevan implícito un problema.

Para describir las actividades cognitivas que engloban estas capacidades, PISA ha optado por elaborar tres *grupos de capacidades*, como se representa en la figura 22.

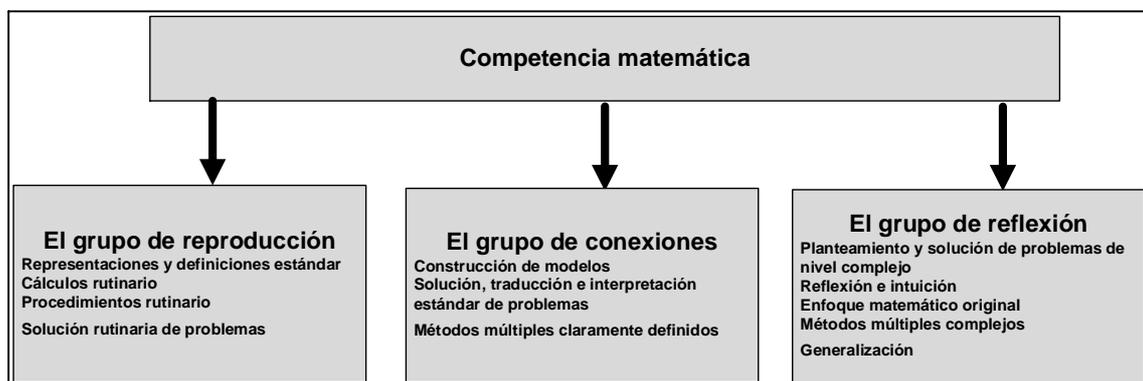


Figura 22: *Capacidades incluidas en la competencia de matemáticas*. Fuente: OCDE (2006b p.112).

4. Competencia Lectora

Las definiciones de la lectura y de la competencia lectora han evolucionado paralelamente a los cambios sociales, económicos y culturales. El concepto de aprendizaje, y en especial el aprendizaje a lo largo de toda la vida, han transformado las percepciones de dicha competencia. De hecho la definición que aporta el programa supera la idea tradicional de competencia lectora como proceso de descodificación y comprensión literal. Para PISA la lectura es un proceso dinámico que incluye muchos factores, entre los que se evalúa la situación de la lectura, la estructura del propio texto y las características de las preguntas que se suscitan sobre el texto.

La competencia lectora es la capacidad de comprender, utilizar y analizar textos escritos para alcanzar los objetivos del lector, desarrollar sus conocimientos y posibilidades y participar en la sociedad.

(OCDE, 2006 p. 48)

En PISA 2009 se ha desarrollado una nueva opción para los países participantes: la lectura de textos electrónicos (Electronic Reading Assessment [ERA]), que trata de evaluar la comprensión lectora de los alumnos en este formato de comunicación. Sólo 19 países han participado en esta pionera opción (del total de 65 en 2009), entre ellos España. El estudio permitirá comparar los resultados en la comprensión lectora sobre textos impresos y electrónicos en los mismos alumnos, que se publicó en junio del 2011 (IE, 2010).

La evaluación ERA abre una posibilidad y un reto para el desarrollo de las evaluaciones internacionales (y nacionales) en un futuro más o menos próximo., ya que en PISA 2015 está previsto desarrollar y aplicar todos los instrumentos mediante este sistema.

4.1. Organización del área de evaluación

4.1.1. Situaciones de la lectura

La situación se entiende como una categorización general de los textos según el uso pretendido por su autor, según la relación con terceros relacionados implícita o explícitamente con el texto y según su contenido general. Los textos de muestra en la evaluación fueron extraídos de una variedad de situaciones con el fin de maximizar la diversidad de contenidos en el estudio de la competencia lectora. Las cuatro variables de

situación están tomadas del *Marco Común Europeo de Referencia para las Lenguas* (MCER) pueden describirse de la siguiente manera:

1. *Lectura para fines públicos*. Es aquella que se practica para tomar parte en actividades sociales o comunitarias; por ejemplo, la lectura de impresos oficiales o de textos informativos sobre acontecimientos públicos. Por lo general, estas actividades comportan un contacto, más o menos anónimo, con otras personas.

2. *Lectura para fines privados (personal)*. Se refiere a la lectura con el fin de satisfacer sus propios intereses, ya sean de orden práctico o intelectual. También se califica así a la lectura encaminada a entablar o conservar relaciones personales con otras personas. En esta categoría se encuentran las cartas personales, así como los textos de ficción, biográficos o informativos que se leen por curiosidad, o como parte de actividades recreativas o de ocio.

3. *Lectura para fines profesionales (entorno laboral)*. Aunque no todos los jóvenes de 15 años tendrán que leer en sus entornos laborales, es importante calibrar la capacidad de una persona de esa edad para introducirse en el mundo del trabajo, pues en la mayoría de los países más del 50 % es probable que pasen a formar parte de la población laboral en un plazo de uno o dos años. Las tareas típicas de esta categoría suelen recogerse bajo la fórmula *leer para hacer* (Sticht, 1975; Stiggins, 1982), ya que están encaminadas al desempeño de una tarea inmediata.

4. *Lectura para fines educativos*. Este tipo de lectura suele formar parte de la adquisición de conocimientos dentro de una tarea educativa más amplia. Con frecuencia, los materiales escritos no los escoge el lector, sino el docente. Normalmente, su contenido ha sido explícitamente elaborado para cumplir una función formativa. Las tareas típicas de este tipo de lecturas se identifican con la función *leer para aprender* (Sticht, 1975; Stiggins, 1982).

4.1.2. *La estructura del texto*

En el núcleo de la evaluación PISA en lectura se sitúa la distinción entre textos continuos y textos discontinuos. Los *textos continuos* están compuestos normalmente por una serie de oraciones que, a su vez, se organizan en párrafos. Tales párrafos pueden hallarse insertos en otras estructuras mayores, como serían los apartados, los capítulos y los libros. Los textos continuos se clasifican primordialmente por su objetivo retórico,

esto es, por el tipo de texto, como la narración, exposición, descripción, argumentación, instrucción, documento o registro e hipertexto.

Mientras que *los textos discontinuos* o documentos se refieren a los cuadros y gráficos, tablas, diagramas, mapas, formularios, hojas informatizadas, convocatorias y anuncios, vales o bonos y certificados.

4.1.3. Características de los ejercicios

Para describir las *características de los ejercicios* de lectura PISA hace una distinción entre los procesos (aspectos) que describen la tarea exigida a los estudiantes; los tipos de ejercicio que especifican la manera exacta en la que los estudiantes han de demostrar su capacidad, y los criterios de corrección que especifican cómo han de evaluarse las respuestas de los estudiantes.

Con el fin de simular situaciones reales de lectura, la evaluación PISA mide los siguientes cinco procesos, figura 23, que deben realizarse para comprender plenamente un texto, ya sea continuo o discontinuo.

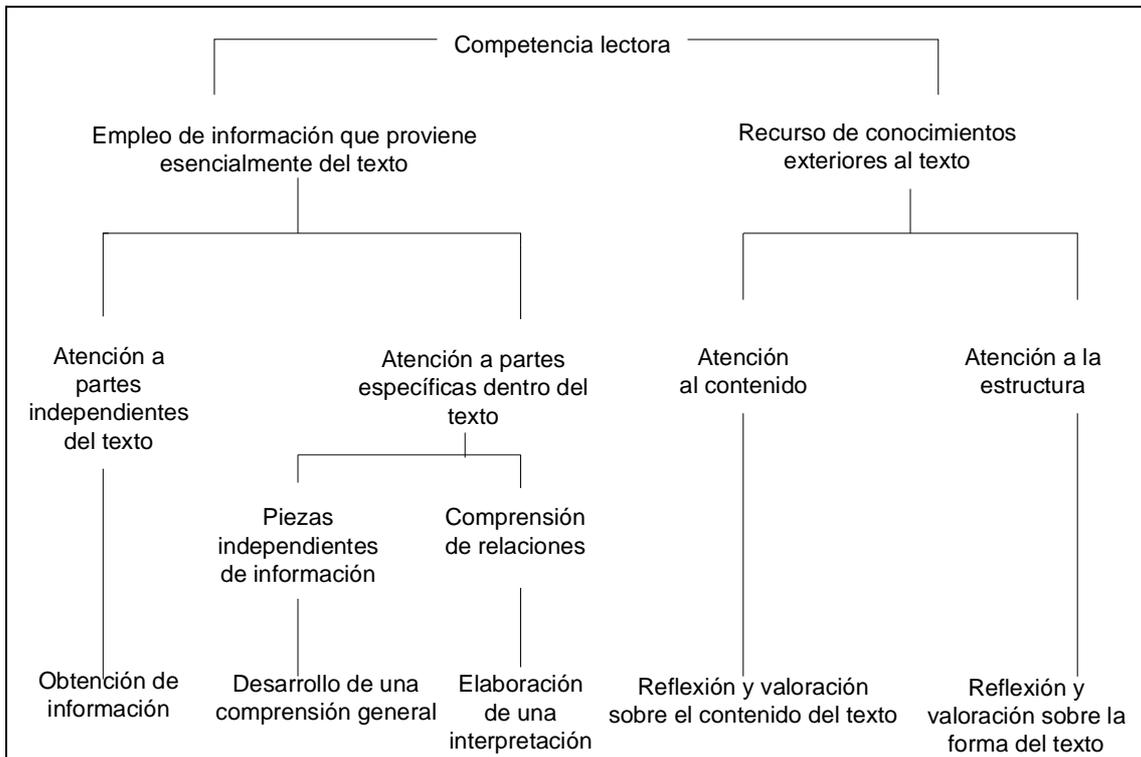


Figura 23: *Subescalas de procesos que componen la competencia lectora.* Fuente: OCDE (2006b p.52).

En cuanto a las características de los ejercicios todas las ediciones de programa han ido planteando cerca del 43 % de los ejercicios de respuesta construida-abierta que

exigían una valoración por parte del corrector o evaluador. Las tareas restantes consistían en ejercicios de respuesta construida-cerrada, que apenas requieren valoración por parte del evaluador, ejercicios de elección múltiple sencilla, en los que los alumnos tienen que elegir una entre varias respuestas alternativas, y finalmente ejercicios de elección múltiple compleja, en los que los estudiantes tienen que escoger más de una respuesta.

Finalmente en cuanto a la *corrección de los ejercicios* depende si su puntuación es dicotómica o politómica como en el caso del resto de competencias evaluadas y que se detallará en este capítulo en la parte de metodología aplicada, como a continuación se muestra.

5. Características metodológicas en PISA

5.1. Diseño muestral

El diseño aplicado en este tipo de encuestas educativas es un diseño estratificado por conglomerados en dos etapas (Martínez-Arias, 2006), a través del cual es posible relacionar variables de alumnado y de centro. En una primera fase se seleccionó una muestra de centros a partir de la lista completa de las organizaciones escolares que constituían la población de interés. En una segunda etapa se tomó una muestra aleatoria simple de alumnos dentro de los centros seleccionados, escogiéndose a 35 centros con población de 15 años. Si en algún centro seleccionado tenía escolarizado a menos de 35 estudiantes, se invitaba a participar a todos los alumnos de 15 años del centro.

Las variables de estratificación fueron explícitas e implícitas, pudiendo variar entre países. La diferencia entre ambos tipos radica en que las variables explícitas dividen a la muestra de centros, mientras que las variables implícitas ordenan los centros dentro de cada estrato explícito. La elección de las variables explícitas depende de su correlación con la prueba de rendimiento. En España, las explícitas fueron la Comunidad Autónoma, la titularidad del centro público/ privado, el tamaño del centro y la modalidad de enseñanza para el País Vasco. La variable de estratificación implícita fue el código postal (OCDE, 2009a).

En el nivel de centro se excluyen a organizaciones que atienden exclusivamente al alumnado con Necesidades Específicas de Apoyo Educativo. En el nivel de alumnado se excluyen a estudiantes con limitaciones en el idioma, con discapacidad motora,

intelectual o emocional. No obstante la ejecución de la prueba se desarrolló en un ambiente inclusivo, de tal forma que PISA desarrolla guías adicionales de aplicación de las pruebas para examinar al alumnado de este tipo.

Las reglas de exclusión de centros y de estudiantes están explicadas en el *Technical Report* (OCDE, 2009b). Se determinó que la tasa de exclusión global intra-centro debía ser inferior a 5% y la tasa de respuesta para los estudiantes en los centros seleccionados debía ser del 80%. En caso contrario se procedía al reemplazamiento. El remplazo fue el centro anterior o posterior con respecto al centro seleccionado inicialmente, un procedimiento que garantizó la equivalencia de las escuelas sustituidas, ya que el marco muestral de centros se organizó según el tamaño de éstos.

Sin embargo los métodos estándar para el análisis estadístico se han desarrollado a partir de muestreo aleatorio simple donde se supone la independencia de las observaciones. En PISA se incumple dicho supuesto. La consecuencia es un fallo en la estimación de la amplitud de los intervalos de confianza obtenidos en los análisis estadísticos (Kish, 1987). La solución pasa por la aplicación de procedimientos especiales, como es el uso *métodos de remuestreo* (Martínez-Arias, 2006), que consisten en obtener múltiples submuestras a partir de la muestra original. Rust y Rao (1996) destacan que el principio común de estos métodos es estimar la variabilidad entre las estimaciones resultantes para obtener el verdadero error muestral de la estimación inicial o de la muestra completa. No hacerlo así, llevaría a una sobreestimación sistemáticas de las varianzas muestrales. PISA sigue el método de replicación repetido balanceado (*Balanced Repeated Replication* [BRR]) con la modificación de Fay (1989) que permite obtener ochenta replicaciones y mejora, por tanto, la precisión del estimador de la varianza.

5.1.1. El método BRR

El método de replicación repetido equilibrado comienza de forma similar a otros métodos de remuestreo. Comienza emparejando sistemáticamente los centros que pertenecen en cada estrato en el orden en el que se seleccionaron, formando *pseudoestratos*. La peculiaridad del método BRR con respecto a otros reside en que dentro de cada pseudoestrato un centro se le atribuye un valor de 0 y al resto de centros se le doblan los pesos en cada una de las réplicas. Así se proporciona un gran número de replicaciones posibles y se genera un conjunto equilibrado de muestras replicadas

según las matrices Hadamard con objeto de evitar cálculos extensos. El número de replicaciones es el número más pequeño múltiplo de cuatro y mayor o igual al número de pseudoestratos. La tabla 18 representa 20 centros, divididos en 10 pseudoestratos, lo que implica 12 replicaciones.

Tabla 18: *Replicaciones del método BRR*. Fuente: OCDE (2006a p.59).

Pseudoestrato	Centro	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12
1	1	2	0	0	2	0	0	0	2	2	2	0	2
1	2	0	2	2	0	2	2	2	0	0	0	2	0
2	3	2	0	0	0	2	0	0	0	2	2	2	0
2	4	0	2	2	2	0	2	2	2	0	0	0	2
3	5	2	0	2	0	0	2	0	0	0	2	2	2
3	6	0	2	0	2	2	0	2	2	2	0	0	0
4	7	2	0	0	2	0	0	2	0	0	0	2	2
4	8	0	2	2	0	2	2	0	2	2	2	0	0
5	9	2	0	2	0	2	0	0	2	0	0	0	2
5	10	0	2	0	2	0	2	2	0	2	2	2	0
6	11	2	0	2	2	0	2	0	0	2	0	0	0
6	12	0	2	0	0	2	0	2	2	0	2	2	2
7	13	2	0	2	2	2	0	2	0	0	2	0	0
7	14	0	2	0	0	0	2	0	2	2	0	2	2
8	15	2	0	0	2	2	2	0	2	0	0	2	0
8	16	0	2	2	0	0	0	2	0	2	2	0	2
9	17	2	0	0	0	2	2	2	0	2	0	0	2
9	18	0	2	2	2	0	0	0	2	0	2	2	0
10	19	2	0	0	0	0	2	2	2	0	2	0	0
10	20	0	2	2	2	2	0	0	0	2	0	2	2

El estadístico se calcula en la muestra completa y posteriormente por cada replicación. Luego, las estimaciones de las replicaciones se comparan con la estimación de la muestra completa para calcular la varianza muestral, tal y como figura en la fórmula.

$$\sigma^2_{(\hat{\theta})} = \frac{1}{G} \sum_{i=1}^G (\hat{\theta}_{(i)} - \hat{\sigma})^2 \quad [5.1]$$

donde G es el número de réplicas; $\hat{\theta}_{(i)}$ representa el estimador estadístico para la muestra replicada y $\hat{\sigma}$ representa el estimador estadístico basado en la muestra completa.

Con este método cada muestra replicada sólo se compone de la mitad de las observaciones disponibles. Esta reducción de muestra podría ser problemática para

estimar un estadístico en una subpoblación *extrema*. Para superar esta desventaja, Fay (1989) sugirió multiplicar los pesos de los centros por un factor k de deflación entre 0 y 1, con un segundo factor de inflación igual a 2 menos k . Por ejemplo si el factor de deflación del peso, llamado k , es igual a 0,6, el factor de inflación del peso será igual a $2-k$, es decir $2-0,6 = 1,4$ (Judkins, 1990). El valor del factor en el estudio es de 0,5 (OCDE, 2009a).

Como ocurre con todos los métodos de replicación, el estadístico de interés se calcula a partir de la muestra entera, y después, de nuevo a partir de cada replicación. Las estimaciones replicadas se comparan entonces con la estimación de la muestra entera para obtener la varianza muestral, como a continuación se muestra en la ecuación 5.2 y sustituyendo los diferentes términos por los valores que toman en PISA, como se ve en la ecuación 5.3.

$$\sigma_{(\hat{\theta})}^2 = \frac{G-1}{G(1-k)^2} \sum_{i=1}^G (\hat{\theta}_{(i)} - \hat{\theta})^2 \quad [5.2]$$

$$\sigma_{(\hat{\theta})}^2 = \frac{G-1}{G(1-k)^2} \sum_{i=1}^G (\hat{\theta}_{(i)} - \hat{\theta})^2 = \frac{1}{80(1-0,5)^2} \sum_{i=1}^{80} (\hat{\theta}_{(i)} - \hat{\theta})^2 = \frac{1}{20} \sum_{i=1}^{80} (\hat{\theta}_{(i)} - \hat{\theta})^2 \quad [5.3]$$

donde k es el factor procedente del método de Fay.

Los cálculos de descriptivos, coeficientes de correlación, coeficientes de regresión, diferencias de medias, porcentajes, desviaciones típicas etc. se calcularon en base a unas macros que se implementaron en el SPSS. 14 para PISA 2006, las cuales incluyeron las 80 replicaciones y la inclusión de los valores plausibles que se verán más adelante.

Sin embargo en los modelos de regresión múltiple multinivel ofrecen la posibilidad de tener en cuenta el hecho de que los alumnos están anidados en clases y centros, y cada factor puede evaluarse por separado cuando se establece la medida de resultado. Sin embargo los programas específicos de modelado multinivel como MLWin o HLM consideran que la muestra de los centros es aleatoria, no tiene en cuenta la información complementaria que utiliza PISA con el fin de reducir varianza muestral. Como consecuencia, las varianzas muestrales estimadas con modelos multinivel siempre serán mayores que las muestras replicadas de Fay.

5.2. Instrumentos de medida

5.2.1. Redacción de ítems y estudios pilotos

La experiencia obtenida a través de las dos evaluaciones anteriores de PISA evidencia la importancia de un trabajo conjunto y cooperativo entre diversos órganos internacionales para llevar a cabo el programa de forma transcultural. Concretamente la evaluación del 2006 se nutre de cinco instituciones consolidadas como CITO (Países Bajos); ILS (Universidad de Oslo, Noruega); la empresa WESTAT y EPS (Estados Unidos); IPN (Universidad de Kiel, Alemania) y NIER (Japón) y coordinadas por ACER (*Australian Council for Educational Research*) (OCDE, 2009b).

Traducir y adaptar los tests a múltiples lenguas es una importante fuente de problemas en las evaluaciones internacionales (Martínez-Arias, 2006; van de Vijver, 2011). Incluso las traducciones pueden producir sesgos (Elosua, 2006) como por ejemplo que los ítems difieran en dificultad y demandas cognitivas, lo que llevaría a comparaciones no válidas (Hambleton, 2002; Hambleton, Merenda y Spielberger, 2005). De tal forma que se aplican procedimientos de doble traducción o *back translation* (nueva traducción a la lengua original). En este sentido PISA aplicó dos versiones fuente; inglés y francés, encontrando una buena aproximación (Adams y Wu, 2002). Además incluye notas de otros países sobre problemas de adaptación, desarrollo de guías detalladas de traducción y adecuación de los contenidos del tests, revisiones después del pilotaje, formación de equipos nacionales, selección y formación de un grupo de verificadores internacionales.

Con la finalidad de mejorar la validez de contenido y validez de constructo de los ítems, PISA desarrolló:

1. Aplicación de paneles de expertos entre los responsables del desarrollo de los instrumentos y posteriormente otros paneles de expertos internacionales;
2. Tests pilotos con muestras de estudiantes representativos de la población objetivo. Incluso presentando versiones alternativas del mismo ítem para identificar la mejor forma de presentación;
3. Desde la aplicación del 2003 se incluye una fase de análisis de los ítems mediante procedimientos de la *psicometría cognitiva*, para mejorar la validez de constructo de las tareas. Permite disponer de una especie de modelo del alumno

en la resolución de la tarea, a través del cual se indaga en las demandas cognitivas (Mislevy, Steinberg y Almond, 2003; Navas y Urdaneta-Durán, en prensa). Además se utilizaron técnicas afines como resolución en voz alta, entrevistas individuales y grupales; y

4. Otros aspectos son la inclusión de ítems de respuesta abierta y extendida, tendencia que se ve en otras evaluaciones internacionales como TIMSS para evaluar procesos superiores (Martin, Mullis, Gregory y Stemler, 2000; Martin, Mullis y Chrostowi, 2004). La inclusión de este tipo de ítems reduce la factibilidad de corrección, la fiabilidad de las puntuaciones y el tiempo requerido para la respuesta del test (Linn, 2004)

Las técnicas de análisis empleadas son cuantitativas tanto procedentes de la Teoría Clásica de los Tests (Muñiz, 2001); Teoría a la Respuesta de Ítems (Muñiz, 1997) y la Teoría de la Generalizabilidad (Brennan, 2002, Cronbach, Gleser Nandan y Rajaratnam, 1972; Martínez-Arias, 1995; OCDE, 2005a). Además de técnicas cualitativas como el escalamiento dual (Nishisato, 1980), también conocido como análisis de correspondencias (Gifi, 1990), cálculos de fiabilidad interjueces o codificadores.

5.2.2. Pruebas aptitudinales

Para que los tests de evaluación del rendimiento sean fiables, presenten una buena validez de contenido y se ajusten bien a las matrices de especificaciones es preciso incluir muchos reactivos (Martínez-Arias, 2006). A su vez es imposible y poco deseable evaluar a todos los alumnos con todos los ítems, entre otros motivos porque generarían fatiga y menor tasa de participación de los centros y consecuentemente sesgos en los resultados y aumento del error muestral. Por estas razones los estudiantes son asignados aleatoriamente a subconjuntos de ítems. A este diseño se conoce como muestreo matricial de ítems (*Matrix Sampling of Items*) (Beaton, 1997; Childs y Jaciw, 2003; Fernández- Alonso y Muñiz, 2011).

Concretamente la forma más común de construcción de formas de tests es el diseño de Bloques Incompletos Balanceados (*Balanced Incomplete Blocks* [BIB]). Frey, Hartig y Rupp (2009) señalaron las siguientes condiciones para este tipo de diseño:

1. Cada cluster debe aparecer una sola vez en cada cuadernillo;
2. Cada cluster debe estar presente en los cuadernillos con la misma frecuencia;

3. Cada cuadernillo debe tener idéntica longitud, lo que significa tener el mismo número de clusteres; y
4. Cada par de clusteres deben figurar juntos en igual frecuencia

Tabla 19: *Distribución de cuadernillos en PISA 2006*. Fuente: OCDE (2009b p.91)

Cuaderno	Cluster 1	Cluster 2	Cluster 3	Cluster 4
1	S1	S11	S10	S18
2	S2	S13	S12	S2
3	S3	S12	S11	S1
4	S4	S14	S13	S3
5	S5	S15	S14	S4
6	S6	S16	S15	S5
7	S7	S17	S16	S6
8	S8	S16A	S17	S7
9	S9	S17A	S16A	S8
10	S18	S10	S17A	S9
11	S1M	S11M	S10	S18
12	S3M	S13M	S12	S2
13	S2	S12	S11M	S1M
14	S4	S14	S13M	S3M
15	M1	M2	R2	R1

Como se especifica en la tabla 19, aparecen trece clusters o grupos, trece cuadernillos y cuatro posiciones posibles, apareciendo cada cluster cuatro veces en las cuatro posiciones posibles. Sin embargo siguiendo a los anteriores autores programas como PISA y estudios de la NAEP aplican un tipo de modelo de BIB llamados diseños de cuadrados Youden (*Youden Square Designs [YSD]*) (Youden, 1937 y 1940). YSD incorporan las cuatro condiciones anteriores más una condición adicional que implica que cada unidad debe aparecer en una posición una vez con igual frecuencia.

Los clusters de S1 A S7 evalúan ciencias; de M1 a M4 evalúan matemáticas y R1 y R2 evalúan lectura. Además de los 13 cuadernillos había uno específico para alumnado con Necesidades Educativas Específicas el cual contenía la mitad de ítems: 50% de ciencias; 25% de lectura y 25% de matemáticas. Las unidades de análisis se estructuran en 13 grupos con un tiempo estimado de 30 minutos de aplicación. Como anteriormente se ha dicho, hay siete grupos de ciencias, dos de lectura y cuatro de matemáticas, distribuidos en 13 cuadernillos, por lo que el efecto de posición queda controlado mediante este tipo de diseño. Cada uno de los cuadernillos contiene cuatro grupos y se asigna a cada alumno uno de estos cuadernillos, dado que la competencia

científica es la competencia central en la aplicación 2006, cada cuadernillo incluye un grupo de ciencias.

En todas las competencias se hace una distinción entre ítems de anclaje e ítems elaborados en *ad hoc* para la aplicación, los primeros sirven para comparar en una misma escala la cohorte del 2003 con la cohorte del 2006, permitiendo una calibración conjunta de los ítems, la comparabilidad de las puntuaciones y, en último término, la equiparación de puntuaciones procedentes de diversas formas de tests.

Todas las unidades de las tres competencias están compuestas de un estímulo (e.g. texto, tabla, gráfico, figura etc.) al que siguen una serie de tareas asociadas a ese estímulo común. De este modo, el estudiante dispone de más tiempo para asimilar un material que luego puede ser utilizado para evaluar diversos aspectos de su rendimiento.

Toda la evaluación se desarrolló mediante una prueba escrita, incluyendo distintos tipos de preguntas. Algunas de ellas requieren que los alumnos escojan respuestas simples que pueden ser cotejadas directamente con una única respuesta correcta, como en los ejercicios de elección múltiple o ejercicios de respuesta construida-cerrada. Otras preguntas, en cambio, poseen un carácter más creativo y requieren que los alumnos elaboren sus propias respuestas. Las tablas 20, 21 y 22 especifican el número de ítems según el formato y contenidos evaluados.

Tabla 20: *Clasificación de ítems por formato y capacidad en ciencias*. Fuente: OCDE (2009b p. 44).

Competencia Científica				
Formato de ítems	Identificación de temas científicos	Explicación de fenómenos científicos	Evidencia de usos científicos	Total
Elección múltiple	9	22	7	38 (35%)
Elección múltiple compleja	10	11	8	29 (27%)
Respuesta cerrada construida	0	4	1	5 (4,63%)
Respuesta abierta construida	5	16	15	36 (33%)
Total	24 (22%)	53 (49%)	31 (29%)	108

Tabla 21: *Clasificación de ítems por formato y capacidad en matemáticas*. Fuente: OCDE (2009b p.45).

Capacidades en Matemática				
Formato de ítems	Reproducción	Conexión	Reflexión	Total
Elección múltiple	5	3	4	12 (25%)
Elección múltiple compleja	0	7	2	9 (19%)
Respuesta cerrada construida	2	2	2	6 (13%)
Respuesta abierta construida	4	12	5	21 (44%)
Total	11 (23%)	24 (50%)	13 (27%)	48

Tabla 22: *Clasificación de ítems por formato y capacidad lectora*. Fuente: OCDE (2009b p.44)

Procesos de la Competencia Lectora				
Formato de ítems	Recuperación de la información	Interpretación de textos	Reflexión y evaluación de textos	Total
Elección múltiple	0	9	0	9 (29%)
Elección múltiple compleja	1	0	0	1 (3%)
Respuesta cerrada construida	6	1	0	7 (23%)
Respuesta abierta construida	3	4	7	14 (45%)
Total	10 (32%)	14 (45%)	7 (23%)	31

5.2.3. Pruebas actitudinales

El dominio afectivo fue evaluado mediante tres escalas actitudinales: Interés en ciencias (*interés*); apoyo a la investigación científica (*apoyo*) y responsabilidad medioambiental (*responsabilidad*), como se ve en la tabla 23.

Para la evaluación del espectro actitudinal se aplicaron escalas tipo Likert de cuatro alternativas, para evitar una respuesta neutral. Además se aplicaron ítems de ordenar opiniones (*match-the-opinion*) para evitar sesgos de deseabilidad social en la escala de *responsabilidad*. Todas las cuestiones actitudinales fueron presentadas después de los ítems de aptitudinales.

Tabla 23: *Clasificación de ítems por actitudes científicas*. Fuente: OCDE (2009b p.24)

	Interés	Apoyo	Responsabilidad	Opinión	Ítems actitudinales
Nº de ítems	44	26	23	12	105

5.2.4. Pruebas de contexto

Los cuestionarios de contexto suelen cumplir tres funciones dentro de las evaluaciones internacionales: 1) definen subgrupos de la población de examinados

permitiendo introducir cualificaciones a los resultados socioculturales entre otros; 2) se examinan las relaciones de este tipo de variables con datos de rendimiento educativo y sus influencias sobre éste, y 3) sus datos se utilizan sobre variables de condicionamiento en la imputación y estimación de los valores plausibles (Martínez-Arias, 2006).

Con el objetivo de recabar información contextual, PISA pide al alumnado y a la dirección de los centros que respondan a un cuestionario de duración aproximada 30 minutos. Esta información contenida se concentra prácticamente en los siguientes puntos:

1. Información del contexto sociocultural y económico de las familias;
2. Diversos aspectos de la vida de los alumnos como actitudes, hábitos y vida en el entorno escolar y familias;
3. Aspectos sobre los centros de enseñanza: Calidad de los recursos humanos y materiales del centro, carácter público o privado de su gestión y financiación; procesos de toma de decisiones y prácticas del personal docente;
4. El contexto de aprendizaje, incluyendo tipo y estructura de las instituciones, tamaño de las clases , implicación de padres;
5. Estrategias de aprendizaje autorregulado, preferencias motivacionales y orientación hacia los propios fines, preferencias por determinados tipos de situaciones de aprendizaje, etc.; y
6. Aspectos del aprendizaje y la instrucción de las ciencias, incluyendo la motivación, implicación y grado de confianza en los estudios científicos por parte de los alumnos, así como el impacto de las estrategias de aprendizaje en los logros relacionados con la enseñanza y aprendizaje de las ciencias

Además se recoge a escala internacional, sin hacer una distinción entre países (a causa de la falta de representatividad), dos cuestionarios adicionales: grado de familiaridad con el uso de ordenadores, resultados que se pueden consultar en un informe de la OCDE (2005) titulado *Are Students Ready for a Technology- Rich world? What Pisa tell us (¿Están preparados los estudiantes para un mundo rico en tecnología?: Lo que nos dicen los estudios de PISA)* y un cuestionario dirigido a los padres que se centra en la ocupación de los padres en relación con las ciencias,

valoraciones de la familia sobre medio ambiente o la importancia de conocimientos y habilidades científicas en el mercado laboral, etc.

5.3. *Construcción de escalas*

El procedimiento para escalar datos de PISA se llevó a cabo mediante el desarrollo de modelos de logit multinomial de coeficientes mixtos (Adams, Wilson y Wang, 1997), que son una generalización del modelo de Rasch y que está implementado en el programa Conquest (Wu, Adams y Wilson, 1997). El modelo permite el análisis de interacciones del ítem con otras variables, así como la posibilidad de generar para cada sujeto una distribución posterior con múltiples resultados. También permite especificar modelos diferentes dentro del mismo test para distintos tipos de ítems. Para ítems de respuesta binaria se especificó el modelo logístico de un parámetro y para los de escala, el modelo de crédito parcial (Masters, 1982; Masters y Wright, 1997). Las pruebas de ajuste de los ítems al modelo, se basan en los residuos o diferencias entre las frecuencias observadas y esperadas, que resume el estadístico INFIT.

Posteriormente se llevó a cabo la calibración de los ítems, en este caso la estimación del parámetro dificultad (b), obteniéndose unos intervalos de confianza a través del cual se determinó el ajuste de los ítems. La calibración se realizó en muestras nacionales (de cada país) y en una muestra de calibración internacional final, formada por 500 sujetos extraídos al azar de los 30 países de la OCDE.

Debido al diseño matricial de ítems como anteriormente se ha visto, se deben utilizar complejos procedimientos para la estimación de las puntuaciones de los sujetos con datos incompletos. La metodología aplicada en PISA y otras evaluaciones están basadas en la teoría de las imputaciones de valores perdidos de Rubin (1987). Concretamente a través de un procedimiento desarrollado por varios autores (Mislevy, 1991; Mislevy, Beaton, Kaplan y Sheehan, 1992; Mislevy, Jonson y Muraki, 1992; Mislevy y Sheehan, 1989). Dichas puntuaciones no son útiles para el diagnóstico de sujetos, pero son muy útiles para la estimación de parámetros poblacionales, que es el objetivo de las evaluaciones internacionales: Tratar de estimar de las características de la población en una variable latente y en relación de la variable latente con otras variables (Adams, 2003).

Por otra parte la aplicación del programa Conquest (Wu, Adams y Wilson, 1997) permite obtener cinco valores plausibles. Este tipo de valores son una representación de

la gama de capacidades que puede suponerse a un alumno. En lugar de estimar directamente la aptitud de los estudiantes (θ), se estima una distribución de probabilidades para la θ . Los valores plausibles son valores aleatorios generados a partir de una distribución a posteriori de valores, con unas probabilidades asociadas, las cuales siguen una distribución normal. De esta distribución se obtienen aleatoriamente cinco valores denominados valores plausibles, que proceden de la propia distribución de cada sujeto. Esto se hace para prevenir el sesgo que se produciría estimando la habilidad a partir de un solo conjunto de ítems de un dominio, por tanto la selección de varios valores es necesaria para estimar la varianza error derivada de la imputación (Martínez-Arias, 2006).

Si θ es el estadístico poblacional y θ_i es el estadístico de interés calculado sobre un valor plausible, entonces:

$$\theta = \frac{1}{M} \sum_{i=1}^M \theta_i, \text{ donde } M \text{ es el número de valores plausibles. [5.4]}$$

Siendo la varianza de imputación:

$$B_M = \frac{1}{M-1} \sum_{i=1}^M (\theta_i - \theta)^2 \quad [5.5]$$

La etapa final corresponde combinar la varianza muestral y la varianza de imputación, como se muestra a continuación:

$$V = U + \left(1 + \frac{1}{M}\right) B_M, \text{ donde } U \text{ es la varianza muestral [5.6]}$$

Los estadísticos poblacionales se estiman usando cada uno de los valores plausibles separadamente. El estadístico poblacional es el promedio de los estadísticos obtenidos con cada uno de los valores plausibles (Willms y Smith, 2008). En la edición 2006 de este estudio hay 45 valores plausibles que como se puede ver se distribuyen atendiendo a las competencias, tabla 24.

Tabla 24: *Representación de los valores plausibles obtenidos en PISA 2006*. Fuente: elaboración propia.

Competencia	Valores plausibles
Competencia matemática	PV1MATH-PV5MATH
Competencia lectora	PV1READ-PV5READ
Competencia científica	PV1SCIE-PV5SCIE
<i>Actitudes</i>	
Interés por temas científicos	PVIINTR-PV5INTR
Apoyo a la investigación científica	PV1SUPP-PV5SUPP
<i>Escalas de las capacidades científica</i>	
Identificar cuestiones científicas	PV1SCIE1-PV5SCIE1
Explicar fenómenos científicos	PV1SCIE2-PV1SCIE2
Realizar pruebas científicas	PV1SCIE3-PV1SCIE3

5.4. Construcción de índices de contexto

Las variables de contexto son simples transformaciones o combinaciones lineales de algunas variables, mientras que las medidas de constructos son estimaciones de rasgos latentes derivadas mediante procedimiento de escalamiento.

Aquellos ítems que fueron destinados a medir constructos latentes, se elaboraron a partir de técnicas psicométricas para analizar la dimensionalidad (análisis factoriales exploratorios y confirmatorios) y su consistencia y estabilidad entre países (análisis confirmatorios multi-grupo). Una vez establecidas las dimensiones, los ítems componentes de las variables latentes fueron calibrados con modelos politómicos de respuesta al ítem y las puntuaciones de los estudiantes individuales fueron estimadas con procedimientos de máxima verosimilitud con 500 alumnos de los 30 países de la OCDE.

6. Resultados de España en PISA

Los resultados globales de las tres competencias se constata que los promedios españoles han oscilado de 2000 a 20009 entre 480 y los 493 puntos. Mientras que los promedios de la OCDE oscilan en esas mismas fechas entre 493 puntos y 501 puntos, como se puede ver en la tabla 25.

Tabla 25: Evolución de los resultados globales en las tres competencias en población española. Fuente: elaboración propia. Fuente: IE (2010a).

		2000	2003	2006	2009
Lectura	Promedio OCDE	500	494	492	493
	España	493	481	461	481
Matemáticas	Promedio OCDE		500	498	496
	España		485	480	483
Ciencias	Promedio OCDE			500	501
	España			488	488

Estas diferencias entre los resultados españoles y los promedios OCDE desaparecen cuando se valoran los porcentajes de los alumnos en los niveles bajos de rendimiento y, por el contrario, alcanzan un valor significativo cuando se comparan los porcentajes de alumnos en los niveles altos de rendimiento (IE, 2010a). Con respecto a la equidad, los resultados españoles han mantenido en los ejercicios correspondientes a las tres competencias valores elevados en las diferentes variables utilizadas para medir la eficacia de los sistemas educativos, para ofrecer igualdad de oportunidades a sus alumnos o para compensar las desigualdades que provienen de los entornos sociales, económicos y culturales.

También se constata que en el sistema educativo español, como en el conjunto de los países de la OCDE, se mantienen las diferencias en el tiempo, con muy ligeras variaciones. Las diferencias en los resultados a favor de las alumnas en comprensión lectora y de los alumnos en matemáticas, mientras que en ciencias los resultados son similares.

7. Planteamiento de los objetivos

Como señala Marchesi (2006a) pocos textos educativos generan tanta expectación, ocupan tantas páginas en los medios de comunicación y son una referencia continua en los debates educativos como son los estudios trienales PISA. Los resultados directos de los conocimientos de los alumnos de cada país se convierten casi en un dogma de fe, mientras que el resto de las informaciones contenidas en el informe pasan a un segundo plano y se perciben más como interpretaciones interesadas que como datos igualmente contrastados. El presente trabajo estudia de forma contextualizada el logro académico del alumnado de quince años en población española, a partir de los datos obtenidos en PISA 2006.

Dicho logro académico se evalúa a través de una evaluación externa atendiendo a los objetivos subyacentes de los sistemas educativos (según los define la sociedad), y no en función de la docencia y aprendizaje de un corpus de conocimientos. Esta forma de enfocar el rendimiento educativo se hace necesaria si se desea fomentar que los centros de enseñanza y los sistemas educativos se centren en los retos que plantea la sociedad contemporánea (OCDE, 2006b). Este logro se concreta en términos de competencias. Una competencia es entendida como la capacidad de utilizar esos conocimientos aprendidos, o dicho de otra forma, su utilidad y aplicabilidad a situaciones de la vida real que deberían poseer todos los jóvenes a los quince años, cuando finalizan o están a punto de finalizar la educación obligatoria (OCDE, 2006b).

El trabajo se plantea a la luz de los estudios del Movimiento teórico-práctico de Eficacia Escolar, ya que es la línea de investigación educativa que más ha incidido en la mejora de la educación en los últimos años (Townsend, 2007). Muestra una clara línea de influencia en políticas educativas generadas a partir de las evaluaciones nacionales e internacionales sobre rendimiento académico (Kyriakides, 2007). Concretamente el impacto de dicho movimiento en PISA se encuentra tanto a nivel teórico como metodológico. En cuestiones teóricas su influencia más directa es la generación de indicadores sobre aspectos de rendimiento educativo y recopilación de información del entorno familiar de los alumnos para comprender mejor los factores que influyen en los resultados educativos. Y además, en la publicación de la situación de cada país en términos de equidad y valor añadido, características que componen la definición de centro o escuela eficaz (Murillo, 2005).

Las encuestas internacionales sobre educación están sobre todo interesadas en la muestra de alumnos, la cual se selecciona dentro de una muestra de centros. La muestra de centros se diseña para optimizar la muestra de los alumnos, esto es, para seleccionar una muestra eficiente de alumnos que minimice el coste de las pruebas. Además considera que las variables de los centros deben medirse como atributos de los estudiantes, más que como elementos en sí mismos (González y Kennedy, 2003; Ainley y Ainley, 2011).

Sin embargo, tal y como plantea la OCDE (2006b), la cada vez mayor importancia y popularidad de los análisis multinivel exige dedicarle mayor atención al nivel de centro y exigen una definición común para lo que se debe considerar un centro educativo en los distintos países evaluados. De hecho a nivel metodológico destaca el

interés común del Movimiento y de la OCDE por la aplicación de modelos multinivel. Los cuales dan cuenta de la estructura escolar jerárquica: alumnos están agrupados en aulas, las aulas están en centros escolares, los centros están incluidos en distritos etc., por tanto resulta pertinente el uso de una estrategia multinivel capaz de recoger de forma precisa el medio natural de los datos (Heck y Scott, 2000). Por otra ofrecen la posibilidad de detectar diferentes fuentes de variación a través de los distintos niveles y mediante la introducción de distintas variables predictoras que se van incluyendo progresivamente en forma de efectos fijos o efectos aleatorios.

Uno de los aportes de este trabajo es la aplicación del conocimiento que ofrece la IEE en la evaluación PISA 2006 en población española. Concretamente la estimación del tamaño del efecto del centro y la consistencia de dicho efecto entre variables de distinta naturaleza tanto aptitudinal como actitudinal y entre distintos grupos de alumnado. De tal forma que se aporta conocimiento al estado de la cuestión sobre la estimación de efectos escolares de dos formas. Por una parte se mide variables aptitudinales que en la literatura no se encuentran con frecuencia porque se suele estudiar la influencia de tales efectos a través de la medición de áreas curriculares más que en variables que se construyen a partir del marco de Competencias Básicas (DeSeCo, Definición y Selección de Competencias, 2003). Por otra, se aporta información sobre la casi inexistente investigación sobre dichos efectos en variables afectivas o no cognitivas (Murillo, 2004 y 2007), variables que también son fundamentales para que centros eficaces colaboren en el desarrollo integral de su alumnado tal y como establece la LOE (2/2006).

Como se señaló la detección de esas variables o factores que inciden en el rendimiento ha sido uno de los objetivos comunes del Movimiento y de la evaluación llevada a cabo por la OCDE. Este trabajo profundiza en el estudio de los indicadores evaluados a través de una organización de un modelo CIPP con variables de distinta naturaleza: contexto; entrada, proceso y producto (Stufflebeam y Shinkfield, 1987) y de forma jerárquica, en dos niveles. Puesto que PISA selecciona por cada centro participante una muestra aleatoria de una población de la misma edad de entre todos los cursos y clases, permite la descomposición de la varianza en sólo dos niveles: una varianza entre centros y una varianza dentro de centros. Esta característica supone una limitación fundamental como es la pérdida de información de un nivel intermedio de aula, el cual hace posible la medición de variables de proceso que tienen que ver con el

proceso de enseñanza-aprendizaje, como comportamiento o metodología docente. Aunque este tipo de variables en el programa se recogen en el nivel uno, a través de opiniones del alumnado.

Desde este trabajo se presenta un modelo comprensivo de eficacia escolar en competencia científica bajo el esquema CIPP (Scheerens, 1990; Creemers, 1994; Murillo, 2007) compuesto por indicadores que se adquieren o se fomentan a lo largo período educativo y que están validados en distintas poblaciones y contextos. Supone una novedad dado que actualmente no se ha realizado una explotación similar con datos PISA 2006. Trabajos que han recurrido a la metodología multinivel con datos del programa en las distintas ediciones: son la competencia lectora (Pajares-Box, 2005), la competencia matemática (Marchesi y Martínez-Arias, 2006; Ruiz de Miguel y Castro-Morera, 2006) y la competencia científica en población del País Vasco (ISEI-IVEI, 2008) y mediante un análisis multinivel que especifica el efecto de políticas y prácticas escolares entre el llamado GIP (*Grupo Iberoamericano de PISA*) entre los que se incluye España (OCDE, 2010).

7.1. Objetivo general

En el trabajo se investiga la influencia de los centros educativos en el desarrollo aptitudinal, a través de la Competencia Lectora, Matemática y Científica, y en el desarrollo actitudinal, mediante el Interés hacia las ciencias y el Apoyo a la investigación científica. Se trata de estimar qué cantidad del rendimiento cognitivo y no cognitivo es atribuible a los efectos de centros. También se comprueba si dicha contribución del sistema educativo se da de forma estable en todas las variables criterio analizadas y si la acción del centro educativo es similar para distintos tipos de estudiantes.

7.2. Objetivos específicos

Objetivo 1. Estimar el tamaño del efecto de centro para las medidas cognitivas (competencia matemática, competencia lectora y competencia científica) y las actitudinales (interés hacia las ciencias y el apoyo a la investigación científica).

Objetivo 2. Calcular la consistencia de los efectos escolares entre las distintas variables aptitudinales (competencia matemática, competencia lectora y competencia científica) y actitudinales (interés hacia las ciencias y el apoyo a la investigación científica).

Objetivo 3. Determinar la eficacia diferencial de los centros en función de diferentes características personales y sociales del alumnado, y para las distintas medidas de rendimiento cognitivo (matemáticas, lectora y científicas) y no cognitivo (interés hacia las ciencias y el apoyo a la investigación científica).

Objetivo 4. Determinar la aportación del contexto y del proceso de enseñanza-aprendizaje asociados a la adquisición de la competencia científica

Objetivo 5. Elaborar un modelo de Eficacia Escolar adecuado a la situación social y educativa científica en educación secundaria española

7.3. Hipótesis

Se plantea una hipótesis de trabajo para cada uno de los objetivos específicos formulados:

Hipótesis 1. Se espera encontrar una magnitud del efecto escolar entre el 15% y 20% de varianza explicada en el nivel dos de centro, como se refleja en la literatura de estudios PISA en población española (IE, 2010; Martínez-Arias y Marchesi, 2006; Pajares-Box, 2005). Mientras que el efecto escolar en variables actitudinales será moderado, igual o inferior al 10% de varianza explicada, de acuerdo a estudios como Fitz-Gibbon (1991) o Murillo (2007).

Hipótesis 2. Se estima una consistencia moderadamente alta de los efectos de centro. Los centros presentarán una mayor consistencia de sus efectos entre los resultados competenciales que entre las actitudes evaluadas (Fitz-Gibbon, 1991; Mortimore et al., 1988; Murillo, 2007)

Hipótesis 3. Se espera que no todos los centros se comporten de forma similar atendiendo a características del alumnado como sexo, índice socioeconómico y cultural familiar, condición de repetidor y de inmigrante, como se ha demostrado en otras cohortes de alumnos (Marchesi y Martínez-Arias, 2006; Pajares-Box, 2005) y en otros estudios de eficacia escolar diferencial (Jesson y Gray, 1991; Murillo, 2008 y 2011, Sammons, Nuttall y Cuttance, 1993; Willms, 1992).

Hipótesis 4. Es previsible encontrar un elevado porcentaje de varianza explicada en variables de contexto, tanto en el nivel de alumnado como de centro. Incluso el porcentaje de variables de contexto será superior a los porcentajes de varianza explicada atribuida a variables de proceso en los dos niveles (Coleman, 1966 y 1986; Marchesi y

Martínez-Arias, 2006; Martín et al., 2008; Murillo, 2008). Lo que implica que indicadores de proceso cuando se evalúen conjuntamente con variables de contexto pierdan su significación por una colinealidad entre ambas (ISEI-IVEI, 2008; Pajares-Box, 2005).

Hipótesis 5. Una vez que se elabora el modelo multinivel final de competencia científica se espera encontrar el impacto del contexto, de la motivación y autoeficacia del alumnado (Martínez-Arias y Marchesi, 2006; OCDE, 2010; Pajares-Box, 2005) y de variables de aula medidas a través del alumnado que dan cuenta del comportamiento docente y sus prácticas como información y concienciación de problemas medioambientales (Creemers y Kyriakides, 2008; Creemers y Kyriakides, 2010; OCDE, 2010). En el nivel dos se espera un mayor efecto de indicadores de características contextuales de la instrucción como el índice socioeconómico de las familias o los recursos educativos del centro (Grisay, 2007; Marchesi y Martínez-Arias, 2006; Murillo, 2007 y 2011) y se encontrará un impacto de variables que tienen que ver dirección de liderazgo y organización de centro en procesos de enseñanza- aprendizaje y evaluación (Cantón y Arias, 2008; Louis, Dretzke, y Wahlstrom, 2010; Waters, Marzano y McNulty, 2003).

8. Índice de tablas

Tabla 13: <i>Contexto de la evaluación de ciencias.</i> Fuente: OCDE (2006a).....	112
Tabla 14: <i>Capacidades científicas en PISA 2006.</i> Fuente: elaboración propia.....	113
Tabla 15: <i>Conocimientos de la competencia científica en PISA 2006.</i> Fuente: elaboración propia.....	114
Tabla 16: <i>Categorías del conocimiento acerca de las ciencias en PISA 2006.</i> Fuente: elaboración propia.....	115
Tabla 17: <i>Áreas de evaluación de las actitudes en PISA 2006.</i> Fuente: elaboración propia.....	116
Tabla 18: <i>Replicaciones del método BRR.</i> Fuente: OCDE (2009b).....	125
Tabla 19: <i>Distribución de cuadernillos en PISA 2006.</i> Fuente: OCDE (2009b)	129
Tabla 20: <i>Clasificación de ítems por formato y capacidad en ciencias.</i> Fuente: OCDE (2009b).	130
Tabla 21: <i>Clasificación de ítems por formato y capacidad en matemáticas.</i> Fuente: OCDE (2009b).....	131

Tabla 22: <i>Clasificación de ítems por formato y capacidad lectora</i> . Fuente: OCDE (2009b)	131
Tabla 23: <i>Clasificación de ítems por actitudes científicas</i> . Fuente: OCDE (2009b)....	131
Tabla 24: <i>Representación de los valores plausibles obtenidos en PISA 2006</i> . Fuente: elaboración propia.	135
Tabla 25: <i>Evolución de los resultados globales en las tres competencias en población española</i> . Fuente: elaboración propia. Fuente: IE (2010).....	136

9. Índice de figuras

Figura 20: <i>Elementos que componen la competencia científica</i> . Fuente: OCDE (2006a).	112
Figura 21: <i>Elementos del área de matemáticas</i> . Fuente: OCDE (2006a).....	117
Figura 22: <i>Capacidades incluidas en la competencia de matemáticas</i> . Fuente: OCDE (2006a).....	119
Figura 23: <i>Subescalas de procesos que componen la competencia lectora</i> . Fuente: OCDE (2006a).	122

CAPÍTULO 6
MÉTODO

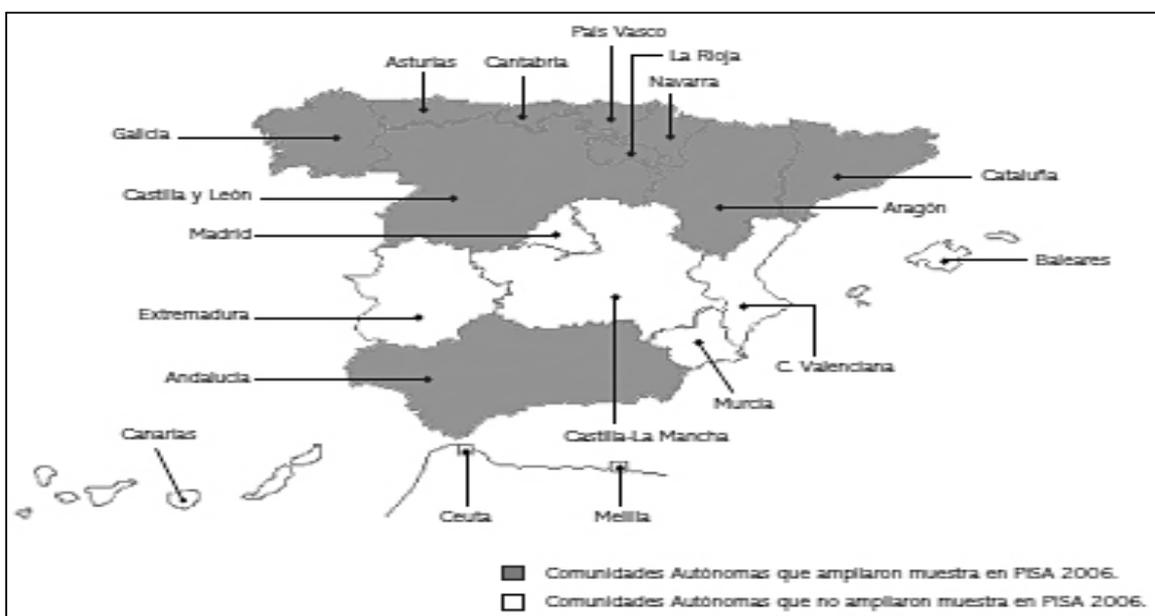
Capítulo 6: Método

1.	Muestra	144
2.	Instrumentos y variables	145
2.1.	Cuestionario de estudiante.....	145
2.1.1.	Variables de contexto/entrada.....	146
2.1.2.	Variables de proceso	149
2.2.	Cuestionario de competencias	150
2.2.1.	Variables resultado.....	151
2.3.	Cuestionario de centro.....	152
2.3.1.	Variables de contexto/entrada.....	152
2.3.2.	Variables de proceso de nivel dos.....	153
3.	Procedimiento	154
4.	Análisis de datos	156
4.1.	Objetivo 1. Estimar el tamaño del efecto de centro	156
4.2.	Objetivo 2. Calcular la consistencia de los efectos escolares entre las distintas variables aptitudinales y actitudinales	157
4.3.	Objetivo 3. Determinar la eficacia diferencial de los centros.....	157
4.4.	Objetivo 4. Determinar la aportación del contexto y del proceso de enseñanza-aprendizaje en la adquisición de la competencia científica	158
4.5.	Objetivo 5. Elaborar un modelo de Eficacia Escolar adecuados a la situación social y educativa científica en educación secundaria española	159
5.	Resumen de tablas	160
6.	Resumen de figuras.....	160

1. Muestra

En la aplicación PISA 2006 los resultados obtenidos no son sólo estatales sino también son representativos de Comunidades Autónomas. Concretamente tal y como recoge la figura 24 diez Comunidades pudieron hacer explotaciones propias de sus datos en el programa. Un hecho sin precedente dado que en la primera edición de 2000 la muestra española de alumnos de 15 años fue estatal y en 2003 tan sólo Castilla y León, Cataluña y el País Vasco obtuvieron resultados propios, dado que aportaron un tamaño de muestra representativo de su población.

Figura 24: *Representación de la participación de España en PISA 2006.* Fuente: IE (2007 p. 19).



La muestra del estudio que se presenta corresponde a las diez Comunidades donde se evaluaron a 17.528 alumnos y cuya media de la edad fue de 15,84 con una desviación típica de 0,29; el 50% son alumnos y el otro 50% son alumnas.

Por lo que respecta al tipo de centro el 56,5% de los 612 centros evaluados son públicos, es decir, son instituciones controladas y dirigidas directa o indirectamente por una autoridad pública de Educación, organismo gubernamental o consejo directivo designado o elegido por el gobierno. Mientras que el 43,5% corresponde a centros controlados o dirigidos directa o indirectamente por una organización no gubernamental, por ejemplo órdenes religiosas, sindicatos, empresas u otras instituciones privadas. En cuanto a la ubicación del centro, la figura 25 refleja una

distribución de los centros en función de su densidad de población, como se puede ver la mayoría de los centros se ubican en núcleos urbanos.

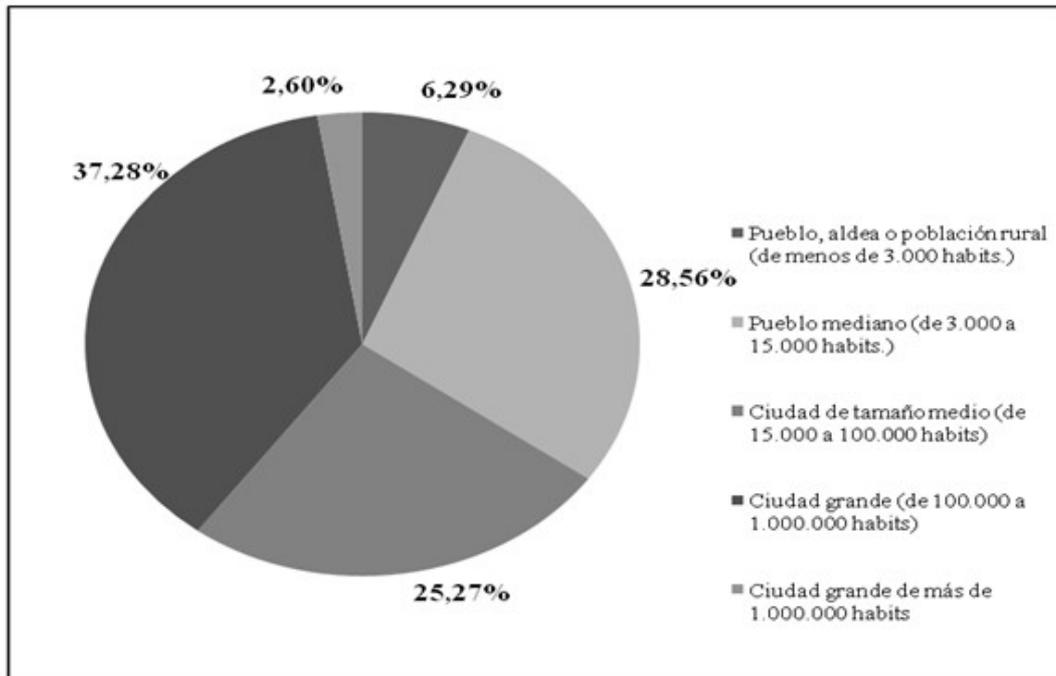


Figura 25: *Distribución de centros según el tipo de población.* Fuente: elaboración propia.

2. Instrumentos y variables

2.1. Cuestionario de estudiante

Este cuestionario estaba compuesto por ocho secciones diferenciadas en contenidos: *sobre ti; tu familia y tu casa; tu opinión sobre la ciencia; el medio ambiente; profesiones relacionadas con la ciencia; tiempo dedicado al estudio de las clases de ciencias; utilización del ordenador*, las cuales recogían los 224 ítems planteados. Algunos de las preguntas eran dicotómicas como el tipo de formación de los padres: título universitario (*sí; no*); otras se valoraban en una escala de frecuencia como el tiempo dedicado al estudio (*nada; menos de dos horas a la semana; más de dos pero menos de cuatro a la semana; más de cuatro pero menos de seis horas a la semana; seis o más horas a la semana*) o, por ejemplo, en los próximos 20 años la contaminación del aire se contestaba en una escala de tres opciones (*mejorarán; seguirán casi igual; empeorarán*) etc. El anexo 2 muestra cómo se valoró cada ítem además de las preguntas que componen cada índice objeto de análisis.

Como indica el informe técnico (OCDE, 2009b) la elaboración de los índices se llevó a cabo mediante la aplicación de modelos de ecuaciones estructurales (*Structural Equation Modelling [SEM]*) para confirmar las dimensiones teóricas esperadas (Kaplan, 2000). Se tuvieron en cuenta diferentes índices para comprobar el ajuste de cada modelo como la raíz media cuadrática (*root-mean square error of approximation [RMSEA]*); la raíz media cuadrática residual (*the root mean square residual [RMR]*) que indican un buen ajuste si el valor obtenido es de 0,05 o inferior; el índice de ajuste comparativo (*the comparative fit index [CFI]*) y el índice de ajuste no normalizado comparativo (*non-normal fit index [NNFI]*) cuya interpretación de un buen ajuste es si sus valores son superiores a 0,95 (OCDE, 2009b; Ferrando y Anguiano-Carrasco 2010). Se realizó una doble estimación del modelo, por una parte se basó en una calibración internacional seleccionando a 500 estudiantes por país y, por otra, se llevó a cabo una calibración propia de cada país.

Para el caso de índices compuestos únicamente por ítems dicotómicos, la estimación de mínimos cuadrados (*Weighted Least Squares [WLS]*) se basó en una matrices policóricas con media y desviación ajustada (*Mean –and –Variance – adjusted WLS estimator [WLSMV]*). La estimación de este tipo de modelo se basó únicamente en la calibración internacional.

Los índices se normalizaron de modo que la media del valor del índice de la población estudiantil de la OCDE toman un valor de cero y una desviación típica de uno. Los valores negativos únicamente indicaban que los encuestados respondieron menos positivamente que lo que lo hizo la media de los países de la OCDE. De esa misma forma, un valor positivo en un índice indica que los encuestados respondieron más favorable que la media de los países de la OCDE.

Los siguientes apartados recogen las variables e índices analizados en el nivel uno de alumnado y nivel dos de centro bajo un esquema CIPP.

2.1.1. Variables de contexto/entrada

La tabla 26 resume todas las variables e índices de contexto-entrada que fueron tratadas en el nivel 1. Existen una serie de variables de diferente tipo como variables dummies¹ tales como ser una estudiante, alumno repetidor etc. que son dicotómicas y

¹ Variable dummy o muda es una variable numérica que transforma en 0 y 1 las categorías de una variable cualitativa para permitir hacer cálculos con los modelo de regresión y los modelos multinivel.

una variable dictomizada como SAMELANG. También se tienen en cuenta dos índices como SCIEPADS que se refiere si alguno de los progenitores desarrolla una ocupación científica y tecnológica² y el índice socioeconómico y cultural (ESCS).

ESCS se compone de las posesiones en el hogar (HOMEPOS) elaborado a partir de WEALTH; CULTPOS; HEDRES y HOMEPOS; por el nivel ocupacional más alto de la familia (HISEI) y el nivel educativo más alto de la familia (PARED). Los valores perdidos en un sólo componente fueron imputados con valores procedentes de análisis de regresión lineal.

Tabla 26: *Variables contexto/entrada de alumnado*. Fuente: elaboración propia.

Nombre	Descripción
FEMALE	Ser alumna (0=alumno;1=alumna).
NATIVE	Estatus de nativo (0=inmigrante; 1=nativo).
REP	Condición de repetidor (0=va al curso modal; 1= repetidor).
SAMELANG	Lengua hablada (0= la prueba está en la misma lengua que habla en casa; 1= la prueba está en distinta lengua a la hablada en casa).
ESCS	Índice socioeconómico y cultural
ESCS2	Índice socioeconómico y cultural cuadrático
SCIEPROF	Profesión relacionada con las ciencias de alguno de los progenitores

El índice de posesiones en el hogar expresa la cantidad de bienes familiares (Buchmann, 2000). El alumnado contestaba a ítems que hacían referencia a poseer distintos bienes adaptados a cada país. De tales reactivos se derivaron cuatro factores como fueron WEALTH que indica posesiones de bienestar; CULTPOSS que indica posesiones culturales; HEDRES se refiere a los recursos educativos del hogar y HOMEPOS es un índice de posesiones. Realmente HOMEPOS resume todos los indicadores de WEALTH, CULTPOSS y HEDRES y añade el número de libros. La tabla 27 se ven los ítems que componen cada índice para países de la OCDE y las propiedades psicométricas del índice ESCS se recogen en la tabla 28.

² Profesión relacionada con las ciencias va más allá de la idea tradicional que considera al científico como alguien que trabaja en un laboratorio o en un ambiente académico. Así una profesión relacionada con la ciencia es cualquiera que implique una formación universitaria en cualquier campo científico. Ejemplos de profesiones: informático, estadístico, arquitecto, profesiones encuadradas en ciencias de la salud, en ciencias sociales, inspectores de seguridad y calidad, controladores aéreos, técnicos de seguridad vial etc. (OCDE, 2009b).

Tabla 27: *Índices de posesiones y contexto del hogar*. Fuente: OCDE (2009b p.316).

Ítem código	En tu casa dispones de ...	WEALTH	CULTPOSS	HEDRES	HOMEPOS
ST13Q01	Un escritorio			x	x
ST13Q02	Habitación propia				x
ST13Q03	Un lugar tranquilo de estudio	x		x	x
ST13Q04	Ordenador para trabajos de clase			x	
ST13Q05	Software educativo			x	x
ST13Q06	Conexión a Internet	x			x
ST13Q07	Tu propia calculadora			x	x
ST13Q08	Libros de literatura		x		x
ST13Q09	Libros de poesía		x		x
ST13Q10	Trabajos artísticos (e.g. cuadros)		x		x
ST13Q11	Libros que ayudan a hacer deberes			x	x
ST13Q12	Un diccionario			x	x
ST13Q13	Un lavavajillas	x			x
ST13Q14	Un DVD	x			x
ST13Q15	Ítem específico de cada país	x			x
ST13Q16	Ítem específico de cada país	x			x
ST13Q17	Ítem específico de cada país	x			x
	¿Cuántos hay en tu casa de...?	ST14			
ST14Q01	Teléfonos móviles	x			x
ST14Q02	Televisiones	x			x
ST14Q03	Ordenadores	x			x
ST14Q04	Coches	x			x
ST15	¿Cuántos libros hay en tu casa?				x

Nota: Las opciones de respuesta de los ítems ST13 fueron sí (1) vs no (2); las opciones de respuesta para las preguntas de ST14 fueron "ninguna", "una" y "tres o más"; las categorías para ST15 fueron recodificadas en tres categorías: de 0-25 libros (0); 26-100 libros (1) y 101 o más (2).

En el Análisis Factorial Confirmatorio (AFC) se obtuvieron índices de ajuste adecuados (RMSEA =0,080, CFI=0,88, NNFI=0,92). Las correlaciones entre los constructos fueron de 0,8 para la relación entre WEALTH y HEDRES; de 0,25 para WEALTH y CULTPOSS y de 0,52 entre CULTPOSS y HEDRES (OCDE, 2009b).

HISEI es otro de los factores que configuran ESCS, procede de la conjunción de dos variables. Por una parte recoge la profesión del padre y de la madre, cuyas respuestas fueron codificadas en cuatro dígitos según los códigos de la escala ISCO (ILO, 1990) y, por otra, tiene en cuenta el índice socioeconómico del estatus ocupacional (ISEI) (Ganzeboom, de Graaf, y Treiman, 1992). De esta unión se obtiene el HISEI que es el estatus ocupacional más alto de los progenitores.

El mayor nivel educativo de los padres expresado en número de años (PARED) es la transformación en número de años a la escala ISCED (OCDE, 1999) cuyas categorías son: "ninguna" (0); (1) "educación primaria"; (2) "primer ciclo de

secundaria”; (3) “segundo ciclo de secundaria” o “formación profesional de grado medio”; (4) “bachillerato” o “formación profesional de grado superior”; (5) “estudios de grado y (6) estudios de posgrado”.

Tabla 28: *Análisis de fiabilidad y validez del ESCS*. Fuente: elaboración propia.

Nombre	Descripción	Alfa	Factor de carga
ESCS		0,69	
HISEI	Estatus ocupacional más alto de los padres		0,84
PARED	El nivel educativo más elevado de los padres		0,82
HOMEPOS	Posesiones en el hogar		0,70

2.1.2. Variables de proceso

La descripción de los índices de proceso de nivel 1 que fueron incluidos como índices de proceso de nivel individual se agruparon bajo diferentes modelos teóricos propuestos por la OCDE (2007b). Las propiedades psicométricas que subyacen a dichos modelos para población española (OCDE, 2009b) figuran en las tablas 29 y 30.

Tabla 29: *Coefficientes de fiabilidad de los índices de proceso de nivel 1*. Fuente: elaboración propia.

Nombre	Descripción	Alfa
<i>Modelo motivación</i>		
INTSCIE	Interés por aprender temas científicos	0,83
JOYSCIE	Gusto por el estudio de las ciencias	0,90
SCIEFUT	Motivación futura para estudiar un grado de ciencias	0,92
INSTSCIE	Motivación instrumental de estudiar ciencias	0,93
<i>Modelo autoeficacia</i>		
SCIEEFF	Autoeficacia en ciencias	0,83
SCSCIE	Autoconcepto en ciencias	0,92
<i>Modelo de orientación hacia carreras científicas</i>		
CARPREP	Preparación que ofrece el centro para estudiar una carrera científica	0,78
CARINFO	Información que ofrece el centro para estudiar una carrera científica	0,80
<i>Modelo creencias científicas</i>		
GENSCIE	Valor general a las ciencias	0,72
PERSCIE	Valor personal a las ciencias	0,79
SCIEACT	Actividades científicas del centro	0,78
<i>Modelo medio ambiente</i>		
ENVAWARE	Preocupación por el medio ambiente	0,77
ENVPERC	Percepción del medioambiente	0,79
ENVOPT	Optimismo por problemas medioambientales	0,78
RESPDEV	Responsabilidad hacia un desarrollo sostenible	0,75

Modelo proceso E/A

SCINTACT	Método de enseñanza de las ciencias	0,77
SCHANDS	Valoración de actividades de laboratorio	0,74
SCINVEST	Iniciativa investigadora de los estudiantes	0,75
SCAPPLY	Aplicación de las ciencias a otros campos	0,77

La tabla 29 señala el Alfa de Cronbach de cada uno de los índices, encontrando la menor consistencia interna en GENSCIE (0,72) y la mayor en INSTSCIE (0,93). La tabla 30 indica distintos índices de ajuste y las correlaciones entre índices de proceso.

En general los modelos creados a partir de la relación entre los constructos presentan adecuados índices de ajuste, con la excepción del compuesto por interés y diversión en las ciencias el cual presenta índices mejorables (RMSEA=0,139 ó NNFI=0,84).

Tabla 30: *Análisis Factorial Confirmatorio de los índices de proceso de nivel 1*. Fuente: elaboración propia.

Nombre	Correlación entre los constructos latentes	RMSEA	RMR	CFI	NNFI
INTSCIE	0,77	0,139	0,068	0,84	0,84
JOYSCIE					
SCIEFUT	0,78	0,097	0,027	0,97	0,97
INSTSCIE					
SCIEEFF	0,46	0,055	0,032	0,97	0,97
SCSCIE					
GENSCIE	0,77	0,08	0,027	0,93	0,93
PERSCIE					
CARPREP	0,45	0,078	0,021	0,95	0,95
CARINFO					
	RESPDEV/ENVAWARE (0,43)				
ENVAWARE	RESPDEV/ENVPERC (0,45)				
ENVPERC	RESPDEV/ENVOPT (-0,18)	0,041	0,022	0,93	0,93
ENVOPT	ENVAWARE/ENVPERC (0,21)				
RESPDEV	ENVAWARE/ENVOPT (-0,3)				
	ENVPERC/ENVOPT (-0,17)				
	SCINTACT/SCHANDS (0,61)				
SCIENTACT	SCINTACT/SCINVEST (0,70)				
SCHANDS	SCINTACT/SCAPPLY (0,66)	0,09	0,051	0,88	0,88
SCINVEST	SCHANDS/SCINVEST (0,78)				
SCAPPLY	SCHANDS/SCAPPLY (0,69)				
	SCINVEST/SCAPPLY (0,68)				

2.2. Cuestionario de competencias

Todas las competencias evaluadas se presentaron en el mismo cuestionario, como se ha visto en el anterior capítulo los ítems están distribuidos en bloques. Un

bloque corresponde a una unidad temática: distintas preguntas planteadas en un contexto o tema común, ejemplo de unidades temáticas se recogen en el anexo 1. Se presentaron ítems que son exclusivos de cada edición y otros que son comunes a los distintos años de aplicación para poder anclar las distintas cohortes de estudiantes. En total 108 preguntas evaluaban competencia científica; 48 matemáticas; 31 lectura y 105 medían actitudes científicas.

El número de unidades o clusters de ítems fueron trece (siete de ciencias; cuatro de matemáticas y dos de lectura). Cada cuadernillo estaba compuesto por cuatro unidades o clusters y la duración de cada unidad era de treinta minutos.

2.2.1. Variables resultado

Cada cuadernillo fue contestado por un 7,69% de la muestra total de alumnado. La distribución de los estudiantes de la OCDE y de los ítems en cada competencia se puede ver en el anexo 3 (OCDE, 2009b pp. 210-214). En dicho anexo aparecen cinco gráficos, tres de las competencias y dos de las actitudes evaluadas, en donde la parte de la derecha expresa la distribución de los estudiantes (representados por Xs) en torno al nivel de competencia estimado en el modelo de Rasch, de tal forma que el alumnado que se encuentra en la cima de la distribución son los más competentes.

En cuanto a la fiabilidad de las escalas, el Alfa de Cronbach en población española fue de 0,90 para matemáticas, de 0,91 para lectura, de 0,92 para ciencias inferior, de 0,89 para interés y de 0,82 para apoyo (OCDE, 2009b p.217).

Para asegurar la validez entre países de los constructos medidos además de llevar a cabo un cuidado proceso de traducción y adaptación a los diferentes idiomas fue necesario que las habilidades, actitudes y percepciones medidas fueran similares en los diferentes contextos evaluados. Las técnicas Psicométricas demostraron una consistente dimensionalidad y validez de constructo a través de los países participantes (OCDE, 2009b). La última edición del Informe Técnico de PISA prueba su modelo multi-dimensional de cinco dimensiones (lectura, matemáticas, ciencias, interés y apoyo científico) a través del ajuste de los datos en TRI mediante el programa Conquest (Wu, Adams y Wilson, 1997).

Además dicho informe incluye análisis de Funcionamiento Diferencial de los Ítems (FDI) mediante modelos TRI entre los países participantes y el género de los estudiantes.

2.3. Cuestionario de centro

El cuestionario de centro recoge las variables e índices contruidos a partir de las respuestas de la dirección. Se compone de 84 ítems organizados en seis secciones: *estructura y organización del centro; personal docente; recursos del centro; estrategias de información sobre los resultados y normas de admisión de alumnos; actividades relacionadas con las ciencias; orientación profesional y estudios posteriores*. El tiempo estimado para su cumplimentación fue de 30 minutos.

2.3.1. Variables de contexto/entrada

Las variables e índices de contexto/entrada tenidos en cuenta en el nivel dos se distribuyen en diferentes submodelos que se presentan en la tabla 31. RURAL y CITY son variables dummies que proceden de una pregunta en donde se pedía que eligieran la opción que mejor describía la población en la que se situaba el centro. Rural se llamó a centros situados en poblaciones que no superan los 15.000 habitantes y ciudad se consideró a aquellos centros localizados en áreas con 15.000 o más habitantes. PRIV es una variable dicotómica que indica que el centro es privado y GOVFUND es un índice que cuantifica las ayudas y subvenciones que recibe el centro y las familias. SCHLSIZE y CLSIZE son variables cuantitativas continuas que cifran el número de estudiantes del centro y se incluye su transformación cuadrática. Desde LOSELE hasta PRESSPA son una serie de variables dicotomizadas. FACS, FACB y FACS proceden de tres análisis factoriales que resumen ítems que preguntaban sobre la autonomía del centro. STRATIO procede de la división entre el tamaño del centro y el número de profesores. El número de profesores con una dedicación de media jornada se les dio un peso de 0,5 mientras que el profesorado a tiempo completo se le dio un 1,0. TCSHORT es un índice construido a partir de la percepción que tiene la dirección de falta de profesorado. Dichos ítems se evaluaron en una escala de cuatro opciones de respuesta: *nada, muy poco, algo y mucho*. En cuanto al número de ordenadores se obtiene dividiendo el número de ordenadores dedicados a la enseñanza entre el número de estudiantes (IRATCOM). Por último SCMATEDU recoge la percepción que tiene la dirección sobre la falta de recursos para la instrucción como falta de conexión a Internet, de programas informáticos, escasez de materiales de consulta de la biblioteca, de medios audiovisuales etc. y su escala de valoración es igual a TCSHORT.

El informe técnico ofrece análisis de consistencia interna de TCSHORT (Alfa = 0,84) y SCMATEDU (Alfa = 0,85) (OCDE, 2009b p.341).

Tabla 31: *Variables de contexto/entrada de nivel 2*. Fuente: elaboración propia.

Nombre	Descripción
Modelo características del centro	
CITY	Situación urbano en el que está localizado el centro (1=urbano; 0=rural)
RURAL	Situación rural en el que está localizado el centro (1=rural; 0=urbano)
PRIV	Financiación privada (1=privada; 0=pública)
PUB*	Índice de ayudas gubernamentales
SCHSIZE*	Tamaño del centro
SCHSIZE2*	Índice del tamaño del centro cuadrático
CLSIZE*	Tamaño del grupo aula
CLIZE2*	Índice del tamaño del grupo aula cuadrático
ESCS centro*	Es la media del ESCS del alumnado del centro
Modelo admisión y contexto de instrucción	
LOSELE	Sin criterios de admisión del alumnado (1=sin criterio de admisión; 0=con criterios de admisión)
HISELE	Criterios de admisión del alumnado (1=con criterio de admisión; 0=sin criterios de admisión)
ABGROUP	Agrupamiento del alumnado por competencia (1=en todas las asignaturas; 0=en algunas asignaturas)
PRESSPA	Expectativas de las familias hacia el centro (1=Sí; 0= No)
Modelo gestión	
FACS*	Autonomía en la gestión del personal (contratación, despido, salario inicial y aumento de sueldo)
FACB*	Autonomía en presupuesto (propuesta y asignación de presupuesto)
FACC*	Autonomía curricular (oferta de cursos, contenidos de cursos y libros de texto)
Modelo de recursos educativos	
STRATIO*	Ratio profesor-alumno
TCSHORT*	Índice de escasez de profesorado
IRATCOM*	Ratio de ordenadores para dar clase
SCMATED*	Índice de recursos educativos del centro

Nota: Las variables con asterisco son continuas.

2.3.2. *Variables de proceso de nivel dos*

Las variables de proceso de este nivel se agrupan en dos submodelos tal y como se presentan en la tabla 32.

Tabla 32: *Variables de proceso de nivel 2*. Fuente: elaboración propia.

Nombre	Descripción
Modelo de Enseñanza-Aprendizaje (E/A)	
TIMELEARN*	Tiempo de aprendizaje total (media de ítems de alumnado)
ANYSCIE*	Aprendizaje de ciencias en el currículo
ENVLEARN*	Actividades del centro relacionadas con el aprendizaje de tópicos medio ambientales (e.g. visitas a museos, parques tecnológicos etc.)
SCIPROM*	Actividades escolares que promueven el aprendizaje científico (e.g. club de ciencia, ferias etc.)
Modelo de evaluación y mejora	
ACC1-ACC3	El centro informa a las familias sobre el rendimiento de sus hijos comparándolo con otros estudiantes; con baremos de la región; estudiantes de otros centros que están en el mismo grado (1=Sí; 0= No)
ACC4-ACC8	Los resultados del alumnado es útil para: publicidad; se evalúa a la dirección; se evalúa al profesorado; se toman decisiones para administrar fuentes educativas; seguimiento por parte de autoridades administrativas (1=Sí; 0= No)

Las variables dicotómicas de evaluación (ACC1- ACC8) se refieren a estrategias de información sobre los resultados que aplica al centro. La variable TIMELEARN se refiere al tiempo de aprendizaje medio del centro que invierte el alumnado. Para ello se les pedía a los estudiantes que cuantificaran el tiempo de horas a la semana que empleaban en estudiar lengua castellana, matemáticas, ciencias entre otras asignaturas. Para contar el tiempo empleado se les pedía que valoraran el tiempo en clases normales en el centro; el tiempo en clases de recuperación, apoyo o particulares y el tiempo en hacer deberes. ANYSCIE es una suma de una serie de preguntas dicotómicas (sí vs no) sobre si ha cursado en el pasado año o en el año de la aplicación diferentes asignaturas relacionadas con las ciencias como biología, física, química etc. tanto en la modalidad optativa como obligatoria. Por último SCIPROM recoge una serie de ítems dicotómicos (sí vs no) que evalúan a la dirección a cerca de la participación del alumnado de 4º de la ESO para aprender temas medioambientales como actividades al aire libre, visitas a museos a centros científicos y tecnológicos etc. según la opinión de la dirección

Por lo que se refiere a ENVLEARN y SCIPROM destacan sus bajos valores de consistencia interna, ya que se encuentra en torno a 0,5 (OCDE, 2009b p.341).

3. Procedimiento

La OCDE desde su página (www.pisa.oecd.org) publica todos los archivos necesarios para hacer investigaciones paralelas a sus publicaciones. A través de la misma página fue posible descargar los cuestionarios de contexto, los libros de códigos, los archivos de sintaxis y macros para leer los datos mediante el SPSS 15. El resultado

fue una base con las respuestas de contexto de todos los países evaluados, en donde se seleccionaron los datos de España cuyo código correspondía en la variable país con el valor 724. Una vez construida una nueva base de datos se seleccionaron los centros y estudiantes de aquellas diez Comunidades que participaron de una forma representativa en el programa. Para llevar a cabo los posteriores análisis multinivel hubo que realizar algunos ajustes como a continuación se expone.

La mayor parte de los indicadores expuestos están incluidos en la misma base de datos y, como se indica en los materiales de referencia de PISA, no tienen que ser recalculados, salvo los casos en los que hubiera omisiones que deben ser tratados por procesos de imputación.

Todas las imputaciones se acompañaron de una variable muda siguiendo el modelo propuesto por Cohen y Cohen (1983). Siguiendo a estos autores se siguen una serie de pasos dependiendo del tipo de variable que se trate:

1. Para las variables continuas del nivel de alumnado los valores perdidos se sustituyeron por el promedio ponderado de la variable en el centro, mientras que el valor imputado para el valor del centro es el promedio de la Comunidad. Los pesos o ponderaciones que se asignaron a cada estudiante para que represente al número de alumnos en la población fue a nivel escolar, proporcionados directamente en la base de datos de PISA con la etiqueta W_FSTWUWT.
2. Para las variables categóricas la identificación se hizo en función de una variable muda. Se asignó 1 si la observación pertenecía a la categoría y 0 en caso contrario. Para las otras variables dummies y dicotómicas, los casos faltantes se imputaron con 0.

Tras lo comentado, la primera tarea en HLM 6.03 [*Hierarchical Linear and Nonlinear Modelling*] fue construir la matriz MDM [*Multivariate Data Matrix*] uniendo las variables de ambos niveles a través del código de identificación del centro. Las variables resultado son los cinco valores plausibles para obtener cinco estimadores (por ejemplo, del intercepto del modelo de regresión), los cuales se promedian para proporcionar el valor final de la estimación de los parámetros del modelo jerárquico deseado.

Para los objetivos cuatro y cinco centrados en construir un modelo multinivel de competencia científica se utilizó los pesos de los centros y de los estudiantes de la base

de datos de PISA en el cálculo de los estimadores del modelo (Marchesi y Martínez-Arias, 2006c).

Se aplicó una segunda estandarización de los coeficientes obtenidos en el HLM dado que este programa no ofrece valores estandarizados. Dicha transformación consistió en multiplicar el valor del coeficiente por el su desviación típica (Willms y Smith, 2005).

4. Análisis de datos

Los siguientes apartados se refieren a los análisis seguidos para operativizar cada objetivo planteado. Todos los pasos que se especifican están encuadrados dentro de una aproximación multinivel utilizando el procedimiento de Máxima Verosimilitud Completa. Las puntuaciones en las variables predictoras cuantitativas fueron centradas en la media global como indican Willms y Smith (2005).

4.1. Objetivo 1. Estimar el tamaño del efecto de centro

El primer objetivo planteado trató de estimar la magnitud del tamaño del efecto de centro para las medidas cognitivas (Competencia Matemática, Competencia Lectora y Competencia Científica) y actitudinales (Interés hacia las ciencias y el Apoyo a la investigación científica).

Para cuantificar el tamaño de los efectos de centro se hizo una distinción entre efectos brutos y netos (Rodríguez-Jimenez y Murillo, 2011). Por una parte se estimaron los *efectos escolares brutos* que no incluyen ninguna variable de ajuste, como es el caso del modelo nulo. El modelo nulo o modelo vacío contiene sólo una variable respuesta y la constante. Este proceso se repite en cada una de las variables output con los cinco valores plausibles. Por otra se calcularon los *efectos escolares netos*, los cuales corresponden a los efectos de centro escolar una vez controlado el ESCS del alumnado y de centro. Para ambos modelos se calculó el impacto del centro mediante el Coeficiente de Correlación Intraclase (CCI).

Efectos escolares brutos:

$$\text{Nivel-1} \quad Y = B_0 + R \quad [6.1]$$

$$\text{Nivel-2} \quad B_0 = G_{00} + U_0 \quad [6.2]$$

Efectos escolares netos:

$$\text{Nivel-1} \quad Y = B_0 + B_1*(ESCS) + R \quad [6.3]$$

$$\text{Nivel-2} \quad B_0 = G_{00} + G_{01}*(ESCS \text{ centro}) + U_0 \quad [6.4]$$

$$B_1 = G_{10}$$

4.2. Objetivo 2. Calcular la consistencia de los efectos escolares entre las distintas variables aptitudinales y actitudinales

Para medir la consistencia del impacto escolar en las distintas escalas se elaboraron las diferencias o residuales - entre la puntuación pronosticada por las variables ajuste con la puntuación empírica obtenida en cada prueba- mediante el procedimiento de Mínimos Cuadrados Ordinarios (MCO). La consistencia se calculó a través de la correlación de Pearson entre cada par de residuales de las distintas competencias.

Las variables de ajuste o predictoras son el ESCS de alumnado y de centro y las variables criterios son las distintas medidas cognitivas y no cognitivas de centro, como se puede ver en las siguientes expresiones, el modelo es igual al anterior modelo de efectos netos. Las variables correlacionadas son los residuales los cuales son estimadores empíricos de Bayes (EB).

$$\text{Nivel-1} \quad Y = B_0 + B_1*(ESCS) + R \quad [6.5]$$

$$\text{Nivel-2} \quad B_0 = G_{00} + G_{01}*(ESCS \text{ centro}) + U_0 \quad [6.6]$$

$$B_1 = G_{10}$$

4.3. Objetivo 3. Determinar la eficacia diferencial de los centros

Para analizar la eficacia diferencial que ejerce el centro se comprobó mediante el análisis de significación de la aportación de las variables de ajuste o contexto en la parte aleatoria tanto de forma individual (un modelo multinivel por cada variable predictora) como en el modelo ajustado global con todas las variables, el cual se recoge en la siguiente ecuación.

Las variables de ajuste en esta ocasión configuran un perfil de ser alumna; inmigrante, repetidora dado que son variables dicotómicas, la pendiente de la recta de regresión toma el valor de uno bajo estas condiciones y con un índice socioeconómico medio (-0,16).

Nivel-1

$$Y = B_0 + B_1*(ESCS) + B_2*(FEMALE) + B_3*(NATIVE) + B_4*(SAMELANG) + B_5*(REP) + R \quad [6.7]$$

$$\text{Nivel-2} \quad B_0 = G_{00} + U_0 \quad [6.8]$$

$$B_1 = G_{10} + U_1 \dots \quad [6.9]$$

$$\dots B_5 = G_{50} + U_5 \quad [6.10]$$

4.4. Objetivo 4. Determinar la aportación del contexto y del proceso de enseñanza-aprendizaje en la adquisición de la competencia científica

La estrategia utilizada para cumplir con este objetivo se centró en el uso de un enfoque aditivo en el que, a partir de una especificación básica de partida como corresponde con el modelo nulo, posteriormente, se fueron incorporando sucesivamente diferentes bloques de variables explicativas.

Tanto para el nivel de alumnado como para el nivel de centro se partió de la estimación del modelo nulo y posteriormente se incorporaron los submodelos de contexto y de proceso. Por último se compararon los porcentajes de varianza explicada aportados por los submodelos con respecto al modelo nulo y con respecto a ellos mismos.

Por otra parte se estimaron diferentes modelos de proceso incluyendo la aportación de las variables significativas de contexto/entrada, dado que la literatura recomienda controlar tales efectos. Por último se generó un nuevo modelo, con aquellas variables e índices significativos en los anteriores análisis para obtener un modelo final en cada nivel. En todos los análisis las variables predictoras se consideraron de efectos fijos. Las siguientes expresiones recogen dos ejemplos de los submodelos planteados en forma de ecuaciones de regresión multinivel. El primero recoge índices de proceso asociados con la actitud hacia el medio ambiente y el segundo corresponde a con características de contexto del centro.

1. Medio ambiente

Nivel-1

$$Y = B_0 + B_1*(ENVAWARE) + B_2*(ENVPERC) + B_3*(ENVOPT) + B_4*(RESPDEV) + R \quad [6.1]$$

Nivel-2

$$B0 = G00 + U0 \quad [6.2]$$

$$B1 = G10 \dots B4 = G40$$

2. Características de centro

Nivel-1

$$Y = B0 + R \quad [6.3]$$

Nivel 2

$$B0 = G00 + G01*(ESCS) + G02*(RURAL) + G03*(CITY) + G04*(SCHSIZE) + G05*(SCHSIZ2) + G06*(CLSIZE) + G07*(CLSIZE2) + G08*(PRIV) + G09*(PUBL) + U0 \quad [6.4]$$

4.5. Objetivo 5. Elaborar un modelo de Eficacia Escolar adecuados a la situación social y educativa científica en educación secundaria española

Para llevar a cabo el objetivo de construir un modelo multinivel de acuerdo al contexto y procesos educativos de la competencia científica para educación secundaria se incluyó en un único análisis todas las variables e índices que resultaron significativas en el nivel de alumnado y de centro como efectos fijos.

Tras obtener un nuevo modelo que incluye todos los indicadores con efectos significativos en ambos niveles se comprobaron los supuestos del modelo (Raudenbush, et al, 2004):

1. El error es homocedástico. Se comprobó que la media de los residuales tomara un valor de cero y su varianza fuera constante en cada uno de los dos niveles. Para la comprobación de la varianza constante se dividió la escala de residuales en cuartiles para comprobar si la varianza era similar en cada cuartil.
2. Los errores tienen una distribución normal para que se puedan inferir los resultados de la muestra a la población. Para comprobar dicho supuesto se calculó la prueba Kolgomorov-Smirnov y se elaboró el gráfico QQ-Plot, que representa el ajuste de los residuales a una distribución normal, para cada uno de los niveles. Si ese gráfico muestra que los errores se distribuyen en torno a la recta diagonal, se cumple el supuesto.

5. Resumen de tablas

Tabla 26: <i>Variables contexto/entrada de alumnado</i> . Fuente: elaboración propia.....	147
Tabla 27: <i>Índices de posesiones y contexto del hogar</i> . Fuente: OCDE (2009a).	148
Tabla 28: <i>Análisis de fiabilidad y validez del ESCS</i> . Fuente: elaboración propia.....	149
Tabla 29: <i>Coefficientes de fiabilidad de los índices de proceso de nivel 1</i> . Fuente: elaboración propia.	149
Tabla 30: <i>Análisis Factorial Confirmatorio de los índices de proceso de nivel 1</i> . Fuente: elaboración propia.	150
Tabla 31: <i>Variables de contexto/entrada de nivel 2</i> . Fuente: elaboración propia.	153
Tabla 32: <i>Variables de proceso de nivel 2</i> . Fuente: elaboración propia.	154

6. Resumen de figuras

Figura 24: <i>Representación de la participación de España en PISA 2006</i> . Fuente: IE (2007).	144
Figura 25: <i>Distribución de centros según el tipo de población</i> . Fuente: elaboración propia.	145

CAPÍTULO 7
RESULTADOS

Capítulo 7: Resultados

1.	Magnitud de los efectos de centro	163
2.	Consistencia de los efectos escolares	165
3.	Eficacia diferencial de los centros	166
4.	Aportación del contexto y del proceso de enseñanza-aprendizaje en la adquisición de la competencia científica	168
4.1.	Nivel de alumnado	169
4.2.	Nivel de centro	173
5.	Modelo de eficacia escolar científica	176
5.1.	Comprobación de supuestos	178
6.	Resumen de tablas	180
7.	Resumen de figuras.....	180

1. Magnitud de los efectos de centro

Para estimar el tamaño del efecto de centro se calculó la correlación intraclase del modelo nulo resultando el modelo de efectos brutos de centro. Posteriormente se calculó el modelo de efectos netos que es aquel que incorpora dos variables predictoras relacionadas con el contexto socioeconómico y cultural de las familias y del centro escolar. La figura 26 muestra la variación del primer modelo en cada una de las escalas evaluadas.

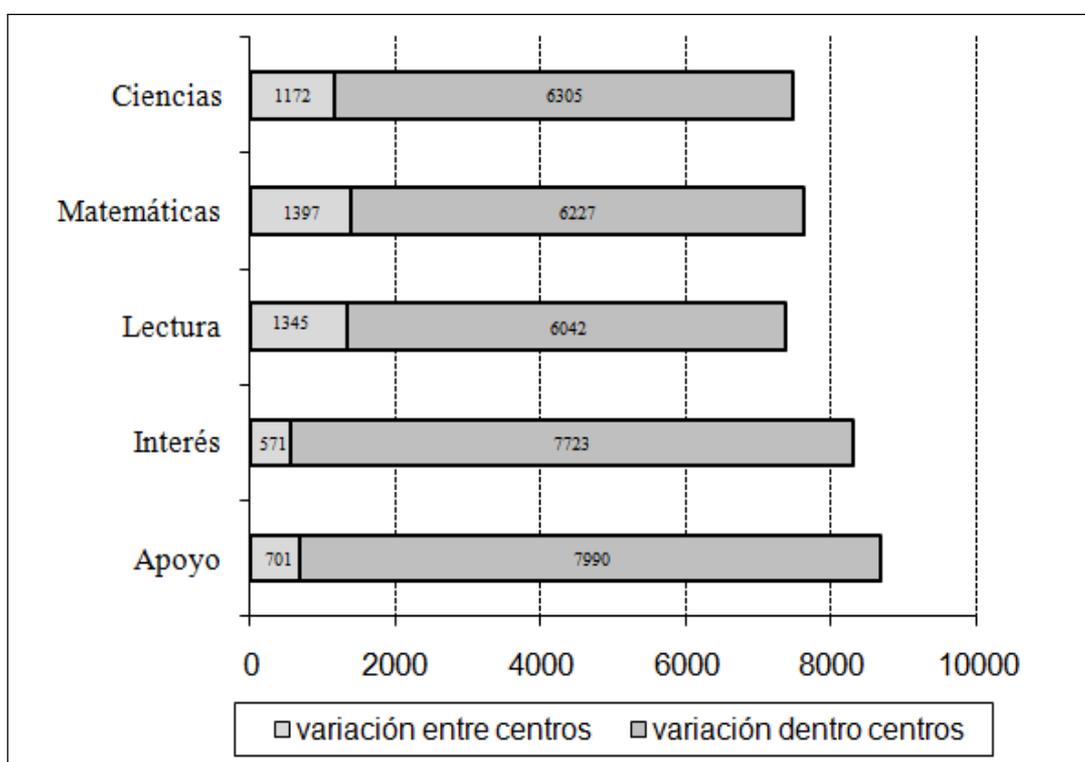


Figura 26: Variación entre y dentro de los centros para el modelo de efectos brutos. Fuente: elaboración propia.

Las variables actitudinales son las que muestran una mayor variabilidad tal y como se representa en la figura 26 dado que la varianza total es mayor que las encontradas en las competencias medidas. Sin embargo dicha cantidad de varianza total no está explicada por variables de tipo de centro ya que como se puede ver en dicha figura o en la tabla 33 es menor el efecto escolar en este tipo de outputs.

Tabla 33: Estimación de los efectos escolares brutos y netos. Fuente: elaboración propia.

Variables	Efectos escolares	
	Brutos	Netos
<i>Pruebas aptitudinales</i>		
Competencia científica	15,67%	11,17%
Competencia matemática	18,33%	11,61%
Competencia lectora	18,21%	13,15%
<i>Pruebas actitudinales</i>		
Interés científico	6,89%	6,56%
Apoyo científico	8,07%	7,96%

El mayor efecto escolar bruto se encuentra en la competencia matemática (18,33%) mientras que la competencia científica es donde se recoge un menor efecto de centro (15,67%). Comparando los efectos escolares netos, una vez eliminada la influencia del ESCS del alumnado y del centro, el impacto de centro se reduce, encontrándose el mayor porcentaje de varianza explicada por el nivel dos en lectura (13,15%). Para las actitudes la influencia del centro es residual, dado que porcentajes de varianza explicada asociada al nivel de centro no llegan a un 10%.

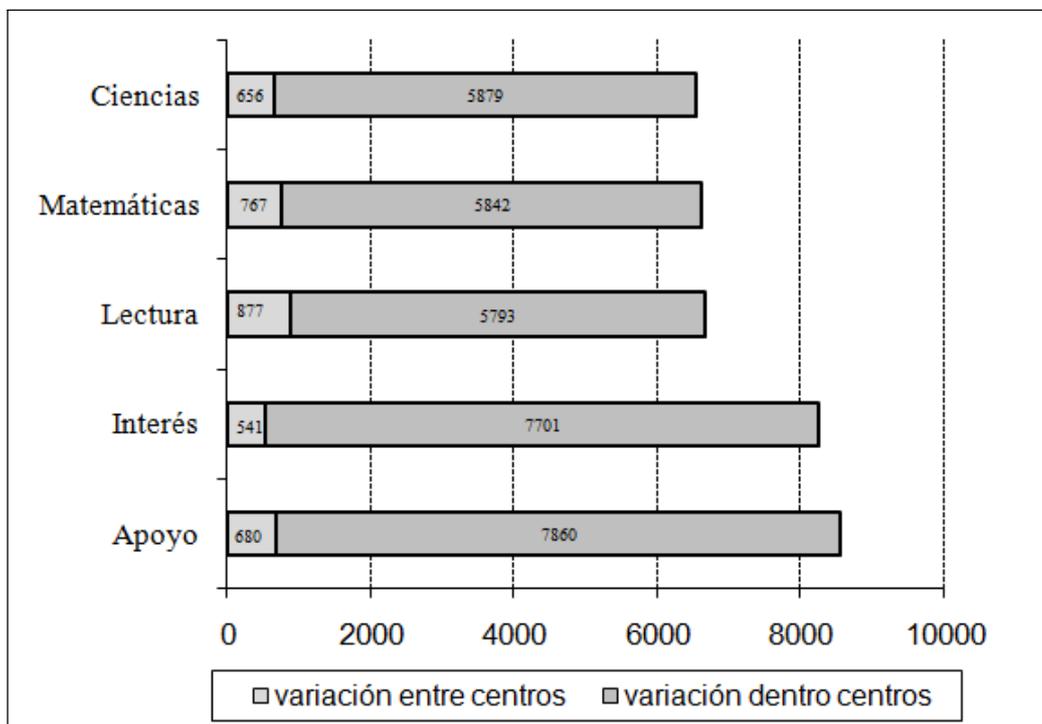


Figura 27: Variación entre y dentro de los centros para el modelo de efectos netos.

Fuente: elaboración propia.

La figura 27 representa la varianza total y la varianza asociada de cada nivel en el modelo de efectos netos de centro. El modelo de efectos netos de centro reduce la

varianza de nivel dos en un 44,03% para ciencias; 45,07% para matemáticas y 34,78% para lectura, lo que indica el alto impacto del ESCS de centro en los resultados, sobre todo en las aptitudes medidas.

La tabla 34 muestra la media y el rango de valores que toma el intercepto con un nivel de significación de 5% en el modelo de efectos netos. Ese valor se encuentra por encima de 500 puntos, media de la OCDE, para todas los modelos con la excepción de lectura (480,80). Destacan las puntuaciones altas en las actitudes, sobre todo el apoyo a la investigación científica (526,64) y el valor que toma el intervalo superior (701,84).

Tabla 34: *Rango de valores plausibles entre los que se sitúa la media de centro.* Fuente: elaboración propia.

	Ciencias	Matemáticas	Lectura	Interés	Apoyo
Media de rendimiento	504,67 (1,54*)	502,17 (1,67*)	480,80 (1,62*)	523,55(1,19*)	526,64 (1,32*)
<i>Rango de valores plausibles¹</i>					
Valor inferior	349,03	347,51	328,45	351,30	351,44
Valor superior	660,30	656,84	633,15	695,80	701,84

Nota: Los valores con asterisco son los errores estándar.

2. Consistencia de los efectos escolares

Para calcular la consistencia de los efectos escolares fue necesario calcular las diferencias o residuales entre el resultado que obtienen los centros y su puntuación pronosticada o, dicho de otra forma, lo que pueden llegar a rendir en función del estatus socioeconómico y cultural, el cual como se ha visto en el objetivo anterior presenta una gran influencia. Concretamente dicha consistencia se comprobó mediante las correlaciones de los residuales EB del nivel de centro como se ve en la tabla 35.

¹ El rango de valores plausibles indica los valores mínimos y máximos entre los que se sitúa la puntuación media del centro con un N.C. del 95% (Raudenbush y Bryk, 2002, p.71). Se construye: $G00 \pm 1,96 (R^{1/2})$

Tabla 35: *Correlaciones entre los residuales de las distintas competencias evaluadas.*

Fuente: elaboración propia.

	Ciencias	Matemáticas	Lectura	Interés	Apoyo
Ciencias	1				
Matemáticas	,762(**)	1			
Lectura	,632(**)	,595(**)	1		
Interés	,186(**)	,082(*)	,042	1	
Apoyo	,344(**)	,217(**)	,203(**)	,546(**)	1

Nota: ** La correlación es significativa al nivel 0,01 (bilateral). * La correlación es significativa al nivel 0,05 (bilateral).

Todas las correlaciones resultaron ser moderadas y en su mayoría significativas. Se encuentra la menor correlación entre las medidas de lectura e interés científico (0,04) y la mayor consistencia entre matemáticas y ciencias (0,76). Es destacable la baja correlación entre las medidas actitudinales y la propia competencia científica: rendimiento en ciencias e interés (0,19) y aptitud científica y apoyo (0,34). Aunque sí se demuestra una consistencia de los efectos de centro entre las tres áreas evaluadas, por tanto centros eficaces en ciencias también lo son en matemáticas y en lectura.

3. Eficacia diferencial de los centros

Para determinar el grado de eficacia de los centros entre diferentes colectivos de alumnado, fue necesario conocer si las diferentes pendientes de las distintas variables predictoras que componen cada recta de regresión son estadísticamente diferentes en cada centro.

Esta modificación implica que cada centro o grupo de nivel dos tome una pendiente propia en las variables predictoras. La tabla 36 muestra las pendientes aleatorias de las variables predictoras que han resultado significativas y por tanto muestran una eficacia diferencial de los centros.

Tabla 36: *Significación de pendientes aleatorias en variables de contexto*. Fuente: elaboración propia.

	Ciencias	Matemáticas	Lectura	Interés	Apoyo
ESCS	No ($p > 0,05$)	Sí ($p < 0,05$)	Sí ($p < 0,05$)	No ($p > 0,05$)	No ($p > 0,05$)
mujer	Sí ($p < 0,05$)	Sí ($p < 0,05$)	Sí ($p < 0,05$)	-	-
Español	Sí ($p < 0,05$)	-			
repetidor	No ($p > 0,05$)	Sí ($p < 0,05$)	No ($p > 0,05$)	-	Sí ($p < 0,05$)
Misma lengua	Sí ($p < 0,05$)	No ($p > 0,05$)	Sí ($p < 0,05$)	-	Sí ($p < 0,05$)

Nota: Los guiones indican que las variables no son significativas en un modelo de efecto fijo.

Las áreas de matemáticas y lectura son las que muestran una mayor eficacia diferencial de los centros en casi todas las características del alumnado evaluadas. De hecho para matemáticas todos los centros no son igualmente eficaces de compensar diferencias de contexto económico y cultural, sexo, nacionalidad y repetición de curso. Con respecto a ciencias ocurre algo similar con la excepción del índice económico y cultural y la condición de repetición dado que todos los centros se comportan de forma similar en estas dos características.

En cuanto a las variables actitudinales no hay diferencias significativas con respecto al sexo en un modelo de efectos fijos. Se encuentran diferencias estadísticamente significativas entre los centros en cuanto al país de origen en el interés por la ciencia. Precisamente en esta escala al alumnado nacido en España se le predicen menos puntos que el alumnado cuyo país de origen, hecho que sólo ocurre en la escala de interés científico. Mientras que con respecto a las actitudes de apoyo científico, la condición de repetición y diferente lengua hablada en casa predicen menos puntos.

La tabla 37 muestra la pendiente promedio en cada una de las variables que resultaron significativas y el intervalo que oscila la pendiente a un nivel de confianza del 95%. Donde se encuentran un mayor grado de variabilidad entre los centros es en las variables de repetición en el apoyo a la investigación científica y ser nativo en las tres competencias. Para lectura destaca la dispersión que muestra la lengua hablada coincide con la lengua de la prueba para lectura.

Tabla 37: *Modelo de pendientes aleatorias*. Fuente: elaboración propia.

	constante	pendiente	S.D.	Int. inferior	Int. superior
<i>Modelo de ciencias</i>	498,74				
mujer		-15,74	9,66	-34,67	3,19
nativo		20,89	23,52	-25,21	66,98
misma lengua		21,88	21,58	-20,42	64,18
<i>Modelo de matemáticas</i>	505,21				
ESCS		10,68	6,43	-1,9228	23,2828
mujer		-22,57	10,03	-42,2288	-2,9112
nativo		23,69	24,88	-25,0748	72,4548
repetidor		-90,45	10,46	-110,9516	-69,9484
<i>Modelo de lectura</i>	450,82				
ESCS		8,51	4,22	0,2388	16,7812
mujer		22,39	11,23	0,3792	44,4008
nativo		12,21	26,72	-40,1612	64,5812
misma lengua		33,61	29,29	-23,7984	91,0184
<i>Modelo de interés</i>	546,17				
nativo		-24,21	26,98	-77,0908	28,6708
<i>Modelo de apoyo</i>	518,34				
repetidor		-21,49	20,66	-61,9836	19,0036
misma lengua		15,601	32,89	-48,8634	80,0654

4. Aportación del contexto y del proceso de enseñanza-aprendizaje en la adquisición de la competencia científica

Para llevar a cabo este objetivo se elaboraron los distintos submodelos en cada nivel y se compararon los porcentajes de varianza explicada en cada uno de ellos. Estos submodelos son los planteados en el anterior capítulo y se agrupan en dos tipos: contexto o proceso. El anexo 4 muestra los estadísticos descriptivos de los indicadores evaluados.

En el nivel de alumnado el modelo de contexto explica el 30,70% de varianza, tabla 38. Los submodelos de proceso también explican porcentajes de varianza similares a dicho porcentaje, incluso el submodelo de proceso compuesto por variables motivacionales es el más explicativo (39,33%) que el submodelo de contexto.

Tabla 38: *Porcentajes de varianza explicada por los submodelos.* Fuente: elaboración propia.

	Modelo Nulo (0)	Modelo contexto (1)	Modelo motivación (2)	Modelo autoeficacia (3)	Carreras científicas (4)	Creencias científicas (5)	Medio ambiente (6)	Modelo proceso E/A (7)
centro	1133,13	690,28	665,89	737,76	795,70	738,01	776,04	751,80
alumnado	6524,37	4521,41	3958,29	4158,05	4575,98	4289,41	4249,19	4331,91
<i>Varianza explicada N1</i>		30,70	39,33	36,27	29,86	34,26	34,87	33,60

El caso contrario se encuentra en el nivel de centro, tabla 39. Los mayores porcentajes de varianza explicada se encuentran en indicadores de contexto. El submodelo de características de centro explica el mayor porcentaje de varianza explicada de centro, 51,67%. No obstante el submodelo que recoge variables de proceso de enseñanza-aprendizaje explica más que submodelos de contexto como admisión y contexto de instrucción (0%) y gestión (9,64%).

Tabla 39: *Porcentajes de varianza explicada por los submodelos de centro.* Fuente: elaboración propia.

	Modelo Nulo (0)	Caracts. Centro(1)	Admisn. y ctxto de instrucc (2)	Gestión del centro (3)	RR del centro (4)	Modelo E/A (5)	Modelo evaluación (6)
centro	1133,13	547,62	1177,69	1023,90	955,29	964,05	1135,54
alumnado	6524,37	6524,94	6302,82	6526,11	6524,93	6523,26	6524,90
<i>Varianza explicada N2</i>		51.67%	0%	9,64%	15,69%	14,92%	0%

Desde el punto de vista teórico los dos tipos de indicadores deben evaluarse conjuntamente para controlar el efecto de las variables de contexto, como se muestra en los siguientes epígrafes.

4.1. Nivel de alumnado

En cuanto a los porcentajes de varianza explicada por los distintos submodelos, todos explican en torno al 36,4 % de la varianza del nivel uno. El modelo menos explicativo es el submodelo compuesto exclusivamente por indicadores de contexto (30,70%) mientras que el submodelo de contexto y medio ambiente (6) es el más explicativo (42,49%).

Tabla 40: Submodelos que componen el nivel 1. Fuente: elaboración propia.

Características	Modelo nulo (0)	Modelo contexto (1)	Modelo motivación (2)	Modelo autoeficacia (3)	Carreras científicas (4)	Creencias científicas (5)	Medio ambiente (6)	Modelo proceso E/A (7)
<i>Intercepto</i>	503,17 (1,60)	503,99 (4,46)	494,81 (4,49)	495,24 (4,36)	503,53 (4,48)	496,60 (4,50)	502,54 (4,17)	508,75 (4,30)
<i>Efectos fijos</i>								
Mujer		-16,28 (1,23)	-16,08 (1,18)	-10,76 (1,17)	-16,67 (1,23)	-12,90 (1,20)	-12,04 (1,16)	-18,9 (1,18)
Nativo		20,86 (3,01)	27,71 (2,99)	25,15 (2,87)	21,62 (3,03)	25,99 (2,98)	17,81 (2,69)	17,75 (2,87)
Repetidor		-84,23 (1,44)	-75,30 (1,45)	-74,77 (1,41)	-82,78 (1,46)	-79,41 (1,45)	-71,22 (1,35)	-78,58 (1,38)
Misma Lengua		21,80 (3,67)	21,06 (3,77)	19,76 (3,60)	21,10 (3,68)	20,56 (3,79)	17,99 (3,38)	18,96 (3,47)
ESCS		12,92 (0,83)	9,99 (0,79)	7,67 (0,76)	12,40 (0,83)	9,63 (0,82)	8,91 (0,71)	12,57 (0,78)
ESCS ²		-1,28 (0,61)	-1,34 (0,60)	-1,25 (0,57)	-1,28 (0,60)	-1,24 (0,6)	-0,84* (0,51)	-1,60 (0,59)
SCIEPADS		5,07 (2,00)	4,58 (1,89)	5,11 (1,92)	4,92 (1,99)	5,91 (1,97)	-0,25 (0,047)	4,81 (1,9)
INTSCIE			6,03 (0,83)					
JOYSCIE			16,25 (0,85)					
INSTSCIE			-1,25* (0,76)					
SCIEFUT			3,42 (0,95)					
SCSCIE				11,65 (0,65)				
SCIEEFF				18,09 (0,58)				
CARINFO					0,94* (0,70)			
CARPREP					7,37 (0,66)			
GENSCIE						11,17 (0,81)		
PERSCIE						5,59 (0,82)		
SCIEACT						6,54 (0,74)		
ENVAWARE							24,43 (0,70)	
ENVPERC							-3,29 (0,74)	
ENVOPT							-8,53 (0,61)	
RESPDEV							10,79 (0,71)	
SCINTACT								-4,21 (0,84)
SCHANDS								0,59* (0,84)
SCINVEST								-20,67 (0,88)
SCAPPLY								14,11 (0,86)
<i>Efectos aleatorios (SD)</i>								
Estudiantes	80,77	67,24	63,98	62,60	66,88	65,01	61,25	64,99
Centros	33,66	26,27	25,13	24,37	25,79	25,27	30,03	24,23
<i>Varianza explicada</i>								
		30,70%	37,26%	39,93%	31,44%	35,23%	42,49%	35,25%
Nivel 1	6524,37	4521,41	4093,54	3919,11	4472,93	4225,90	3751,98	4224,58
Nivel 2	1133,13	690,28	631,64	594,09	665,17	638,53	901,91	587,22

Nota: Los índices con asteriscos son no significativos ($p > 0,05$)

Al introducir variables de contexto algunos indicadores de proceso dejaron de ser significativos como: CARINFO que indica el grado de información que sobre carreras científicas; SCHANDS es la valoración de experimentos en el laboratorio. Tampoco resultó significativo INSTSCIE que se refiere a la motivación instrumental o la utilidad que encuentran en el estudio de ciencias, incluso presenta un efecto negativo que contradice el marco teórico.

Tabla 41: *Modelo final de nivel 1*. Fuente: elaboración propia.

Características	Modelo nulo (0)	Modelo final: alumnado
	<i>Intercepto</i>	503,17 (1,60)
	<i>Efectos fijos</i>	499,29 (4,07)
Modelo de contexto (1)	Mujer	-11,74 (1,09)
	Nativo	20,75 (2,61)
	Repetidor	-62,01 (1,30)
	MismaLengua	15,74 (3,29)
	ESCS	5,13 (0,71)
	ESCS ²	-1,45 (0,53)
	SCIEPADS	4,32 (1,74)
Modelo motivación (2)	INTSCIE	1,05 (0,78)*
	JOYSCIE	9,48 (0,83)
	SCIEFUT	3,14 (0,80)
Modelo autoeficacia (3)	SCSCIE	7,40 (0,71)
	SCIEEFF	8,89 (0,61)
	CARPREP	-1,97 (0,61)*
Modelo de creencias científicas (4)	GENSCIE	5,67 (0,74)
	PERSCIE	-4,43 (0,75)
	SCIEACT	-3,29 (0,76)
Modelo medio ambiente (5)	ENVAWARE	15,71 (0,75)
	ENVPERC	-3,90 (0,71)
	ENVOPT	-7,72 (0,59)
	RESPDEV	7,07 (0,74)
Modelo de enseñanza/aprendizaje (6)	SCINTACT	-4,31 (0,79)
	SCINVEST	-15,68 (0,75)
	SCAPPLY	5,46 (0,76)
<i>Efectos aleatorios (SD)</i>		
Estudiantes	80,77	56,92
Centros	33,66	30,71
<i>Varianza explicada</i>		50,34%
Nivel 1	6524,37	3239,92
Nivel 2	1133,13	943,11

Nota: Los índices con asteriscos son no significativos ($p > 0,05$)

El modelo final de alumnado donde se recoge en un único análisis todas las variables que resultaron significativas se muestra en la tabla 41. Se pronostican 499,29

puntos al alumnado que toma un valor de cero en las variables dummies y obtienen una puntuación media en los índices contruidos. Todas las variables de contexto son significativas. Se encuentra que la característica de ser nativo, realizar la prueba en la misma lengua que habla en casa son predictores de una mayor puntuación en ciencias. Mientras que a los estudiantes repetidores y ser alumna se les pronostican menos puntos.

El ESCS predice un incremento de 5,23 puntos en la competencia a medida que aumenta una desviación estándar dicho indicador (1,02). Este último índice socioeconómico también se consideró como predictor cuadrático, el cual toma un valor negativo lo que indica que si se representa dicho gradiente gráficamente tomaría una curvatura convexa (extremos hacia abajo), por lo que las diferencias educativas tienden a disminuir a medida que el ESCS progresa.

Entre los índices de proceso pierden significación para predecir éxito en ciencias la preparación que tiene el alumnado sobre información de carreras científicas (CARPREP) y el interés por el estudio de esta área (INTSCIE). Los predictores que más puntuación predicen tienen que ver con la motivación, la autoeficacia y el cuidado al medio ambiente. El entusiasmo para estudiar temas científicos (JOYSCIE), hace predecir un aumento de 9,48 puntos en el rendimiento por cada desviación típica (1,00) que aumenta JOYSCIE. La percepción de eficacia (SCIEEFF) incrementa el rendimiento 9,51 puntos por aumento en una desviación (1,07) y de forma similar ocurre con el autoconcepto (SCSCIE). SCSCIE predice 7,55 puntos en ciencias por aumento en una desviación (1,02).

El submodelo de medio ambiente presenta valores significativos en todos sus indicadores. Disponer de concienciación sobre problemas medioambientales (ENVAWARE) predice un aumento en el rendimiento de 14,92 por aumento en una desviación (0,95). De forma similar ocurre con el acuerdo sobre un desarrollo sostenible (RESPDEV), a medida que aumenta una desviación (0,91), se predicen 6,43 puntos más en ciencias.

Por último el modelo 6 de enseñanza-aprendizaje presenta efectos significativos pero en algún indicador el impacto es contradictorio. Se encuentra el mayor efecto positivo en la percepción que tiene el alumnado en percibir la aplicación de las ciencias en otros campos (SCIEAPPLY), de hecho el aumento en una desviación (0,98) de

SCIEAPPLY la competencia científica aumenta 5,35 puntos. Sin embargo a mayor iniciativa investigadora por parte del alumnado hace predecir un efecto negativo en el rendimiento. Concretamente se pronostica una disminución de 15,37 puntos en el rendimiento a medida que aumenta la posibilidad de iniciar un experimento por parte del alumnado, la desviación de dicho índice es (0,98).

4.2. Nivel de centro

En el nivel dos también se han elaborado incluyendo indicadores de contexto y proceso, tabla 42.

El submodelo 2 de admisión y contexto de la instrucción no presenta significación en ninguna de los indicadores y la varianza explicada es nula. El submodelo 1 de características de centro es el que explica más cantidad de varianza del nivel dos, en torno al 52%. Concretamente las características de centro con más poder predictivo son el índice socioeconómico y cultural medio de centro (ESCS) y el tamaño de la clase (CLSIZE). El ESCS medio de centro predice un aumento del rendimiento en ciencias de 24,30 puntos por desviación típica (0,55). Mientras que el tamaño de clase hace predecir 12, 52 puntos menos por incremento en una desviación (0,77).

El submodelo 3 de gestión de centro explica un 9,64% de la varianza de centro y la variable con efecto significativo es la autonomía del centro en la gestión de personal (FACS). El submodelo 4 de recursos de centro explica en torno al 15,69% de la varianza de centro, la variable significativa que más efecto presenta es el ratio de alumnos- profesor que tiene el centro, se predicen 12,25 puntos más a medida que aumenta una desviación (4,36).

Por lo que respecta a los submodelos, que analizan conjuntamente el contexto y proceso, son los que más porcentaje de varianza aportan gracias a la incursión de índices de contexto. De hecho el mayor aporte procede del predictor ESCS de centro. Sin embargo otros índices de contexto cuando se evalúan conjuntamente con indicadores proceso pierden su significación como el tamaño de clase; la condición de ruralidad; autonomía de la gestión de personal y ratio profesor - alumno. Ninguno de los predictores del modelo de enseñanza-aprendizaje ni de evaluación resultaron significativos con la excepción de tiempo dedicado al aprendizaje medio que hace aumentar la competencia científica en 7,64 puntos por desviación típica (1,29).

Tabla 42: *Submodelos que componen el nivel 2.* Fuente: elaboración propia.

Características	Modelo nulo	Caracts. Centro (1)	Admisn. y cxto de instrucc (2)	Gestión del centro (3)	RR del centro (4)	Modelo E/A (5)	Modelo evaluación (6)
<i>Intercepto</i>	503,17 (1,60)	501,07 (7,35)	503,88 (4,28)	504,01 (1,55)	503,22 (1,50)	501,57 (1,57)	502,73 (2,93)
<i>Efectos fijos</i>							
RURAL		8,44 (3,45)				4,99 (2,84)*	4,81 (2,90)*
URBANO		3,69 (3,10)*					
PRIV		0,87 (2,98)*					
PUBL		0,13 (0,07)*					
SCHOSIZE		2,73(1,21)				1,13 (0,42)	1,06 (0,43)
SCHOSIZE ²		-0,09(0,06)*					
CLSIZE		-16,26(7,74)				-1,78 (1,28)*	-1,98(1,48)*
CLSIZE ²		1,65 (0,88)*					
ESCS CENTR		44,19 (3,32)				39,37 (2,99)	42,53 (2,98)
LOSELE			0,77 (4,27)*				
HISELE			-8,25 (18,30)*				
ABGROUP			-4,14 (4,34)*				
PRESSPA			5,34 (3,19)*				
FACS				9,59 (1,66)		-0,37 (2,47)*	-0,37(2,41)*
FACB				1,22 (1,66)*			
FACC				0,67 (1,73)*			
STRATIO					2,81 (0,37)	-0,29(0,64)*	0,07(0,62)*
TCSHORT					-4,77 (2,02)	-3,08 (1,65)*	-4,01 (1,66)
IRATCOM					-20,6(20,53)*		
SCMATED					2,68 (1,57)*		
ENVLEARN						-0,65 (1,23)*	
SCIPROM						1,63 (1,44)*	
ANYSCIE						0,09 (0,14)*	
TIMLEARN						5,93 (1,12)	
ACC1							3,07 (2,59)*
ACC2							1,73 (4,76)*
ACC3							-3,53(5,34)*
ACC4							-0,95(3,68)*
ACC5							-4,87(4,26)*
ACC6							-4,53(2,65)*
ACC7							0,06 (2,57)*
ACC8							1,16 (2,65)*
<i>Efectos aleatorios (SD)</i>							
Estudiantes	80,77	80,78	79,39	80,78	80,78	80,77	80,78
Centros	33,66	23,40	34,32	31,99	30,91	22,47	23,46
<i>Varianza explicada</i>		51.67%	0%	9,64%	15,69%	55,45%	51,42%
Nivel 1	6524,37	6524,94	6302,82	6526,11	6524,93	6525,36	6525,73
Nivel 2	1133,13	547,62	1177,69	1023,90	955,29	504,83	550,49

Nota: Los índices con asteriscos son no significativos ($p > 0,05$)

Los siguientes resultados muestran los efectos de indicadores de contexto y proceso que resultaron significativas en el modelo que, tabla 43. Algunos como la condición de rural, el tamaño de clase, la autonomía del centro en la gestión de su personal y el ratio profesor-alumno no se incluyen porque dejan de ser significativos cuando se introducen todas las variables de forma simultáneamente.

Tabla 43: *Modelo final de nivel 2*. Fuente: elaboración propia.

Características	Modelo nulo (0)	Modelo final N2: centro
<i>Intercepto</i>	503,17 (1,60)	503,42 (1,20)
<i>Efectos fijos</i>		
	SCHOSIZE	0,82 (0,36)
	ESCS CENTR	37,86 (2,62)
	TCSHORT	-3,28 (1,67)
	TIMLEARN	5,77 (1,09)
<i>Efectos aleatorios (SD)</i>		
Estudiantes	80,77	80,78
Centros	33,66	22,49
<i>Varianza explicada</i>		55,35%
Nivel 1	6524,37	6525,63
Nivel 2	1133,13	505,94

Nota: Los índices con asteriscos son no significativos ($p > 0,05$)

Se predicen 503,42 puntos en ciencias para aquellos centros que tienen un valor medio en los predictores. El porcentaje de varianza explicado de nivel macro es del 55,35% y todas las variables resultaron significativas. El índice socioeconómico y cultural medio del centro es el predictor que mayor efecto presenta. ESCS predice un aumento del rendimiento en ciencias de 20,82 puntos por desviación típica (0,55) que aumenta el contexto socioeconómico y cultural. Después del estatus económico el indicador que más predice es el tiempo medio dedicado al aprendizaje hace aumentar la competencia científica en 7,44 puntos por desviación típica (1,29). Por último el tamaño del centro y la falta de profesorado muestran un efecto significativo pero menor con respecto a los dos indicadores anteriores. El tamaño de centro hace predecir 3,23 puntos más a medida que aumenta una desviación (3,94) y con respecto a la percepción de la escasez de profesorado se da una relación negativa, de hecho predice una reducción del rendimiento de 2,36 puntos por aumento en una desviación (0,72) en dicho índice.

El mayor número de variables que resultaron significativas corresponden al nivel de alumnado. Sin embargo en términos de varianza explicada los índices incluidos en el nivel de centro explican un 64,22% de la varianza del nivel dos mientras que el porcentaje de varianza explicada por el nivel uno es de 51,09%.

El intercepto calculado es de 496,85 puntos. En cuanto al nivel de alumnado las características de contexto evaluadas destacan los efectos positivos de ser nativo y que la prueba se desarrolle en la lengua hablada en el hogar. Por el contrario estudiantes repetidores y ser alumna se les pronostican menos puntos. Existen también índices como el contexto socioeconómico o que algunos de los progenitores se dediquen a una profesión científica con menor poder predictivo pero asociado al rendimiento de forma positiva. El ESCS cuadrático toma un valor negativo por lo que si se representa un gradiente gráficamente tomaría una curvatura convexa que indica que las diferencias educativas tienden a disminuir a medida que el ESCS progresa.

En este mismo nivel destaca el compromiso con el medio ambiente (ENVAWARE) que hace pronosticar 15,24 puntos más por desviación que aumenta (0,95) o el entretenimiento en el estudio de ciencias (JOYSCIE) que hace pronosticar 10,46 puntos más por desviación (1,00). No obstante existen otras variables que presentan un efecto negativo que es esperable como aquellos que creen que los problemas medio ambientales mejorarán (ENVOPT) se les predice una menor puntuación, en torno a 7,39 puntos menos por cada desviación que aumenta dicha consideración (0,99). Sin embargo otros como la iniciativa investigadora de los alumnos (SCINVEST) que hace que se pronostiquen 15,93 puntos menos en la competencia científica por cada desviación que aumenta dicho índice (0,98).

En el nivel de centro el mayor efecto se da en el índice socioeconómico del centro (ESCS centro) el cual hace predecir 4,28 puntos más en el rendimiento por desviación típica (0,55). De forma menos acusada están las horas de tiempo de enseñanza-aprendizaje que invierte el alumnado, concretamente se pronostican 4,55 puntos más a medida que aumentan las horas de estudio en una desviación (1,29).

5.1. Comprobación de supuestos

Una vez elaborado el modelo final se comprobaron los supuestos del modelo: errores homocedásticos y los errores presentan una distribución normal en cada nivel medido.

Los resultados indican que los errores son aproximadamente homocedásticos en cada nivel dado que la media es nula y la varianza es similar entre los valores medios y, también, presentan varianzas semejantes los valores extremos, como se ve en la tabla 45.

Tabla 45: *Descriptivos de los términos de error por niveles.* Fuente: elaboración propia.

	Media	Varianza
Nivel alumnado	0,000	
Inferior al 25%		879,540
Entre el 25% y 50%		114,100
Entre el 50% y 75%		111,660
Superior a 75%		722,761
Nivel centro	0,000	
Inferior al 25%		109,049
Entre el 25% y 50%		11,517
Entre el 50% y 75%		11,791
Superior a 75%		82,179

La prueba de Kolgomorov-Smirnov indica que los errores se distribuyen de forma normal en los dos niveles. Dado que en los errores del nivel de alumnado no presentan valores significativos en dicha prueba ($p= 0,186$) y tampoco en el nivel de centro ($p= 0,518$). Además se presentan dos gráficos Q-Q en donde los errores siguen se distribuyen en torno a la recta diagonal como se pueden ver en las figuras 28 y 29.

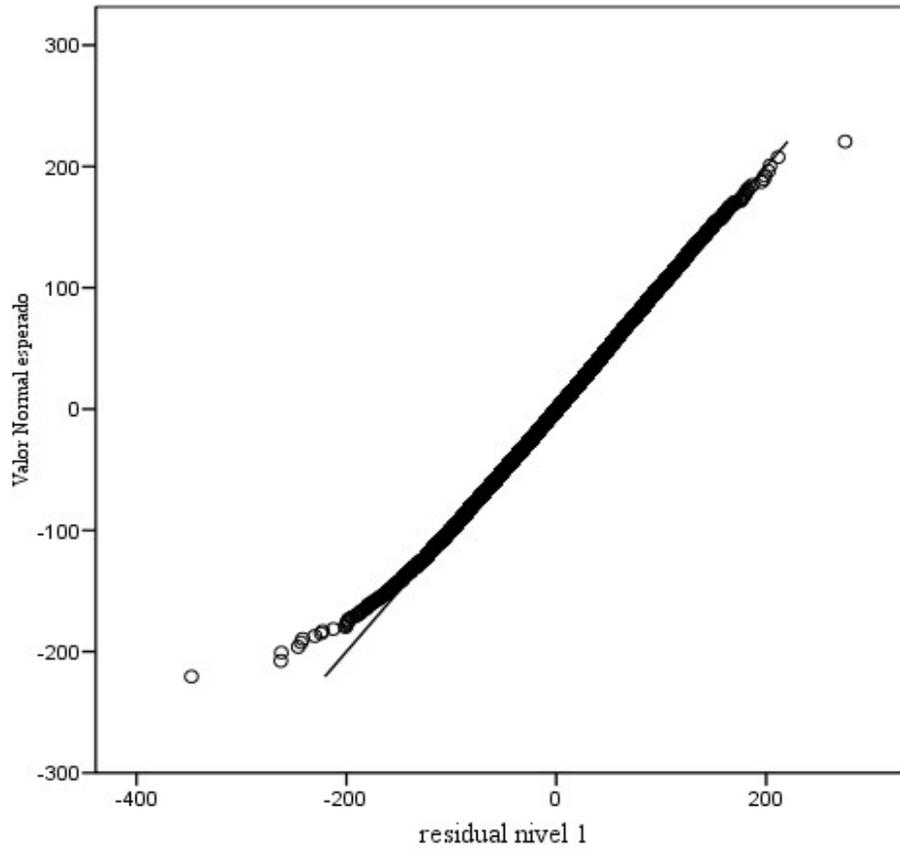


Figura 28: *Distribución de los errores de nivel de alumnado.* Fuente: elaboración propia.

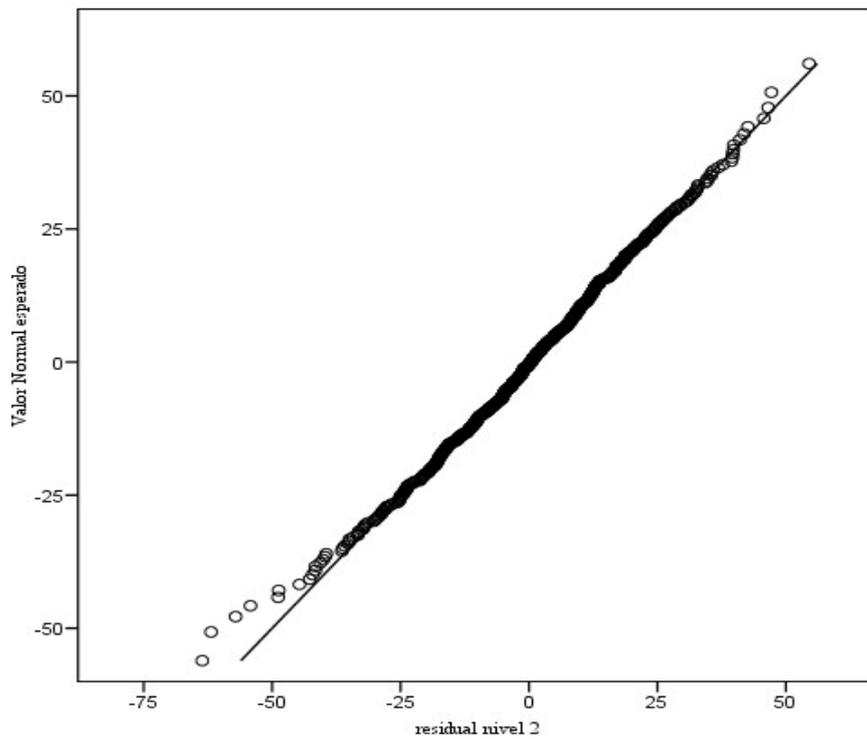


Figura 29: *Distribución de los errores de nivel de centro.* Fuente: elaboración propia.

6. Resumen de tablas

Tabla 33: <i>Estimación de los efectos escolares brutos y netos.</i> Fuente: elaboración propia.....	163
Tabla 34: <i>Rango de valores plausibles entre los que se sitúa la media de centro.</i> Fuente: elaboración propia.	165
Tabla 35: <i>Correlaciones entre los residuales de las distintas competencias evaluadas.</i> Fuente: elaboración propia.	166
Tabla 36: <i>Significación de pendientes aleatorias en variables de contexto.</i> Fuente: elaboración propia.	167
Tabla 37: <i>Modelo de pendientes aleatorias.</i> Fuente: elaboración propia.	168
Tabla 38: <i>Porcentajes de varianza explicada por los submodelos.</i> Fuente: elaboración propia.....	169
Tabla 39: <i>Porcentajes de varianza explicada por los submodelos de centro.</i> Fuente: elaboración propia.	169
Tabla 40: <i>Submodelos que componen el nivel 1.</i> Fuente: elaboración propia.....	170
Tabla 41: <i>Modelo final de nivel 1.</i> Fuente: elaboración propia.	171
Tabla 42: <i>Submodelos que componen el nivel 2.</i> Fuente: elaboración propia.....	174
Tabla 43: <i>Modelo final de nivel 2.</i> Fuente: elaboración propia.	175
Tabla 44: <i>Modelo multinivel final.</i> Fuente: elaboración propia.	176
Tabla 45: <i>Descriptivos de los términos de error por niveles.</i> Fuente: elaboración propia.	178

7. Resumen de figuras

Figura 26: <i>Variación entre y dentro de los centros para el modelo de efectos brutos.</i> Fuente: elaboración propia.	163
Figura 27: <i>Variación entre y dentro de los centros para el modelo de efectos netos.</i> Fuente: elaboración propia.	164
Figura 28: <i>Distribución de los errores de nivel de alumnado.</i> Fuente: elaboración propia.....	179
Figura 29: <i>Distribución de los errores de nivel de centro.</i> Fuente: elaboración propia.	179

CAPÍTULO 8
DISCUSIÓN

Capítulo 8: Discusión

1.	Discusión y conclusiones.....	183
----	-------------------------------	-----

1. Discusión y conclusiones

Conocer la influencia de los centros educativos en el proceso de aprendizaje de aptitudes y de actitudes ha sido la finalidad de este trabajo. Por lo que determinar las propiedades de los efectos escolares (magnitud, consistencia y eficacia diferencial) en tres competencias básicas (ciencias, matemáticas y lectura) y en dos actitudes (interés y apoyo a la investigación científica); y elaborar un modelo comprensivo y multinivel sobre la competencia científica en población española fueron los objetivos desarrollados a lo largo del anterior capítulo para llegar al propósito señalado.

En general los datos obtenidos apoyan las hipótesis planteadas. Los efectos de los centros escolares encontrados en la adquisición de competencias básicas son los que se planteaban en la primera hipótesis y son muy similares a los obtenidos en diversos estudios como Pajares-Box (2005) o Marchesi y Martínez-Arias (2006c) y en la recopilación de 26 investigaciones sobre la magnitud de los efectos escolares (Teddlie, Reynolds y Sammons, 2000). Sin embargo una vez que se incorporan en los análisis multinivel, variables de socioeconómicas del alumnado y del centro, con una gran influencia sobre el rendimiento como se ha recogido en la literatura de la IEE, los mayores efectos escolares se encuentran en la competencia lectora. Resultado que contradice al estudio de Mandeville y Anderson (1987), los cuales encontraron que los estudiantes adquirirían prácticamente todo el conocimiento de lengua en el entorno familiar. No obstante dicho resultado es similar al encontrado por Rodríguez-Jiménez y Murillo (2011). Estos últimos autores también encuentran un mayor efecto de centro en lectura una vez controlada la influencia del estatus socioeconómico, por lo que la lectura es una competencia que no está tan determinada por la condición económica.

Por lo que respecta a las variables actitudinales a pesar que son los outputs que muestran una mayor variabilidad, ésta se atribuye en mayor medida a indicadores de alumnado y no tanto a variables de centro como se planteaba en la primera hipótesis y como ocurre en otros trabajos (Fitz-Gibbon, 1991 y Murillo, 2007).

La consistencia de los efectos escolares obtenida entre las diferentes medidas valida parcialmente la segunda hipótesis. Tal y como se planteaba los centros escolares presentan una mayor consistencia de los efectos entre las competencias que entre las actitudes, pero se esperaba una mayor correlación. Los resultados encontrados son similares a la propuesta por Thomas, Sammons y Mortimore (1997). Estos autores

encuentran una congruencia baja y moderada. Entre los datos obtenidos destacan las relaciones moderadas entre las distintas competencias medidas y relaciones particularmente bajas entre la competencia científica y las actitudes que subyacen a tal competencia. Este último resultado no confirma el planteamiento teórico propuesto por la OCDE (2006b). Se podría concluir como otros autores (Knuver y Brandsma, 1993; Mortimore et al., 1988 y Murillo, 2004 y 2007) que los efectos escolares medidos como *outputs* académicos son independientes de los estimados en *outputs* actitudinales, aunque es necesario profundizar en el estudio de esta propiedad. De hecho son prácticamente inexistentes los trabajos que estudian la consistencia de los efectos escolares en población española de secundaria a través de pruebas actitudinales estandarizadas.

Las características de entrada capaces de incidir en el logro académico que se han encontrado en la historia de la eficacia escolar de forma unánime son el rendimiento previo y las condiciones socioeconómicas (Levine y Lezotte, 1990). PISA no aporta una medida de rendimiento previo por lo que se incluyó en este trabajo la condición de repetidor, la cual es un predictor de fracaso escolar y un estigma que marca el futuro recorrido escolar (Gomolla y Radtke, 2002; Jimerson, Anderson y Whipple, 2002). Otras variables de contexto tenidas en cuenta fueron el sexo y el estatus de inmigrante, que resultaron significativas en trabajos como Pajares-Box (2005); Marchesi-Martínez-Arias (2006); Murillo (2008 y 2011).

La tercera hipótesis planteaba que no todos los centros se comportan de manera similar atendiendo a las anteriores características, los resultados sustentan dicha idea. Las áreas de matemáticas y lectura son las que muestran una mayor eficacia diferencial de los centros en casi todas las características del alumnado evaluadas. Con respecto a ciencias ocurre algo similar con la excepción del índice económico y cultural y la condición de repetición. Todos los centros se comportan de forma similar en estas dos características. En cuanto a la enseñanza de actitudes científicas los centros no son igual de eficaces para fomentar actitudes científicas. Concretamente no todos los centros se comportan de forma similar para generar un interés científico teniendo en cuenta el origen de su alumnado. Algo semejante ocurre con actitudes de apoyo científico, ya que los centros no compensan las diferencias atribuidas a la condición de repetición y al hecho de que sus estudiantes hablen en casa una lengua distinta a la prueba.

Los dos últimos objetivos se centraron en construir el modelo multinivel adecuado a las características del entorno y al proceso de enseñanza-aprendizaje de la competencia científica, de estos dos objetivos se desprenden dos hipótesis que en general se demuestran. Analizando de forma aislada cada uno de los modelos propuestos destaca en el nivel de alumnado el poder predictivo de las variables de contexto, como se planteaba en la cuarta hipótesis y tal como se lleva encontrando desde los resultados de Coleman (1966). Sin embargo en el nivel de alumnado se han encontrado modelos de variables de proceso que presentan porcentajes de varianza explicada similares o superiores al atribuido a variables de contexto. Tal es el caso de la importancia de variables motivacionales como se recoge en estudios como Creemers (1994) o en los resultados de la OCDE (2010). En el nivel de centro las características de centro explican el mayor porcentaje de varianza explicada, sobre todo el índice socioeconómico y cultural del centro. Los recursos del centro, concretamente la falta de profesorado es un predictor para obtener una baja puntuación en ciencias. La relevancia de estos dos indicadores señala la importancia del contexto socioeconómico como indicó Murillo (2011) y la importancia de los recursos educativos como recogen Grisay (1996); Marchesi y Martínez-Arias (2006c) y Murillo (2007). Por tanto la hipótesis cuarta que recoge la importancia del contexto se sustenta claramente en el nivel de dirección.

Desde el punto de vista teórico y tras conocer los resultados encontrados en los anteriores objetivos propuestos, el contexto debe controlarse cuando se evalúan modelos de proceso (Brookver et al., 1979). De hecho diversos índices dejan de ser significativos por efectos de colinealidad como se planteaba en la cuarta hipótesis. El grado de información que dispone el alumnado sobre carreras científicas, la valoración de hacer experimentos en el laboratorio y la utilidad en el estudio de ciencias pierden su significación en el nivel de alumnado. Este hecho también ocurre en el nivel de centro dado que todos los índices pierden su significación con la excepción del tamaño del centro, el estatus socioeconómico del centro, la escasez de profesorado y el tiempo dedicado a la enseñanza-aprendizaje.

En el modelo final los indicadores incluidos explican un elevado porcentaje de varianza estimada en cada nivel. El intercepto estimado en población española se sitúa en valores cercanos a la media de la OCDE. En general los indicadores con mayor poder predictivo son los que se plantean en la quinta hipótesis. La excepción es el bajo

impacto que ejerce en la competencia científica las variables de dirección que tienen que ver con procesos de enseñanza-aprendizaje y evaluación.

En dicho modelo todas las variables de contexto medidas en el nivel de alumnado fueron significativas. Se encuentra que la característica de ser nativo predice y realizar la prueba en la misma lengua que se habla en casa hace aumentar la puntuación en la prueba. Mientras que ser estudiante repetidor hace pronosticar menos puntos con respecto al alumnado matriculado en su curso modal; ser alumna hace pronosticar menos que ser alumno. El índice socioeconómico familiar predice un incremento en la competencia a medida que aumenta dicho indicador, aunque las diferencias en valores extremos son menores. Efecto que también lo obtiene Pajares-Box (2005) en la competencia lectora.

Entre los índices de proceso destaca el entusiasmo por estudiar temas científicos; la percepción de eficacia y el autoconcepto como variables predictoras del éxito en el rendimiento científico, como se encuentra en Martínez-Arias y Marchesi, 2006; Pajares-Box, 2005). Asimismo haber desarrollado una concienciación sobre problemas medioambientales y sobre la necesidad de un desarrollo sostenible son actitudes que favorecen un éxito en el estudio de ciencias, como plantea la OCDE (2006b) y OCDE (2010). Por último el modelo de enseñanza-aprendizaje presenta efectos significativos pero en algún indicador el impacto es contradictorio. El mayor efecto se encuentra en la percepción que tiene el alumnado sobre la aplicación de las ciencias en otros campos del conocimiento, sin embargo a mayor iniciativa investigadora que propone el centro se predice un efecto negativo en el rendimiento. Resultados de este tipo hace plantearse si los estudiantes tienen dificultades para entender lo que se les pregunta en los cuestionarios de contexto tal y como plantean Hopfenbeck y Maul (2011). En el nivel de centro los índices que predicen un mayor rendimiento son el estatus socioeconómico y cultural del centro y el tiempo medio dedicado al aprendizaje. El tiempo que dedica el centro al aprendizaje se ha encontrado en otros estudios como Rutter (1979) y Stringfield y Teddlie (1989). Una vez elaborado el modelo final se comprobó que los errores eran aproximadamente homocedásticos y presentaban una distribución normal de los mismos en cada nivel medido.

El trabajo expuesto es una explotación de los datos que proceden del programa PISA. Este hecho implica ventajas como la posibilidad de disponer de una muestra representativa de la población española de educación secundaria en un momento

próximo a finalizar estudios obligatorios y seleccionada mediante un cuidadoso proceso de muestreo aleatorio; disponer de datos procedentes de instrumentos de evaluación fiables y validados en diferentes contextos culturales; obtener una serie de medidas de desempeño procedentes de un enfoque novedoso de evaluación como es el enfoque de las competencias y analizarlo de una forma contextual a través de una serie de indicadores de diverso tipo.

Sin embargo este tipo de explotación de datos conlleva una serie de limitaciones. En primer lugar la falta de un diseño multinivel y longitudinal con capacidad de estimar la estabilidad de los efectos encontrados, comparar distintas cohortes de estudiantes en conocimiento científico y analizar las propiedades de los efectos escolares. La segunda limitación es no disponer de más variables relacionadas con el rendimiento como, por ejemplo, disponer de una medida de rendimiento previo obtenido en otro tipo de evaluaciones internas como notas académicas o procedentes de otras evaluaciones externas al centro como evaluaciones de diagnóstico. En este sentido también destaca la falta de un nivel de grupo-aula o departamental medido a través del profesorado, nivel que como se ha visto en la literatura es donde se encuentra gran parte de los efectos escolares (Creemers y Kyriakides, 2006 y 2008; Sammons, Thomas y Mortimore, 1997). No obstante la OCDE no obvia la importancia de dicho nivel y ha comenzado a desarrollar nuevos programas internacionales donde se evalúa el comportamiento docente de forma exclusiva mediante el programa TALIS (IE, 2010d).

Como líneas de trabajo se plantea la elaboración de un modelo multinivel longitudinal que compare las dos cohortes de alumnado medidas en dos aplicaciones PISA como actualmente se puede realizar sobre la competencia lectora, dado que la evaluación de esta área por parte de la OCDE se ha completado. Por otra parte se plantea desarrollar una evaluación interna entre los centros para que ellos mismos se midan en indicadores de eficacia y mejora entre los distintos agentes que componen su comunidad educativa. Esta autoevaluación tendría un gran valor estratégico porque supone ser un complemento a evaluaciones externas, validando experiencias de eficacia y creando nuevas posibilidades de factores de eficacia. Además permitiría proporcionar experiencias y conocimientos útiles en proyectos de mejora adaptados al propio centro.

A modo de síntesis y haciendo referencia al estudio por el que se inició el movimiento de la Eficacia Escolar se puede afirmar que *“el centro sí importa”*. Esto es, los centros educativos en España juegan un papel fundamental en los procesos de

enseñanza-aprendizaje entre el alumnado de quince años próximo a finalizar estudios obligatorios. Este trabajo demuestra la importancia del sistema educativo a través de una serie de tests de desempeño que corresponden a la evaluación externa PISA guiada por la OCDE y en un marco teórico que no corresponde con el currículo nacional.

Haciendo uso de los conocimientos provenientes de la línea de investigación de la eficacia escolar se analizaron las propiedades de magnitud, consistencia y eficacia diferencial de los centros. Los resultados demuestran el efecto del centro en el desarrollo de las competencias medidas como fueron ciencias, matemáticas y lectura. Mientras que dicho efecto es menor en el desarrollo de actitudes, concretamente aquellas relacionadas con las ciencias. Además el efecto del centro se mantiene o es consistente entre las distintas competencias medidas, por lo que centros eficaces en ciencias también lo son en matemáticas y en lectura. No obstante ese efecto de los centros en el rendimiento de su alumnado no es igual, sobre todo cuando se trata de rebajar diferencias atribuidas a características del alumnado como el sexo, la repetición, ser inmigrante o el estatus socioeconómico familiar. De hecho no todos los centros se comportan de forma similar o no muestran valores de equidad.

Por otra parte se probaron distintos indicadores de contexto y proceso evaluados entre el alumnado y la dirección a través de un diseño multinivel. Dicho modelo tuvo como fin de obtener un diagnóstico de los factores que predicen rendimiento científico. Los resultados indicaron la importancia de características individuales del alumno como sexo, procedencia, ser repetidor y las características del contexto como el estatus socioeconómico familiar y escolar. Además se demostró la influencia de variables de proceso entre el alumnado como la motivación, la eficacia y la concienciación de problemas medio ambientales, las cuales se adquieren a través de prácticas escolares. También resultó ser un predictor de éxito en la competencia científica el tiempo medio que dedica el centro al aprendizaje de su alumnado.

A su vez los resultados presentados plantean futuras líneas de trabajo como probar la estabilidad de un modelo multinivel referido a la competencia científica en otras cohortes de alumnado o desarrollar una evaluación interna que implique diversos agentes educativos y proponga líneas de mejora contextualizadas a la propia realidad del centro.

REFERENCIAS

- Ackerman, P. (1996). Intelligence as process and knowledge: An integration for adult development and application. En W. Rogers y A. Fisk (Eds.), *Aging and skill performance: Advances in theory and applications*. Hillsdale, NJ: Erlbaum.
- Adams, R. J. (2003). Response to cautions on OECD's recent educational survey (PISA). *Oxford Review of Education*, 29, 38-389.
- Adams, R. J., Wilson, M., y Wang, W. C. (1997). The multidimensional random coefficients multinomial logit model. *Applied Psychological Measurement*., 21, 1-23.
- Adams, R. J., y Wu, M. L. (2002). *PISA 2000. Technical Report*. París: OECD.
- Ainley, M. A., y Ainley, J. B. (2011). A culture perspective on the structure of student interest in science. *International Journal of Science Education*, 33(1), 51-71.
- Aitkin, M., y Longford, N. (1986). Statistical modelling issues in school effectiveness studies. *Journal of the Royal Statistical Society*, 149(Series A), 1-43.
- Alker, H. R. (1969). A typology of ecological fallacies. En M. Dogan y S. Rokkan (Eds.), *Quantitative ecological analysis in the school science*. Cambridge Mass: MIT Press.
- Allport, G. (1937). *Personality: A psychological interpretation*. Nueva York: Holt.
- APA. (2005). *Publication Manual of the American Psychological Association*. Washington, DC: American Psychological Association.
- Artelt, C., Schiefele, U., y Scheneider, S. (2001). Predictors of reading literacy. *European Journal of Psychology of Education*, 16(3), 262-383.
- Ato, M., y Vallejo, G. (2007). *Diseños experimentales en Psicología*. Madrid: Pirámide.
- Askew, M., y William, D. (1995). Recent research in mathematics education. . En *Office for standars in education*. (Vol. 53, pp. 5-16). Londres: HMSO.
- Austin, G. R. (1978). *Process evaluation: A comprehensive study of outliers*. Baltimore: Maryland Department of Education.
- Barrueco, A. (1992). *Calidad y eficacia en los centros educativos: análisis de las aprotaciones de un modelo explicativo*. Tesis no publicada, Universidad de Salamanca.
- Beaton, A. (1997). Item sampling. En J. P. Keeves (Ed.), *Educational research, methodology and measurement: An international handbook* (2. ed., pp. 976-984). Cambridge: Pergamon.

- Bickel, R. (2007). *Multilevel analysis for applied research: It's just regression!* Nueva York: Guilford Press.
- Blakey, L. S., y Heath, A. F. (1992). Differences between comprehensive schools: Some preliminary findings. En D.Reynolds y P.Cuttance (Eds.), *School Effectiveness. Research policy and practice* (pp. 121-133). Londres: Casell.
- Blickle, G. (1996). Personality traits, learning strategies and performance. *European Journal of Personality*, 10, 337-352.
- Blosser, P. (1984). *Attitude research in science education*. Columbus: ERIC Clearinghouse for Science, Mathematics and Environment Education.
- Bolívar, A. (1999). *Cómo mejorar los centros educativos*. Madrid: Síntesis.
- Bonvin, P., Bless, G., y Schuepbach, M. (2008). Grade retention: Decision-making and effects on learning as well as social and emotional development. *School Effectiveness & School Improvement*, 19(1), 1-19.
- Borich, G. (2009). *Effective teaching methods*. Upper Saddle River, NJ: Merrill Pub Co.
- Bosker, R. J., Kremers, E. J., y Lugthart, E. (1990). School and struction effects on mathematics achievement. *School Effectiveness and School Improvement*, 1(4), 233-248.
- Brandsma, H. P. (1993). *Characteristics of primary schools and the quality of education*. Croninger: RION.
- Brennan, R. L. (2002). *Elements of generalizability theory*. Nueva York: Springer.
- Brimer, A. (1978). *Sources of difference in school achievement*. Londres: NFER.
- Brookover, W. B., Beady, C., Flood, P., Schewitzer, J., y Wisenbaker, J. (1979). *School social system and student achievement: Schools can make a difference*. Nueva York: Praeger.
- Brookover, W. B., y Scheneider, J. M. (1975). Academic environments and elementary school achievement. *Journal research and development in education*, 9, 82-91.
- Brookver, W. B., y Lezzote, L. W. (1977). *Changes in school characteristics coincident with changes in student achievement*. East Lansing: Institute for Research on Teaching, Michigan State University.
- Brophy, J., y Good, T. L. (1986). Teacher behavior and student achievement. En M. C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 328-375). Nueva York: Macmillan.

- Brown, A. (2009). *A guide to effective instruction*. Boston MA: Houghton Mifflin Company.
- Bryk, A., y Raudenbush, S. (1992). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage Publications.
- Buchmann. (2002). Measuring family background in international studies of education: Conceptual issues and methodological challenges En A. C. Porter y M. Gomolla (Eds.), *Methodological advances in cross-national surveys of educational achievement* (pp. 150-197). Washington, D.C.: National Academy Press.
- Bybee, R. (1997). Towards an understanding of scientific literacy. En W. Gräber y C. Bolte (Eds.), *Scientific Literacy: An international symposium*. Kiel: IPN
- Caffrey, E., Fuchs, D., y Fuchs, L. S. (2008). The predictive validity of dynamic assessment: A review. *Journal of Special Education*, 41(4), 254-270.
- Cantón, I., y Arias, A. R. (2008). La dirección y el liderazgo: Aceptación, conflicto y calidad. *Revista de Educación*, 345, 229-254.
- Carroll, J. B. (1963). A model for school learning. *Teachers College Record*, 64(8), 723-733.
- Casanova, M. A. (2007). Hacia la selección de indicadores críticos en los centros docentes. *Avances en Supervisión Educativa*, 7.
- Castejón, J. L. (1994). Estabilidad de diversos índices de eficacia de centros educativos. *Revista de Investigación Educativa*, 24, 45-60.
- Castejón, J. L. (1996). *Determinantes del rendimiento académico de los estudiantes y de los centros educativos: Modelos y factores*. Alicante: Editorial Club Universitario.
- Cazden, C. B. (1986). Classroom discourse. En M. C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 432-463). Nueva York: Macmillan.
- Chamorro-Premuzic, T., y Furnham, A. (2003). Personality predicts academic performance: Evidence from two longitudinal university samples. *Journal of Research in Personality*, 37, 319-338.
- Childs, R., y Jaciw, A. P. (2003). Matrix sampling of items in large-scale assessments. *Practical assessment research and evaluation* 8(16). Consultado el 23/05/2011, en <http://PAREonline.net/getvn.asp?v=8&n=16>
- Cohen, J., y Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

- Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., et al. (1966). *Equality of educational opportunity*. Washington: US Government Printing Office.
- Coleman, J. S., Hoffer, T., y Kilgore, S. B. (1981). *Public and private schools*. Chicago: National Opinion Research Center.
- Coleman, J. S., Hoffer, T., y Kilgore, S. B. (1982). *High school achievement: Public, catholic and other private schools compared*. Nueva York: Basic.
- Connell, N. (1996). *Getting off the list: School improvement in New York City*. Nueva York City: Educational Priorities Panel.
- Costa, P. T. J., y McCrae, R. R. (1997). Longitudinal stability of adult personality. En R. Hogan, J. Johnson y S. Briggs (Eds.), *Handbook of personality psychology* (pp. 269-290). San Diego: Academic Press.
- Cotton, K. (1995). *Effective school practices: A research synthesis. 1995 updated*. Portland, OR: Northwest Regional Educational Laboratory.
- Creemers, B., y Kyriakides, L. (2006). Critical Analysis of the Current Approaches to Modelling Educational Effectiveness: The importance of establishing a dynamic model. *School Effectiveness and School Improvement*, 17(3), 347-366.
- Creemers, B., y Reezigt., G. J. (1997). School Effectiveness and school improvement: Sustaining links. *School Effectiveness and School Improvement*, 8(4), 396-429.
- Creemers, B., Stoll, L., Reezigt., G. J., y Team., E. (2007). Effective school improvement - ingredients for success: The results of an international comparative study of best practice case studies. En T. Townsend (Ed.), *International Handbook of school effectiveness and improvement* (pp. 825-838). Dordrecht: Springer.
- Creemers, B. P. M. (1992). School effectiveness and effective instruction. The need for a further relationship. En J. Bashi y Z. Sass (Eds.), *School effectiveness and improvement*. Jerusalem: Hebrew University Press.
- Creemers, B. P. M. (1994). *The effective classroom*. Londres: Casell.
- Creemers, B. P. M. (1996). The school effectiveness knowledge base. En D. Reynolds, R. Bollen, B. P. M. Creemers, D. Hopkins, L. Stoll y N. Largerweij (Eds.), *Making good schools* (pp. 36-58). Londres: Routledge.
- Creemers, B. P. M., y Kyriakides, L. (2006). Critical analysis of the current approaches to modelling educational effectiveness: The importance of establishing a dynamic model. *School Effectiveness y School Improvement*, 17(3), 347-366.

- Creemers, B. P. M., y Kyriakides, L. (2008). *Dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools*. Londres: Routledge.
- Creemers, B. P. M., y Kyriakides, L. (2010). Explaining stability and changes in school effectiveness by looking at changes in the functioning of school factors. *School Effectiveness y School Improvement*, 21(4), 407-427.
- Creemers, B. P. M., y Reezigt, G. J. (1996). School level conditions affecting the effectiveness of instructions. *School Effectiveness y School Improvement*, 7, 197-228.
- Creemers, B. P. M., y Scheerens, J. (1994). Developments in the educational effectiveness research programme. *International Journal of Education Research*, 21(2), 125-139.
- Creemers, B. P. M., Stringfield, S., y Guldmond, H. (2002). The quantitative data. En D. Reynolds, B. Creemers, S. Stringfield, C. Teddlie y E. Schaffer (Eds.), *World class schools: International perspectives on school effectiveness* (pp. 33-55). Londres: Routledge, Falmer.
- Cronbach, L. J., Gleser, G. C., Nanda, N., y Rajaratnam, N. (1972). *The dependability of behavioral measures: The generalizability theory for scores and profiles*. Nueva York.
- Crone, L. J., Lang, M., y Franklin, B. J. (1994). *Achievement measures of school effectiveness: Comparison of consistency across years*. New Orleans.
- Crone, L. J., Lang, M., Franklin, B. J., y Hallbrook, A. (1994). Composite versus component score: Consistency of school effectiveness classification. *Applied Measurement in Education* 7(4), 303-321.
- Darling-Hammond, L. (2000). Teacher quality and student achievement: a review of state policy evidence. *Journal*. Consultado el 16/05/11, en <http://epaa.asu.edu/ojs/article/view/392>
- de Jong, R., Westerhof, K. J., y Kruiter, J. H. (2004). Empirical evidence of a comprehensive model of school effectiveness: A multilevel study in mathematics in the first year of junior general education in the Netherlands. *School Effectiveness y School Improvement*, 15, 3-31.
- Demetriou, A., Kyriakides, L., y Avraamidou, C. (2003). The missing link in the relations between intelligence and personality. *Journal of Research in Personality*, 37(6), 547-581.

- den Brok, P., Brekelmans, M., y Wubbels, T. (2004). Interpersonal teacher behaviour and student outcomes. *School Effectiveness y School Improvement*, 15, 407-442.
- D'Haenens, E., van Damme, J., y Onghena, P. (2010). Multilevel exploratory factor analysis: Illustrating its surplus value in educational effectiveness research. *School Effectiveness y School Improvement*, 21(2), 209-235.
- Dosey, J. A., McCrone, S. A., y O'Sullivan, C. (2006). Problem solving in the PISA and TIMSS 2003. Consultado el 11/07/11, en <http://nces.ed.gov/pubsearch>
- Doyle, W. (1990). Classroom knowledge as a foundation for teaching. *Teacher college record*, 9(3), 347-360.
- Draper, D. (1995). Inference and hierarchical modeling in the social sciences. *Journal of Educational Statistics*, 20(2), 115-148.
- Durán, D. (2004). La reacción de los profesores a las evaluaciones externas de los establecimientos escolares: La proposición de un modelo. *Education Policy Analysis Archives*, 12(9).
- Edmonds, R. (1979). Effective schools for the urban poor. *Educational Leadership*, 37(1), 15-24.
- Ellis, A. (1975). *Sucess and failure: A summary of findings and recommendations for improving elementary reading in Massachussetts city schools*. Waterton, MA: Educational Research Corporation.
- Elousua, P. (2006). Funcionamiento diferencial del ítem en la evaluación internacional PISA. Detección y comprensión. *RELIEVE*, 12(002), 247-259.
- Everston, C., Anderson, C., Anderson, L., y Brophy, J. (1980). Relationships between classroom behaviour and student outcomes in junior high math and English classes. *American Educational Research Journal.*, 17, 43-60.
- Fay, R. E. (1989). Theoretical applications of weighting for variance calculation. *Proccedings of the selection on survey research methods of the american statistical association*, 212-217.
- Fensham, P. J. (2000). Time to change drivers for scientfic literacy. *Canadian Journal of Science, Mathematics and Technology Education* 2, 9-24.
- Fernández-Alonso, R., y Muñiz, J. (2011). Diseño de cuadernillos para la evaluación de las Competencias Básicas. *Aula Abierta*, 39(2), 3-34.
- Ferrando, P. J., y Anguiano-Carrasco, C. (2010). El análisis factorial como técnica de investigación en Psicología. *Papeles del Psicólogo*, 31(1), 18-33.

- Filp, J. (ed.). (1984). *La educación prescolar mirada desde la escuela*. Santiago de Chile: CIDE.
- Fitz-Gibbon, C. T. (1991). Multilevel modelling in an indicator system. En S. W. Raudenbush y J. D. Willms (Eds.), *Schools classrooms and pupils: International studies of schooling from multilevel perspective* (pp. 67-83). San Diego, CA: Academic Press.
- Fitz-Gibbon, C. T. (1996). *Monitoring education: Indicators, quality and effectiveness*. Londres: Cassell.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34, 906-911.
- Fraser, B. J. (1991). Two decades of classroom environment research. En B. J. Fraser. y H. J. Walberg. (Eds.), *Educational environments: evaluations, antecedents and consequences* (pp. 3-29). Oxford: Pergamon.
- Fraser, B. J., Walberg, H. J., Welch, W. W., y Hattie, J. A. (1987). Synthesis of educational productivity research. *International Journal of Educational Research*, 11(2), 145-152.
- Frey, A., Hartig, J., y Rupp, A. (2009). An NCME instructional module on booklet design in large-scale assessments of student achievement: Theory and practice. *Educational Measurement Issues and Practice*, 28(3), 39-53.
- Fuentes, A. (1986). *Procesos funcionales y eficacia de la escuela. Un modelo causal*. Madrid: Universidad Complutense de Madrid.
- Fullan, M. (1991). *The new meaning of educational change* Londres: Cassell.
- Fuller, B., y Clarke, P. (1994). Raising school effects while ignoring the culture? Local conditions and the influence of classroom tools, rules and pedagogy. *Review of educational research*, 64(1), 119-157.
- Gage, N. L. (1963). *Handbook of research on teaching*. Chicago: Rand McNally.
- Ganzeboom, H. B. G., De Graaf, P. M., y Treiman, D. J. (1992). Standard international socio-economic index of occupational status. *Social Science Research*, 21, 1-56.
- García-Durán, M. (1991). *Investigación evaluativa sobre las variables pedagógicas que discriminan entre los centros de EGB de alto y bajo rendimiento en el medio rural de la provincia de Cádiz*. Tesis doctoral no publicada, UNED.
- Gardner, P. L. (1975). Attitudes to science. A review. *Studies in science education*, 2, 1-41.

- Gardner, P. L. (1984). Students' interest in Science and Technology: An international overview. En M. Lehrke., L. Hoffmann. y P. L. Gardner (Eds.), *Interests in science and technology education* (pp. 15-84). Kiel: IPN
- Gauld, C., y Hukins, A. A. (1980). Scientific attitudes: A review. *Studies in Science Education*, 7, 129-161.
- Gaviria, J., Martínez-Arias, R., y Castro, M. (2004). Un estudio multinivel sobre los factores de Eficacia Escolar en países en desarrollo: El caso de los recursos en Brasil. *Education Policy Analysis Archives*. Consultado el 06/10/2008, en <http://epaa.asu.edu/epaa/v12n20/>.
- Gifi, A. (1996). *Nonlinear multivariate analysis*. Nueva York: John Wiley & Sons Ltd.
- Goldstein, H. (1987). Multilevel covariance component models. *Biometrika*, 74, 430-431.
- Goldstein, H. (1995). *Multilevel statistical models*. Londres: Edward Arnold.
- Goldstein, H. (2001). *League tables and Schooling*. Londres: Institute of Education, University of Londres.
- Goldstein, H. (2003). *Multilevel statistical model (3rd edn)* Londres: Arnold.
- Goldstein, H., y Sammons, P. (1997). The influence of secondary and junior schools on sixteen year examination performance: A cross-classified multilevel analysis. *School Effectiveness y School Improvement*, 8, 219-230.
- Gómez-Dacal, G. (1989). *Los determinantes de la productividad de la enseñanza*. Tesis doctoral no publicada, Universidad de Barcelona.
- Gomolla, M., y Radtke, F. O. (2002). *Institutionelle Diskriminierung. Die Herstellung ethnischer Differenz in der Schule*. Opladen: Leske und Budrich.
- González, E. J., y Kennedy, A. M. (2003). *PIRLS 2001. User guide for the international database*. Boston College: Chesnut Hill.
- González-Galán, A. (2000). *Calidad, eficacia y clima en centros educativos. Modelos de evaluación y relaciones causales*. Tesis doctoral no publicada, Universidad Complutense de Madrid.
- Good, T. L., Wiley, C. R. H., y Florez, I. R. (2009). Effective teaching: An emerging synthesis. En L. J. Saha y A. G. Dworkin (Eds.), *International handbook of research on teachers and teaching* (pp. 803-815). Nueva York: Springer.
- Gräber, W., y Bolte, C. (1997). *Scientific literacy: An international symposium*. Kiel: IPN

- Gray, J., McPherson, A. E., y Raffe, D. (1983). *Reconstructions of secundary education: Theory, myth and practice since the war*. Londres: Routledge and Kegan Paul.
- Grisay, A. (1996). *Evolution of cognitive and effective development in lower secondary school*. Enschede: University of Twente.
- Hallinger, P., y Heck, R. B. (1998). Exploring the principal's contribution to school effectiveness: 1980-1995. *School Effectiveness y School Improvement*, 9(2), 157-191.
- Hambleton, R., Merenda, P. F., y Spielberger, C. D. (2005). *Adapting educational and psychological tests for cross-cultural assessment*. Nueva Jersey: IEA Lawrence Erlbaum Associates.
- Hambleton, R. K. (2002). Adapting achievement tests into multiples languages for international assesment. En A. C. Porter y A. Gamoran (Eds.), *Methodological advances in cross-national surveys of educational achievment* (pp. 58-79).
- Hanuskek, E. A. (1986). The economics of schooling: Production and efficiency in public schools. *Journal of economic lierature*, 24, 1141-1177.
- Hanuskek, E. A. (1997). Assessing the effects or school resources on student performance: An update. *Educational evaluation and policy analysis*, 19(2), 141-164.
- Hargreaves, A., Lieberman, A., Fullan, M., y Hopkins, D. (eds). (1998). *International Handbook of Educational Change* Dordretch: Kluver Academic Publishers.
- Heck, R. K., y Scott, L. T. (2000). *An introduction to multinivel modeling techniques*. Mahwah: Lawrence Erlbaum Associates.
- Hextall, J., y Mahony, P. (1998). *Effective teachers effective schools* Londres: Biddles.
- Hill, P. W., Rowe, K. J., y Jones, T. (1995). Factors affecting student's educational progress: Multinivel modelling of educational effectiveness. Congress for School Efectiveness and Improvement. Leewarden.
- Hopfenbeck, T. N., y Maul, A. (2011). Examining evidence for the validity of PISA learning strategy scales based on sutdent response processes. *Routledge*, 11, 95-121.
- Hopkins, D., y Reynolds, D. (2002). *The past, present and future of school improvement*. Londres: DfES.

- Hox, J. J., y Mass, C. J. M. (2005). Multilevel analysis. En K. Kempf-Leonard (Ed.), *The encyclopedia of social measurement* (pp. 785-793). San Diego: Elsevier Academic Press.
- IE. (2007). *PISA 2006. Programa para la evaluación internacional de alumnos de la OCDE. Informe español*. Madrid: Ministerio de Educación. Secretaría de Estado de Educación y Formación Profesional. Dirección General de Evaluación y Cooperación Territorial.
- IE. (2010a). *PISA 2009. Programa para la Evaluación Internacional de los Alumnos. OCDE. Informe Español*. Madrid: Ministerio de Educación. Secretaría de Estado de Educación y Formación Profesional. Dirección General de Evaluación y Cooperación Territorial.
- IE. (2010b). *Evaluación General de Diagnóstico 2009. Educación Primaria. Cuarto Curso. Informe de Resultados* Madrid: Ministerio de Educación. Secretaría de Estado de Educación y Formación Profesional. Dirección General de Evaluación y Cooperación Territorial.
- IE. (2010c). *ICCS 2009. Estudio internacional de civismo y ciudadanía IEA. Informe español*. Madrid: Ministerio de Educación. Secretaría de Estado de Educación y Formación Profesional. Dirección General de Evaluación y Cooperación Territorial.
- IE. (2010d). *TALIS. Estudio Internacional sobre la Enseñanza y el Aprendizaje. Informe español 2009*. Madrid: Ministerio de Educación. Secretaría de Estado de Educación y Formación Profesional. Dirección General de Evaluación y Cooperación Territorial.
- ILO. (1990). *ISCO-88: International Standard Classification of Occupations. International Labour*. . Geneva: International Labour Office.
- ISEI-IVEI. (2008). *Informe final de la evaluación PISA 2006. Proyecto para la evaluación internacional de los estudiantes de 15 años en Ciencias, Matemáticas y Lectura. Resultados en Eusakadi*. Bilbao: Gobierno Vasco.
- Jansen, J. D. (1995). Effective schools? Comparative education. *Comparative Education Review* 31(2), 181-200.
- Jencks, C. S., Smith, M., Acland, H., Bane, M. J., Cohen, D., Gintis, H., et al. (1972). *Inequality: A reassessment of the effect of family and schooling in America*. Nueva York: Basic Books.

- Jesson, D., y Gray, J. (1991). Slants on slopes: Using multi-level models to investigate differential school effectiveness and its impact on pupil's examination results. *School Effectiveness y School Improvement*, 2(3), 230-247.
- Jimerson, S. R., Anderson, G. E., y Whipple, A. (2002). Winning the battle and losin the war: Examining the relation between grade retention and dropping out of high school. *Psychology in the Schools*, 39, 441-457.
- Jöreskog, K. G., y Sörbom, D. (1993). *LISREL* Lincolwood, IL: Scientific Software International, inc.
- Judkins, D. R. (1990). Fay's Method of Variance Estimation. *Journal of Official Statistics*, 6(3), 223-239.
- Kaplan, L. S., y Owings, W. A. (2002). *Teacher quality, teaching quality and school improvement*. Bloomington, IN: Phi Delta Kappa.
- Kappa, P. D. (1980). *Why some urban schools succeed?* Bloomington, IN: Phi Delta Kappa.
- Kim, J.-S. (2009). Multilevel analysis: An overview and some contemporary issues. En R. E. Millap. y A. Maydeu-Olivares. (Eds.), *The SAGE handbook of quantitative methods in Psychology*. Londres: Sage Publications.
- Kim, J.-S., y Frees, E. W. (2005). Fixed effects estimation in multilevl models. Consultado el 18/05/11, en <http://research.bus.wisc.edu/jfrees>
- Kim, J.-S., y Fress, E. W. (2006). Omitted variables in multilevel methods. *Psychometrika*, 71, 659-690.
- Kish, L. (1965). *Survey sampling*. Nueva York: Wiley.
- Kish, L. (1987). *Statistical design for research*. Nueva York Wiley.
- Klopfer, L. E. (1976). A structure for the affective domain in relation to science education. *Science Education*, 60(3), 299-312.
- Knuver, A. W., y Brandsam, H. P. (1993). Cognitive and affective outcomes in school efectiveness research. *School Effectiveness and School Improvement*, 4(3), 189-204.
- Knuver, A. W., y Brandsma, H. P. (1993). Cognitive and Affective Outcomes in School Effectiveness Research. *School Effectiveness and School Improvement*, 4(3), 189-204.
- Kochan, S. E., Tashakkori, A., y Teddlie, C. (1996). You can't judge a high school by test data alone: Constructing an alternative indicator of secondary school

- effectiveness. American Educational Research Association Annual meeting. Nueva York, NY.
- Kreft, I. (1987). *Models and methods for the measurement of school effects*. Tesis doctoral no publicada, Universidad de Amsterdam.
- Kreft, I., y De Leeuw, J. (1998). *Introducing multilevel modeling*. Newbury Park, CA: Sage.
- Kyriakides, L. (2002). A research based model for the development of policy on baseline assessment *British Educational Research* 28(6), 805-826.
- Kyriakides, L. (2005). Extending the comprehensive model of educational effectiveness by an empirical investigation. *School Effectiveness y School Improvement*, 16, 103-152.
- Kyriakides, L. (2007). Generic and differentiated models of educational effectiveness. Implications for the improvement of educational practice. En T. Townsend (Ed.), *International handbook of school effectiveness and improvement* (pp. 41-57). Dordrecht: Springer
- Kyriakides, L., y Creemers, B. (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), 183-105.
- LaForgia, J. (1988). The affective domain related to science education and its evaluations. *Science Education*, 72(4), 407-421.
- Lavonen, J., y Laaksonen, S. (2009). Context of teaching and learning school science in Finland: Reflections on PISA 2006 results. *Journal of Research in Science Teaching*, 46, 922-944.
- Levine, D. U., y Lezotte, L. W. (1990). *Unusually effective schools: A review and analysis of research and practice*. Madison, WI.: National Center for Effective Schools Research and Development.
- Lezotte, L. W., Edmonds, R., y Ratner, G. (1974). *A final report: Remedy for school failure to equitable basic school skills*. East Lansing: Michigan State University.
- Linn, R. L. (2002). The measurement of student achievement in international studies. En A. C. Porter y A. Gamoran (Eds.), *Methodological advances in cross-national surveys of educational achievement* (pp. 27-57). Washington, D.C.: National Academy Press.
- Linn, R. L. (2004). Assessment and accountability. *Educational Researcher*, 29, 4-14.

- LOCE (2002). *Ley Orgánica de Calidad de la Educación*. Madrid: MEC.
- LOE (2006). *Ley Orgánica de Educación*. Madrid: MEC.
- LOGSE. (1990). *Ley de Ordenación General del Sistema Educativo*. Madrid: MEC.
- LOPEG (1995). *Ley Orgánica de Participación, Evaluación y Gobierno*. Madrid: MEC.
- Longford, N. (1998). *Random coefficient models*. Oxford: Clarendon Press.
- Louis, K. S., Dretzke, B., y Wahlstrom, K. (2010). How does leadership affect student achievement? Results from a national US survey. *School Effectiveness y School Improvement*, 21(3), 315-336.
- LLECE (2001). *Primer estudio internacional comparativo sobre Lenguaje, Matemáticas y factores asociados, para alumnos del tercer y cuarto grado de la educación básica. Informe técnico*. Santiago de Chile: OREALC/UNESCO.
- MacBeath, J., y Mortimore, P. (Eds.). (2001). *Improving School Effectiveness*. Buckingham: Open University Press.
- Madden, J. V., Lawson, D. R., y Sweet, D. (1976). *School Effectiveness Study*. Sacramento, CA.: State of California Department of Education.
- Mandeville, G. K. (1988). School effectiveness indicators revisited: Cross-year stability *Journal of Education Measurement*, 24(3), 203-216.
- Mandeville, G. K., y Anderson, L. W. (1987). The stability of school effectiveness indices across grade level and subjects areas. *Journal of Educational Measurement*, 24(3), 203-216.
- Marchesi, A. (2006a). El informe PISA y la política educativa en España. *Revista de Educación, extraordinario 2006*, 337-355.
- Marchesi, A. (2006b). Cambio educativo y evaluación de escuelas. *Transatlántica educación*, 1, 43-54.
- Marchesi, A., y Martínez-Arias, R. (coords.). (2006). *Escuelas de éxito en España. Sugerencias e interrogantes a partir del informe PISA 2003*: Fundación Santillana.
- Marks, J., Cox, C., y Pomian-Srzednicki, M. (1983). *Standars in english schools*. Londres: London National Council for Eduational Standars.
- Martín, E., Martínez-Arias, R., Marchesi, A., y Pérez, E. M. (2008). Variables the predict academic achievement in the spanish compulsory second education system: A longitudinal multi-level analysis. *The Spanish Journal of Psychology*, 11, 400-413.

- Martin, M. O., Gregory, K. D., y Stemler, S. E. (2000). *Technical Report. Third international mathematics and Science study: 1999*. Boston College: Chestnut Hill, MA.
- Martin, M. O., Mullis, I. V. S., Gregory, K. D., Hoyle, C., y Shen, C. (2000). *Effective schools in Science and Mathematics*. Boston, MA.: TIMSS International Study Center. Boston College.
- Martin, M. O., Mullis, L. V. S., y Chrostowski, S. J. (2004). *TIMSS 2003. Technical report*. Boston College: Chestnut Hill, MA.
- Martin, M. O., Mullis, L. V. S., y Chrostowski, S. J. (2005). *PISA 2003 data analysis manual: SPSS users*. Paris: OCDE.
- Martínez-Arias, R. (1995). *Psicometría: Teoría de los tests psicológicos y educativos*. Madrid: Síntesis.
- Martínez-Arias, R. (2006). La metodología de los estudios PISA. *Revista de Educación, n° extraordinario*, 111-119.
- Martínez-Arias, R. (2010). La evaluación del desempeño. *Papeles del psicólogo, 31(1)*, 85-92.
- Martínez-Arias, R., Gaviria-Soto, J. L., y Castro-Morera, M. (2008). Concepto y evolución de los modelos de valor añadido en educación. *Revista de Educación, 348*, 15-45.
- Masters, G. N., y Wright, B. D. (1997). The partial credit model. En W. J. van der Linder y R. K. Hamilton (Eds.), *Handbook of modern item response theory* (pp. 101-122). Nueva York/Berlin/Heidelberg: Springer.
- Mayer, V. J. (ed.). (2002). *Global science literacy*. Dordrecht: Kluwer Academic Publishers.
- Mayeske, G. W., Wisler, C. E., Beaton, A. E., Weinfield, E. O., Cohen, W. M., Okada, T., et al. (1972). *A study of our nations's schools*. Washington, D.C.: US Department of Health Education and Welfare.
- McIntosh, R. G. (1968). Equal Educational Opportunity. *Harvard Educational Review, 301-308*.
- Millán, M. (1978). *La eficacia en la educación escolar*. Tesis doctoral no publicada, Universidad de Valencia.
- Millar, R., y Osborne, J. (1998). *Beyond 2000: Science education for the future*. Londres: King's College Londres School of Education.

- Miller, S. K. (1985). Research on exemplary schools: An historical perspective. En G.R. Austin. y H. Garber (Eds.), *Research on exemplary schools* (pp. 3-30). Orlando, Fl: Academic Press.
- Mislevy, R. J. (1991). Randomization-based inference about latent variables from complex samples. *Psychometrika*, 56, 177-196.
- Mislevy, R. J., Beaton, A., y Sheehan, K. (1992). Estimating population characteristics from sparse matrix samples of item responses. *Journal of Education Measurement*, 29(2), 131-161.
- Mislevy, R. J., y Sheehan, K. M. (1989). The role collateral information about examinees in item parameter estimation. *Psychometrika*, 54, 661-679.
- Mislevy, R. J., Steinberg, L. S., y Almond, R. G. (2003). On the structure of educational assessments. *Interdisciplinary Research and Perspectives*, 1, 3-67.
- Mislevy, R. K., Johnson, E. G., y Muraki, E. (1992). Scaling procedures in NAEP. *Journal of Education Statistics*, 17, 131-154.
- Mortimore, P. (1991). The nature and findings of school effectiveness research in primary sector. En S. Riddell y S. Brown (Eds.), *School Effectiveness research: Its messages for school improvement* (pp. 9-19). Edinburgh: HMSO.
- Mortimore, P., Sammons, P., Stoll, L., Lewis, D., y Ecob, R. (1986). *The ILEA school project: Summary report*. Londres Research and Statistics, Inner Londres Educational Authority.
- Mortimore, P., Sammons, P., Stoll, L., Lewis, D., y Ecob, R. (1988). The effects of school membership on pupil's educational outcomes. *Research Papers in Education*, 3(1), 3-26.
- Mujis, D., y Reynolds, D. (2000). School efectiveness and teacher effectiveness in Mathematics: Some preliminary findings from the evaluation of the Mathematics Enhancement Programme (Primary). *School Effectiveness y School Improvement*, 11, 273-303.
- Mujis, D., y Reynolds, D. (2001). *Effectiveness teaching*. Londres: Sage.
- Muñiz, J. (1997). *Introducción a la teoría de respuesta a los ítems*. Madrid: Pirámide.
- Muñiz, J. (2001). *Teoría clásica de los tests*. Madrid: Pirámide.
- Muñiz, J., Fidalgo, A. M., García-Cueto, E., Martínez, R., y Moreno, R. (2005). *Análisis de los ítems*. Madrid: Muralla.

- Muñoz-Repiso, M., Cerdán, J., Murillo, F. J., Calzón, J., Egido, I., García, R., et al. (1995). *Calidad de la educación y eficacia de la escuela. Estudio sobre la gestión de los recursos educativos*. Madrid: CIDE.
- Muñoz-Repiso, M., y Murillo, J. (coord.). (2003). *Mejorar las escuelas, mejorar los resultados. Investigación europea sobre mejora de la eficacia escolar*. Bilbao: Mensajero.
- Murillo, F. J. (1996). ¿Son eficaces nuestras escuelas? *Cuadernos de Pedagogía*, 246, 66-72.
- Murillo, F. J. (2011). Mejora de la eficacia escolar en Iberoamérica. *Revista Iberoamericana de Educación*, 55, 49-83.
- Murillo, F. J., Martínez-Garrido, C. A., y Hernández-Castilla, R. (2011). Decálogo para una enseñanza eficaz. *Journal*, 9(1), 6-27. Consultado el 26/04/11, en <http://www.rinace.net/reice/numeros/artsvol9num1/art1/art1.pdf>.
- Murillo, F. J., y Román, M. (2011). School infrastructure and resources do matter: Analysis of the incidence of school resources on the performance of Latin American students. *School Effectiveness y School Improvement*, 22(1), 29-50.
- Murillo, F. J. (coord.). (2003). *La investigación sobre eficacia escolar en Iberoamérica. Revisión del estado de la cuestión*. Bogotá: Convenio Andrés Bello.
- Murillo, F. J. (coord.). (2007). *Investigación iberoamericana sobre eficacia escolar*. Bogotá: Convenio Andrés Bello.
- Murillo, F. J. (coord.). (2008). Hacia un Modelo de Eficacia Escolar. Estudio Multinivel sobre los Factores de Eficacia de las Escuelas Españolas. 6(1). Consultado el 18/09/08, en <http://www.rinace.net/arts/vol6num1/art1.pdf>.
- Murillo, F. J. (2001). Eskola eraginkortasunaren hobekuntzarakp europako eredu baterantz. *Euskal Herriko Unibertsitateko Hezkuntza Aldizkaria*, 25, 7-25.
- Murillo, F. J. (2004). *Aportaciones de la investigación sobre eficacia escolar. Un estudio multinivel sobre los efectos escolares y los factores de eficacia de los centros docentes de primaria en España*. Universidad Complutense de Madrid, Madrid.
- Murillo, F. J. (2005). *La investigación sobre eficacia escolar*. Barcelona: Octoedro.
- Murphy, J. (1992). School effectiveness and school restructuring: Contributions to educational improvement. *School Effectiveness and School Improvement*, 3(2), 90-109.

- Murphy, J., Hallinger, P., y Mesa, R. P. (1985). School effectiveness: Checking progress and assumptions and developing a role for state and federal government. *Teachers College Record*, 86(4), 615-641.
- Navas, M. J., y Urdaneta-Durán, E. J. (en prensa). PISA y el triángulo de la evaluación. *Psicothema*.
- Nevo, D. (1995). *School-based evaluation. A dialogue for school improvement*. Oxford: Pergamon.
- New York State Department of Education (1974). *School factors influencing reading achievement: A case study of two inner city schools*. Albany, N.Y.: Division of Education and Evaluation.
- New York State Department of Education (1976). *Three strategies for studying the effects of school process*. Albany, N.Y.: Bureau of School Programs Evaluation.
- Nicalaidou, M., y Petridou, A. (2011). Evaluation of CPD programme: Challenges and implications for leader and leadership development. *School Effectiveness y School Improvement*, 22(1), 51-85.
- Nishisato, S. (1980). *Analysis of categorical data: Scaling and its applications*. Toronto: University of Toronto.
- Nuttall, D. L., Goldstein, H., Prosser, T., y Rabash, J. (1989). Differential school effectiveness. *International Journal of Educational Research*, 7(13), 769-776.
- OCDE. (1999). *Measuring student knowledge and skills: A new framework for assesment*. París: OECD publications.
- OCDE. (2003). *Definiton and Selection of Competencies: Theoretical and conceptual foundations (DeSeCo), summary of the final report <<key competencies for a successfullife and a well-functioning society>>*. París: OECD publications.
- OCDE. (2003). *The PISA 2003 assesment framework. Mathematics, Reading, Science and Problem Solving knowledg and skills*. París: OECD publications.
- OCDE. (2005). *Are Students Ready for a Technology-Rich World? What PISA Studies Tell Us*. París: OECD publications.
- OCDE. (2004a). *Innovation in the knowledge economy: Implications for education and learning*. París: OECD publications.
- OCDE. (2007a). *Evidence in education: Linking research and policy*. París: OECD publications.

- OCDE. (2004b). *Learning for tomorrow 's world. First results from PISA 2003*. París: OECD publications.
- OCDE. (2006a). *PISA 2003. Manual de análisis de datos. Usuarios de SPSS*. Madrid: INECSE.
- OCDE. (2006b). *PISA 2006. Marco de la evaluación. Conocimientos y habilidades en Ciencias, Matemáticas y Lectura*. Madrid: Santillana Educación, S.L.
- OCDE. (2007b). *PISA 2006. Science competencies for tomorrow 's world*. París: OECD publications.
- OCDE. (2009a). *PISA. Data Analysis Manual. SPSS, second edition*. París: OECD publications.
- OCDE. (2009b). *PISA 2006. Technical Report* París: OECD publications.
- OCDE. (2010). *Iberoamérica en PISA 2009. Informe regional*. Madrid: Santillana.
- Olsen, R. V., y Lie, S. B. (2011). Profiles of students' interest in science issues around the world: Analysis of data from PISA 2006. *International Journal of Science Education*, 33(1), 97-120.
- Orlich, D. C., Harder, R. J., Callahan, R. C., Trevisan, M. S., y Brown, A. H. (2010). *Teaching strategies: A guide to efective instruction*. Boston, MA: Wadsworth.
- Pajares-Box, R. (2005). *Resultados en España del estudio Pisa 2000. Conocimientos y destrezas de los alumnos de 15 años*. Madrid: INECSE.
- Pedró, F. (2010). La innovación sistémica en educación: Implicaciones políticas y organizativas. En A. Manzanares-Moya (Ed.), *Organizar y dirigir en la complejidad. Instituciones educativas en evolución* (pp. 675-688). Wolter Kluwer España, S.A.
- Pellegrino, J. W., Chudowsky, N., y Glaser, R. (2001). *Knowing what students know: The science and design of educational assesment*. Washington, DC: National Academy Press.
- Pennycuick, D. (1993). *School effectiveness in developing countries. A summary of the research evidence*. Londres: Overseas Development Agency.
- Peña-Suárez, E., Fernández-Alonso, R., y Muñiz, J. (2009). Estimación del valor añadido de los centros escolares. *Aula Abierta*, 37(1), 3-18.
- Plowden Committee (1967). *Children and their Primary schools*. Londres: HMSO.
- Porter, A. (1991). Creating a system of school process indicators. *Educational Evaluation and Policy Analysis*, 13(1), 13-29.

- Prais, S. J. (1983). Formal and informal teaching a further reconsideration of Professor Bennet's statistics. *Journal of the Royal Statistical Society ser. A* (146), 163-169.
- Purkey, S. C., y Smith, M. S. (1983). Effective schools: A review. *Elementary School Journal*, 4, 427-452.
- Raudenbush, S., Bryk, A., y Congdon, R. (2005). HLM 6.05. Chicago: Scientific Software International.
- Raudenbush, S., Bryk, A. S., Cheong, Y. F., y Congdon, R. (2004). *HLM 6: Linear and non linear modeling* Lincoldwood.
- Raudenbush, S., Bryk, T., y Congdon, R. (2000). HLM 6. Hierarchical Linear and Nonlinear Modeling.
- Raudenbush, S. W. (1989). The analysis of longitudinal, multilevel data. *International journal of educational research*, 13(7), 721-739.
- Raudenbush, S. W., y Bryk, A. S. (1986). A hierarchical model for studying school effects. *Sociology of Education*, 59(1-17).
- Raudenbush, S. W., y Bryk, A. S. (2002). *Hierarchical lenieal models: Applications and data analysis methods* (2ª ed.). Thousand Oaks, CA: Sage.
- Raudenbush, S. W., Bryk, A. S., y Congdon, R. (2005). HLM 6.05. Chicago: Scientific Software International.
- Reynolds, D. (1975). When teachers and pupils refuse a truce: The secondary school and the generation of delinquency. En G. Mingham. y G. Pearson. (Eds.), *British Working Class Youth Culture* (pp. 124-137). Londres: Routledge.
- Reynolds, D. (1976). The delinquent school. En P.Woods (Ed.), *The process of schooling* (pp. 124-178). Londres: Routledge.
- Reynolds, D. (2007). School Effectiveness and School Improvement (SESI): Links with the international standars/accountability agenda. En C. Teddlie y D. Reynolds (Eds.), *The international handbook of school effectiveness research* (pp. 471-485). Nueva York: Springer.
- Reynolds, D., Bollen, R., Creemers, B., Hopkins, D., Stoll, L., y Largerweik, N. (1996). *Making good schools. Linking school effectiveness and school improvement*. Londres: Routledge.
- Reynolds, D., Creemers, B., Hopkins, D., Stoll, L., y Bollen, R. (1997). *Las escuelas eficaces. Claves para mejorar la enseñanza*. Madrid: Aula XXI/Santillana.

- Reynolds, D., Creemers, B., Stringfield, S., Teddlie, C., y Schaffer, G. (2005). *World class schools: International perspectives on school effectiveness*. Londres: Routledge.
- Reynolds, D., Hopkins, D., y Stoll, L. (1993). Linking school effectiveness knowledge and school improvement practice: Towards a synergy. *School Effectiveness and School Improvement*, 4(1), 37-58.
- Reynolds, D., Mujis, D., y Treharne, D. (2003). Teacher evaluation and teacher effectiveness in the United Kingdom. *Journal of Personnel Evaluation in Education*, 17(1), 83-100.
- Reynolds, D., y Sullivan, M. (1979). Bringing schools back. In L. Barton (Ed.), *Schools, Pupils and Deviance* (pp. 89-129). Driffield: Nafferton.
- Reynolds, D., y Teddlie, C. (2000). The future agenda for school effectiveness research. En C.Teddlie. y D.Reynolds (Eds.), *The international handbook of School Effectiveness Research*. Londres: Falmer Press.
- Riddell, A. R. (1989). An alternative approach to the study of school effectiveness in third world countries. *Comparative Education Review*, 33(4), 481-497.
- Roberts, K. H., y Burstein, L. (1980). *New directions in methodology: Aggregation issues in organizational science*. San Francisco: Jossey-Bass.
- Roberts, M. (ed.). (2001). *Scientific literacy: Towards balance in setting goals for school science program*. Ottawa: Science Council of Canada.
- Robison, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review*, 15, 351-357.
- Rodríguez-Díez, B. (1990). *Modelo para la evaluación externa de los centros escolares de EGB*. Tesis no publicada, Universidad de Oviedo.
- Rodríguez-Gómez, G. (1991). *Investigación evaluativa en torno a los factores de eficacia escolar de los centros públicos de EGB*. Tesis no publicada, UNED.
- Rodríguez-Jiménez, O. R., y Murillo, F. J. (2011). Estimación del efecto escuela para Colombia. *Magis, Revista Internacional de Investigación en Educación*, 3(6), 299-316.
- Roman, M. (2008). Investigación latinoamericana sobre enseñanza eficaz, ILEE. In UNESCO (Ed.), *Eficacia escolar y factores asociados en América Latina y Caribe* (pp. 209-255). Santiago de Chile UNESCO.
- Romero, S., Ordóñez, X. G., López-Martín, E., y Navarro, E. (2009). Análisis de la estructura cognitiva de la competencia científica en PISA 2006 mediante el

- método de las distancias de mínimo-cuadráticas: El caso español. *Psicothema*, 21(4), 568-572.
- Rosenshine, B. (1983). Teaching functions in instructional programs. *Elementary school journal*, 89, 421-439.
- Rosenshine, B., y Stevens, R. (1986). Teaching functions. En M. C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 376-391). Nueva York: Macmillan.
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. Nueva York: Wiley
- Ruiz de Miguel, C., y Castro-Morera, M. U. (2006). Un estudio multinivel basado en PISA 2003: Factores de eficacia escolar en el área de matemáticas., 14(29).
- Rust, K. F., y Rao, J. N. K. (1996). Variance estimation for complex surveys using replication techniques. *Statist. Meth. Medical Research*, 5, 283-310.
- Rutter, M., Mortimore, P., Ouston, J., y Maughan, B. (1979). *Fifteen thousand hours*. Londres: Open Books.
- Rutter, M., Mortimore, P., Ouston, J., y Maughan, B. (1993). Differential school effectiveness: Results from a reanalysis of the inner Londres education Authority's Junior School Project Data. *British Educational Research Journal*, 4, 381-405.
- Rychen, D. S., y Salganik, L. H. (2003). *Key Competencies for a Successful Life and a Well-Functioning Society*. Göttingen: Hogrefe & Huber.
- Sammons, P., Hillman, J., y Mortimore, P. (1995). *Key characteristics of effective schools: A response of school effectiveness research*. Londres: OFSTED.
- Sammons, P., Nuttall, D. L., y Cuttance, P. (1993). Differential school effectiveness: Results from reanalysis of the inner Londres Education Authority's Junior School Project Data. *British Educational Research Journal*, 4, 381-405.
- Sammons, P., Thomas, S., y Mortimore, P. (1997). *Forging links: Effective schools and effective departments*. Londres: Paul Chapman Publishing.
- Saunders, L. (1999). A brief history of educational "value added": How did we get to where we are. *School Effectiveness and School Improvement*, 10(2), 233-256.
- Saunders, L. (2000). Value added: Telling the truth about schools' performance? *Topic papers. The world bank group*.
- Scheerens, J. (1990). School effectiveness reserch and the development of process indicators of school functioning. *School Effectiveness and School Improvement*, 1(1), 61-80.

- Scheerens, J. (1992). *Effective schooling research, theory and practice*. Londres: Cassell.
- Scheerens, J. (2001). School effectiveness in developing countries. *School Effectiveness y School Improvement*, 12(4), 353-358.
- Scheerens, J., y Bosker, R. (1997). *The foundations of educational effectiveness*. Oxford Pergamon.
- Scheerens, J., y Creemers, B. P. M. (1989). Towards a more comprehensive conceptualization of school effectiveness. En B. P. M. Creemers, T. Peters y D. Reynolds (Eds.), *School Effectiveness and school improvement* (pp. 265-278). Lisse: Swtes yZeitlinger.
- Schweitzer, J. (1984). Characteristics of effective schools. American Educational Research Association Annual meeting. Nueva Orleans.
- Schibeci, R. A. (1984). Attitude to science: An update. *Studies in Science Education*, 11, 26-59.
- Sheffield, P., y Saunders, S. (2002). Using the british education index to survey the field of educational studies. *British Journal of Educational Studies*, 50(1), 165-183.
- Shepard, L. A. (1989). Why we need better assesment. *Educational Leadership*, 46(2), 4-8.
- Shuerger, J. M., y Kuma, D. L. (1987). Adolescent personality and school performance: A follow-up study. *Psychology in the Schools*, 24, 281-285.
- Slater, R., y Teddlie, C. (1992). Toward a theory of school effectiveness and leadership. *School Effectiveness and School Improvement*, 3(4), 242-257.
- Slavin, R. E. (1987). A theory of school and classroom organization. *Educational Psychologist*, 22(2), 89 -108.
- Slavin, R. E., Madden, N. A., Dolan, L. J., Wasik, B. A., Ross, S., Smith.I., et al. (1996). Sucess for all: A summary research. *Journal of Education for Students Placed at Risk*, 1(1), 41-76.
- Smith, D. J., y Tomlinson, S. (1989). *The school effect: A study of multi-racial comprehensives*. Londres: Policy Studies Institute.
- Smith, M. (1972). Equality of educational opportunity: the basic findings reconsidered. En F. Mosteller. y D. P. Moyniban. (Eds.), *On equality if educational opportunity*. Nueva York: Vintage Books.
- Snijders, T. A. B., y Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modelling*. Londres: SAGE.

- Snow, R. E., y Lohman, D. E. (1984). Toward a theory of cognitive aptitude for learning from instruction. *Journal of Educational Psychology*, 76, 347-376.
- Spartz, J. L., Valdes, A. L., McCornick, W. L., Myers, J., y Geppert, W. J. (1977). *Delaware educational accounting system case studies: Elementary school grades 1-4*. Dover: Delaware department of public instruction.
- Stenmark, J. K. (1992). *Mathematics assesment: Myths, models, good questions and practical suggestions*. Reston, VA.: NCTM.
- Sternberg, R. J. (1988). Mental self-government: A theory of intellectual styles and their development. *Human Development*, 31, 197-224.
- Sticht, T. G. (ed.). (1975). *Reading for working: A functional literacy anthology*. Alexandria, VA.: Human Resources Research Organization.
- Stiggins, R. J. (1982). An analysis of the dimensions of job-related reading. *Reading World*, 82, 237-247.
- Stoll, L. (2009). Capacity building for school improvement or building capacity for learning? *A changing landscape*. *Journal of Educational Change*, 10, 115-127.
- Stoll, L., y Fink, D. (1996). *Changing our schools. Linking school effectiveness and school improvement*. Buckingham: Open University Press.
- Stringfield, S., y Slavin, R. E. (1992). A hierarchical longitudinal model for elementary school effects. En B. P. M. Creemers. y G. J. Reezigt. (Eds.), *Evaluation of Educational Research* (pp. 35-69). Groningen: ICO.
- Stringfield, S., y Teddlie, C. (1989). The first three phases of the Louisiana school effectiveness study. En B. P. M. Creemers, T. Peters y D. Reynolds (Eds.), *School effectiveness and school improvement*. Lisse: Swets y Zeitlinger.
- Stufflebeam, D. L., y Shinkfield, A. J. (1987). *Evaluación sistemática. Guía teórica y práctica*. Barcelona: Centro de Publicaciones del MEC y Piados Ibérica.
- Teddlie, C. (1984). *The Louisiana school effectiveness study: Phase two 1982-1984*. Baton Rouge, Luisiana: Louisiana State Department of Education.
- Teddlie, C., Kirby, P. C., y Stringfield, S. (1989). Effective versus ineffective schools: Observable differences in the classroom. *American journal of education*, 97(3), 221-236.
- Teddlie, C., y Reynolds, D. (2000). *The International Handbook of School Effectiveness Research*. Londres: Falmer Press.
- Teddlie, C., Reynolds, D., y Sammons, P. (2000). The methodology and scientific properties of school effectiveness research. En C. Teddlie y D. Reynolds (Eds.),

- The International Handbook of School Effectiveness Research* (pp. 55-133). Londres: Falmer Press.
- Teddlie, C., y Stringfield, S. (1993). *Schools make a difference: lessons learned from a ten year study of school effects*. Nueva York: Teachers College Press.
- Thomas, S., Sammons, P., y Mortimore, P. (1997). Stability and consistency in secondary school's effects on student's GCSE outcomes over three years *School Effectiveness y School Improvement*, 8(2), 169-197.
- Towsend, T., y Otero, G. (1999). *The global classroom: Engaging students in third millenium schools*. Melbourne: Hawker Brownlow.
- Towsend, T. (ed). (2007). *International handbook of school effectiveness and improvement*. Nueva York: Springer.
- Tristán-López, A. (coord.) (2008). *Análisis multinivel de la calidad educativa en México ante los datos PISA 2006*. México: INEE.
- Turner, R. (2006). El programa internacional para la evaluación de los alumnos PISA. Una perspectiva general. *Revista de Educación, extraordinario 2006*, 45-74.
- Turner, R. (2006). El programa para la evaluación internacional de alumnos (PISA). Una perspectiva general. *Revista educación, Número extraordinario.*, 45-74.
- UNESCO. (1993). *International forum on scientific and technological literacy for all: Final Report*. París: UNESCO.
- Vallejo, G., Arnau, J., y Bono, R. (2008). Construcción de modelos jerárquicos en contextos aplicados. *Psicothema*, 20(004), 830-838.
- van der Vijver, F. J. R. (2011). New developments in methodology for cross-cultural research. *XII Congreso de Metodología de las Ciencias Sociales y de la Salud*. San Sebastián.
- Venezky, R., y Winfield, L. (1979). *Schools that succeed beyond expectations in teaching reading*. Newark: University of Delaware.
- Virgilio, I., Teddlie, C., y Oescher, J. (1991). Variance and context differences in teaching at differentially effective schools. *School Effectiveness and School Improvement*, 2(2), 152-168.
- Wainer, H., Bradlow, E. T., y Wang, X. (2007). *Testlet response theory and its applications*. Nueva York. : Cambridge University Prress.
- Wang, M. C., Haertel, G. D., y Walberg, H. J. (1993). Toward a knowledge base for school learning. . *Review of educational research*, 63(3), 249-294.

- Waters, T., Marzano, R. J., y McNulty, B. (2003). *Balanced leadership: What 30 years of research tells us about leadership on student achievement*. Consultado el 20/05/11, en <http://mcrel.org>
- Weber, G. (1971). *Inner-city children can be taught to read: Four successful schools*. Washington, DC: Council for Basic Education.
- Webster, W. J., Mendro, L., y Almaguer, T. O. (1994). Effectiveness indices: A value-added approach to measuring school effect. *Studies in Educational Evaluation*, 20, 113-145.
- Welch, A. R. (ed.). (2000). *Third world education: Quality and equality*. Nueva York: Garland Pub.
- Wilder, G. (1977). Five exemplary reading programs. En J. T. Guthrie (Ed.), *Cognition, curriculum and comprehension*. (pp. 57-68). Newark, DE.: International Reading Association.
- Willms, J. D., y Raudenbush, S. W. (1989). A longitudinal hierarchical linear model for estimating school effects and their stability. *Journal of Education Measurement*, 26(3), 209-232.
- Wilms, D. (1992). *Monitoring school performance: A guide for educators*. Londres: Falmer.
- Wilms, D., y Smith, T. (2005). *A manual for conducting analyses with data from TIMSS and PISA*. Consultado el 31/05/11, en http://www.datafirst.uct.ac.za/wiki/images/e/e7/TIMSS_1995-99_Manual.pdf
- Wittrok, M. (1986). *Handbook of research teaching* New Cork: MacMillan.
- Wu, M. L., Adams, R. J., y Wilson, M. R. (1997). *Conquest: Multi-aspect test software [programa informático]*. Camberwell, Australia: Australian Council for Education Research.
- Youden, W. J. (1937). Use of incomplete block replications in estimating tobacco-mosaic virus. *Contributions from Boyce Thompson Institute.*, 9, 41-48.
- Youden, W. J. (1940). Experimental designs to increase accuracy of greenhouse studies. *Contributions from Boyce Thompson Institute.*, 11, 219-228.

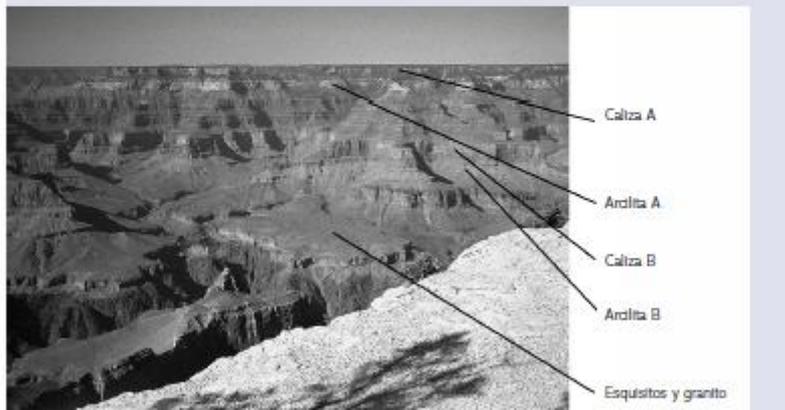
||| ANEXOS

Se recogen los estímulos, los ítems y los criterios de corrección. El texto y las preguntas sombreadas se refieren a la evaluación de las actitudes relacionadas con la competencia científica.

EL GRAN CAÑÓN

El Gran Cañón está situado en un desierto de los Estados Unidos. Es un cañón muy largo y profundo que contiene muchos estratos de rocas. En algún momento del pasado, los movimientos de la corteza terrestre levantaron estos estratos. Hoy en día el Gran Cañón tiene 1,6 km de profundidad en algunas zonas. El río Colorado fluye por el fondo del cañón.

Mira la siguiente foto del Gran Cañón, tomada desde su orilla sur. En las paredes del cañón se pueden ver los diferentes estratos de rocas.



Pregunta 7: EL GRAN CAÑÓN S426Q07

Cada año unos cinco millones de personas visitan el parque nacional del Gran Cañón. Existe preocupación por el deterioro que está sufriendo el parque debido al elevado número de visitantes.

¿Es posible responder las preguntas siguientes mediante una investigación científica? Marca con un círculo la respuesta, Sí o No, para cada pregunta.

¿Es posible responder esta pregunta mediante una investigación científica?	¿Sí o No?
¿Qué cantidad de erosión se produce por la utilización de las pistas forestales?	Sí / No
¿El parque es tan bello como lo era hace 100 años?	Sí / No

EL GRAN CAÑÓN; puntuación de la pregunta 7

Máxima puntuación

Código 1: Las dos respuestas son correctas: Sí, No, en este orden.

Ninguna puntuación

Código 0: Otras respuestas.

Código 9: Sin respuesta.

Pregunta 3: EL GRAN CAÑÓN S426Q03

La temperatura en el Gran Cañón varía de menos de 0 °C a más de 40 °C. Aunque la zona es desértica, las grietas de las rocas a veces contienen agua. ¿De qué manera estos cambios de temperatura y la presencia de agua en las grietas de las rocas contribuyen a acelerar el desmenuzamiento de las rocas?

- A El agua congelada disuelve las rocas calientes.
- B El agua cementa a las rocas entre sí.
- C El hielo pule la superficie de las rocas.
- D El agua congelada se dilata en las grietas de las rocas.

EL GRAN CAÑÓN; puntuación de la pregunta 3

Máxima puntuación

Código 1: D. El agua congelada se dilata en las grietas de las rocas.

Ninguna puntuación

Código 0: Otras respuestas.

Código 9: Sin respuesta.

Pregunta 5: EL GRAN CAÑÓN 5426Q05

En el estrato de caliza A del Gran Cañón se encuentran muchos fósiles de animales marinos, como almejas, peces y corales. ¿Qué sucedió hace millones de años para que aparezcan estos fósiles en este estrato?

- A Antiguamente los habitantes transportaban alimentos marinos desde el océano a esta área.
- B En otro tiempo, los océanos eran más violentos, y olas gigantes arrastraban criaturas marinas hacia el interior.
- C En esa época, la zona estaba cubierta por un océano que más tarde se retiró.
- D Algunos animales marinos vivieron una vez sobre la tierra antes de emigrar al mar.

EL GRAN CAÑÓN; puntuación de la pregunta 5**Máxima puntuación**

Código 1: C. En esa época, la zona fue cubierta por el mar y más tarde se retiró.

Ninguna puntuación

Código 0: Otras respuestas.

Código 9: Sin respuesta.

Pregunta 10S: EL GRAN CAÑÓN 5426Q10S

¿En qué medida estás de acuerdo con las afirmaciones siguientes?

Marca sólo una casilla en cada fila.

	Totalmente de acuerdo	De acuerdo	En desacuerdo	Totalmente en desacuerdo
a) El estudio sistemático de los fósiles es importante.	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
b) Las medidas de protección contra el deterioro de los parques naturales deben basarse en datos científicos.	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4
c) El estudio científico de los estratos geológicos es importante.	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4

EL EJERCICIO FÍSICO

El ejercicio físico practicado con regularidad, pero con moderación, es bueno para la salud.



Pregunta 1: EL EJERCICIO FÍSICO s493Q01

¿Cuáles son los beneficios del ejercicio físico practicado con regularidad? Marca con un círculo la respuesta Sí o No para cada afirmación.

¿Es lo siguiente un beneficio del ejercicio físico practicado con regularidad?	¿Sí o No?
El ejercicio físico ayuda a prevenir las enfermedades del corazón y los problemas circulatorios.	Sí / No
El ejercicio físico hace que tengas una dieta saludable.	Sí / No
El ejercicio físico ayuda a prevenir la obesidad.	Sí / No

EL EJERCICIO FÍSICO; puntuación de la pregunta 1**Máxima puntuación**

Código 1: Las tres respuestas son correctas: Sí, No, Sí, en este orden.

Ninguna puntuación

Código 0: Otras respuestas.

Código 9: Sin respuesta.

Pregunta 3: EL EJERCICIO FÍSICO S493Q03

¿Qué sucede cuando se ejercitan los músculos? Marca con un círculo la respuesta, Sí o No, para cada afirmación.

¿Sucede esto cuando se ejercitan los músculos?	¿Sí o No?
Los músculos reciben un mayor flujo de sangre.	Sí / No
Se forma grasa en los músculos.	Sí / No

EL EJERCICIO FÍSICO; puntuación de la pregunta 3**Máxima puntuación**

Código 1: Las dos respuestas son correctas: Sí, No, en este orden.

Ninguna puntuación

Código 0: Otras respuestas.

Código 9: Sin respuesta.

Pregunta 5: EL EJERCICIO FÍSICO S493Q05 - 01 11 12 99

¿Por qué respiras más fuerte cuando haces ejercicio físico que cuando tu cuerpo está en reposo?

.....

EL EJERCICIO FÍSICO; puntuación de la pregunta 5

Máxima puntuación

Código 11: Para disminuir la cantidad de dióxido de carbono, que ha «aumentado», Y para suministrar «más» oxígeno al cuerpo. [No debe aceptarse aire por dióxido de carbono u oxígeno].

- Cuando haces ejercicio necesitas más oxígeno y produces más dióxido de carbono. La respiración sirve para esto.
- Respirar más rápido permite que entre más oxígeno en la sangre y que se elimine más dióxido de carbono.

Código 12: Para disminuir la cantidad de dióxido de carbono del cuerpo, que ha «aumentado» O para aportar «más» oxígeno al cuerpo pero no ambas.[No debe aceptarse aire por dióxido de carbono u oxígeno].

- Porque debemos deshacernos del dióxido de carbono que se forma.
- Porque los músculos necesitan oxígeno. [Implica que el cuerpo necesita más oxígeno cuando se hace ejercicio (utilizando los músculos)].
- Porque el ejercicio físico consume oxígeno.
- Se respira más fuerte porque llega más oxígeno a los pulmones. [Está mal expresado, pero reconoce que hay una mayor aportación de oxígeno].
- Como se utiliza bastante energía, el cuerpo necesita el doble o el triple de aire y también necesita eliminar el dióxido de carbono. [El Código 12 se asigna por la segunda frase, la cual implica que el cuerpo debe eliminar más dióxido de carbono de lo normal. La primera frase no contradice a la segunda, pero sola, recibiría el Código 01].

Ninguna puntuación

Código 01: Otras respuestas.

- Para que entre más aire en los pulmones.
- Porque los músculos consumen más energía. [No es suficientemente precisa].
- Porque el corazón late más rápido.
- Porque el cuerpo necesita oxígeno. [No se refiere a la necesidad de más oxígeno].

Código 99: Sin respuesta.

Índice	Ítems
Nivel 1	
	¿Estás interesado en aprender sobre los siguientes temas científicos? Escala: mucho (1); bastante (2); poco (3); nada (4)
INTSCIE	<ol style="list-style-type: none"> 1. Física 2. Química 3. Biología vegetal 4. Biología humana 5. Astronomía 6. Geología 7. La forma en que los científicos diseñan sus experimentos 8. Requisitos para formular una explicación científica
	¿Estás de acuerdo con las siguientes afirmaciones? Escala: mucho (1); bastante (2); poco (3); nada (4)
JOYSCIE	<ol style="list-style-type: none"> 1. Me lo paso bien aprendiendo temas científicos 2. Me gustan los libros de ciencia 3. Disfruto haciendo problemas de ciencias 4. Me encanta adquirir nuevos conocimientos relacionados con la ciencia 5. Estoy interesado en aprender temas de ciencias
	¿Estás de acuerdo con las siguientes afirmaciones? Escala: mucho (1); bastante (2); poco (3); nada (4)
INSTSCIE	<ol style="list-style-type: none"> 1. Hacer un esfuerzo en las clases de ciencias merece la pena porque me ayudará a trabajar en lo que quiero dedicarme en un futuro 2. Lo que aprendo en las clases de ciencias es importante para lo que quiero estudiar más adelante 3. Estudio ciencias porque sé que va a ser útil para mí 4. Merece la pena estudiar las asignaturas de ciencias porque lo que aprendo mejorará las perspectivas de mi carrera profesional 5. Aprender muchas cosas de ciencias me ayudará para conseguir un trabajo
	¿Estás de acuerdo con las siguientes afirmaciones? Escala: mucho (1); bastante (2); poco (3); nada (4)
SCIEFUT	<ol style="list-style-type: none"> 1. Me gustaría trabajar en una profesión que tuviera que ver con la ciencia 2. Me gustaría estudiar una carrera de ciencias en la universidad 3. Me gustaría dedicar mi vida a la ciencia de alto nivel 4. Cuando sea mayor. Me gustaría trabajar en proyectos científicos
	¿Serías capaz de realizar sin ayuda las tareas siguientes? Escala: con facilidad (1); con un poco de esfuerzo (2); con bastante esfuerzo (3); no podría (4)
SCIEEFF	<ol style="list-style-type: none"> 1. Reconocer el problema científico que subyace en un artículo periodístico sobre un tema de salud 2. Explicar por qué ocurren terremotos con más frecuencia en unas zonas que en otras 3. Describir el papel de los antibióticos en el tratamiento de una enfermedad 4. Identificar el problema científico asociado a la eliminación de basuras 5. Predecir cómo los cambios ambientales podrían afectar a la supervivencia de ciertas especies 6. Interpretar información científica contenida en las etiquetas de los productos alimenticios 7. Debatir si la aparición de nuevos datos puede hacerte cambiar la opinión sobre la posibilidad de vida en Marte 8. Identificar la mejor de dos explicaciones sobre la formación de la lluvia ácida
	¿Estás de acuerdo con las siguientes afirmaciones? Escala: mucho (1); bastante (2); poco (3); nada (4)
SCSCIE	<ol style="list-style-type: none"> 1. Aprendería con facilidad contenidos avanzados de ciencias 2. Normalmente contesto bien a las preguntas en los exámenes de ciencias 3. Aprendo temas de ciencias rápidamente 4. Se me da bien los temas de ciencias 5. En las clases de ciencias, entiendo muy bien los conceptos 6. Entiendo con facilidad los conceptos nuevos de ciencias
	¿Estás de acuerdo con las siguientes afirmaciones? Escala: mucho (1); bastante (2); poco (3); nada (4)
GENSCIE	<ol style="list-style-type: none"> 1. Los avances en ciencia y tecnología suelen mejorar las condiciones de vida de las personas 2. La ciencia es importante para ayudarnos a comprender el mundo natural que nos rodea 3. Los avances en ciencia y tecnología ayudan a mejorar la economía 4. La ciencia tiene un valor para la sociedad 5. Los avances en ciencia y tecnología suelen proporcionar beneficios sociales

	<p>¿Estás de acuerdo con las siguientes afirmaciones? Escala: mucho (1); bastante (2); poco (3); nada (4)</p>
PERSCIE	<ol style="list-style-type: none"> 1. Algunos conceptos me ayudan a ver cómo me relaciono con otras personas 2. Cuando sea adulto, utilizaré la ciencia de muchas maneras 3. La ciencia es muy importante para mí 4. Considero que la ciencia me ayuda a entender las cosas que me rodean 5. Cuando termine mi estudios tendré muchas oportunidades de utilizar la ciencia
	<p>¿Con qué frecuencia haces lo siguiente? Escala: con mucha frecuencia (1); habitualmente (2); a veces (3); nunca o casi nunca (4)</p>
SCIEACT	<ol style="list-style-type: none"> 1. Ver programas científicos en televisión 2. Pedir prestado o comprar libros sobre ciencia en general 3. Visitar páginas web dedicadas a temas científicos 4. Escuchar programas de radio sobre avances en ciencia 5. Leer revistas científicas o artículos de periódico sobre ciencias 6. Asistir a un taller sobre ciencias
	<p>¿Estás informado sobre los siguientes temas medioambientales? Escala: nunca he oído hablar de este tema (1); he oído hablar de este tema, pero no sabría explicar de qué se trata (2); sé algo sobre este tema y podría explicarlo a grandes rasgos (3); conozco este tema y sabría explicarlo bastante bien (4)</p>
ENVAWARE	<ol style="list-style-type: none"> 1. Incremento de los gases de efecto invernadero en la atmósfera 2. Uso de organismos modificados genéticamente 3. Lluvia ácida 4. Residuos nucleares 5. Consecuencias de la tala bosques para dar otra utilización al suelo
	<p>¿Constituyen los siguientes temas medioambientales una preocupación importante para ti y para otras personas? Escala: Es una preocupación importante para mí y para otras personas (1); es una preocupación para otras personas en mi país pero no para mí (2); es una preocupación importante sólo para personas de otros países (3); no es una preocupación importante para nadie (4)</p>
ENVPERC	<ol style="list-style-type: none"> 1. Contaminación de aire 2. Escasez de energías 3. Extinción de plantas y animales 4. Tala de bosques para dar otra utilización al suelo 5. Escasez de agua 6. Residuos nucleares
	<p>En los próximos 20 años ¿crees que mejorarán o empeorarán los problemas relacionados con los siguientes temas medio ambientales? Escala: mejorarán (1); seguirán casi igual (2); empeorarán (3)</p>
ENVOPT	<ol style="list-style-type: none"> 1. Contaminación de aire 2. Escasez de energías 3. Extinción de plantas y animales 4. Tala de árboles para darle otro uso a la tierra 5. Escasez de agua 6. Desechos nucleares
	<p>¿Estás de acuerdo con las siguientes afirmaciones? Escala: mucho (1); bastante (2); poco (3); nada (4)</p>
RESPDEV	<ol style="list-style-type: none"> 1. Es importante realizar controles periódicos de los gases de escape de los automóviles para utilizar su circulación 2. Me molesta malgastar energía cuando se un uso innecesario de los aparatos electrónicos 3. Estoy a favor de que haya leyes que regulen las emisiones de las fábricas, aunque ello conlleve un incremento del precio de los productos 4. Debería reducirse al mínimo el uso de envases de plástico para disminuir la cantidad de basura 5. Se debería obligar a las fábricas a demostrar que eliminan sus desechos peligrosos en condiciones totalmente seguras 6. Estoy a favor de que haya leyes que protejan el hábitats de las especies en peligro de extinción 7. Siempre que sea posible se debe producir electricidad a partir de fuentes renovables, aunque salga más cara
	<p>¿Estás de acuerdo con las siguientes afirmaciones? Escala: mucho (1); bastante (2); poco (3); nada (4)</p>
CARPREP	<ol style="list-style-type: none"> 1. Las asignaturas que se imparten en mi centro proporcionan a los alumnos las destrezas y los conocimientos iniciales para una profesión relacionada con la ciencia 2. Las asignaturas de ciencias de mi centro proporcionan a los alumnos las destrezas y los conocimientos iniciales para una profesión relacionada con la ciencia 3. Las materias que estudio me dan habilidades y conocimientos básicos para una profesión relacionada con las ciencias 4. Mis profesores me proporcionan las destrezas y conocimientos para ejercer una profesión científica

	<p>¿Estás informado sobre los siguientes temas? Escala: mucho (1); bastante (2); poco (3); nada (4)</p>
CARINFO	<ol style="list-style-type: none"> 1. Las profesiones relacionadas con la ciencia que ofrece el mercado laboral 2. Dónde encontrar información sobre carreras relacionadas con las ciencias 3. Los pasos que los alumnos deben seguir para orientarse hacia una profesión relacionada con la ciencia 4. Los patronos o las empresas que contratan a gente con carreras relacionadas con la ciencia
	<p>¿Con qué frecuencia se dan las siguientes actividades en las clases de ciencias? Escala: en todas las clases (1); en la mayoría de las clases (2); en algunas clases (3); nunca o casi nunca (4)</p>
SCINTACT	<ol style="list-style-type: none"> 1. A los alumnos se les da la oportunidad de exponer sus ideas 2. Los alumnos pasan tiempo en el laboratorio realizando experimentos prácticos 3. A los alumnos se les pide que piensen en cómo investigar en el laboratorio un problema de ciencias 4. A los alumnos se les pide que apliquen un tema de ciencias a los problemas de la vida diaria 5. En las clases se tiene en cuenta las opiniones de los alumnos sobre los temas 6. Se pide a los alumnos que obtengan conclusiones sobre un experimento que hayan realizado 7. El profesor realiza experimentos a modo de demostración
	<p>¿Cuándo hay clase de ciencias con qué frecuencia ocurren las siguientes acciones? Escala: en todas las clases (1); en la mayoría de las clases (2); en algunas clases (3); nunca o casi nunca (4)</p>
SCHANDS	<ol style="list-style-type: none"> 1. Los estudiantes invierten tiempo haciendo experimentos 2. Los estudiantes deben plantear tópicos que se puedan investigar en el laboratorio 8. Se pide a los alumnos que obtengan conclusiones sobre un experimento que hayan realizado 3. Los alumnos hacen experimentos siguiendo las instrucciones de un profesor
	<p>¿Cuándo hay clase de ciencias con qué frecuencia ocurren las siguientes acciones? Escala: en todas las clases (1); en la mayoría de las clases (2); en algunas clases (3); nunca o casi nunca (4)</p>
SCINVEST	<ol style="list-style-type: none"> 1. A los alumnos se les permite diseñar sus propios experimentos 2. A los alumnos se les da la oportunidad de elegir sus propias investigaciones 3. A los alumnos se les pide que realicen una investigación para demostrar sus propias ideas
	<p>¿Cuándo hay clase de ciencias con qué frecuencia ocurren las siguientes acciones? Escala: en todas las clases (1); en la mayoría de las clases (2); en algunas clases (3); nunca o casi nunca (4)</p>
SCAPPLY	<ol style="list-style-type: none"> 1. El profesor explica cómo una fórmula o idea científica puede aplicarse a varios fenómenos diferentes (p.ej., movimiento de los objetos, sustancias con propiedades similares) 2. El profesor usa la ciencia para ayudarnos a entender el mundo más allá del centro escolar 3. El profesor explica con claridad la relevancia de conceptos científicos en la vida de las personas 4. El profesor utiliza ejemplos de aplicaciones tecnológicas para mostrar que la ciencia es importante para la sociedad
Nivel 2	
	<p>¿Se ve dificultada la capacidad del centro para proporcionar enseñanza por alguno de factores siguientes? Escala: mucho (1); bastante (2); poco (3); nada (4)</p>
TCSHORT	<ol style="list-style-type: none"> 1. Falta de profesorado cualificado de Ciencias 2. Falta de profesorado cualificado de Matemáticas 3. Falta de profesorado cualificado de Lengua 4. Falta de profesorado cualificado de otras asignaturas
	<p>¿Se ve dificultada la capacidad del centro para proporcionar enseñanza por alguno de factores siguientes? Escala: mucho (1); bastante (2); poco (3); nada (4)</p>
SMATEDU	<ol style="list-style-type: none"> 1. Falta de técnicos de laboratorio 2. Falta de otro personal de apoyo 3. Inadecuación o escasez de material de enseñanza (p.ej., libros de texto) 4. Inadecuación o escasez de material del laboratorio de Ciencias 5. Inadecuación o escasez de ordenadores para la enseñanza 6. Inadecuación o ausencia de conexión a Internet 7. Inadecuación o escasez de programas informáticos para la enseñanza 8. Inadecuación o escasez de materiales de consulta en la biblioteca 9. Inadecuación o escasez de medios audiovisuales
	<p>¿Se realizan las siguientes actividades en el centro para promover la implicación de los alumnos de 4º de la ESO con las ciencias? Escala: sí (1), no (2)</p>
SCIPROM	<ol style="list-style-type: none"> 1. Talleres de ciencias 2. Exposiciones científicas 3. Concursos científicos 4. Proyectos científicos extracurriculares (incluidos trabajos de investigación) 5. Excursiones y viajes de estudio

¿Se organizan las siguientes actividades en el centro para que los alumnos de 4º de ESO tengan oportunidad de aprender sobre temas medioambientales?

Escala: sí (1), no (2)

ENVLEARN

1. Clases o actividades al aire libre
 2. Salidas a museos
 3. Salidas a centros científicos o tecnológicos
 4. Proyectos científicos extracurriculares (incluidos los trabajos de investigación)
 5. Conferencias o seminarios (p.ej., trayendo ponentes invitados)
-

Estadísticos descriptivos N1

Variable	N	M	D.T.	Mínimo	Máximo
ST_WGT55	17505	0.18	0.07	0.05	0.70
PV1SCIE	17505	506.16	85.87	4.76	920.67
PV2SCIE	17505	506.26	85.87	52.34	924.40
PV3SCIE	17505	506.14	86.35	4.76	884.28
PV4SCIE	17505	506.06	86.50	4.76	883.35
PV5SCIE	17505	506.58	86.75	82.20	952.39
ESCS	17505	-0.16	1.02	-5.07	2.81
ESCS2	17505	1.07	1.24	0.00	25.70
FEMALE	17505	0.50	0.50	0.00	1.00
NATIVE	17505	0.94	0.24	0.00	1.00
MNATIVE	17505	0.01	0.08	0.00	1.00
SAMELANG	17505	0.95	0.21	0.00	1.00
REP	17505	0.31	0.46	0.00	1.00
PARED	17505	0.13	0.34	0.00	1.00
INTSCIE	17505	-0.18	1.00	-3.14	3.29
JOYSCIE	17505	-0.18	0.99	-2.15	2.06
INSTSCIE	17505	0.07	1.07	-2.10	1.82
SCIEFUT	17505	0.10	1.07	-1.42	2.27
SCIEEFF	17505	-0.04	1.07	-3.77	3.22
SCSCIE	17505	0.01	1.02	-2.36	2.24
GENSCIE	17505	0.32	0.99	-3.66	2.19
PERSSCIE	17505	0.04	0.97	-3.08	2.53
SCIEACT	17505	-0.18	0.99	-1.69	3.38
ENVAWARE	17505	0.15	0.95	-3.44	3.01
ENVPERC	17505	0.54	0.87	-4.11	1.39
ENVOPT	17505	0.14	0.99	-1.61	2.85
RESPDEV	17505	0.28	0.91	-4.00	2.30
CARPREP	17505	0.05	0.97	-2.92	1.96
CARINFO	17505	0.03	0.98	-2.44	2.53
SCINTACT	17505	-0.02	0.95	-2.51	2.47
SCHANDS	17505	-0.35	0.98	-2.10	2.91
SCINVEST	17505	-0.16	0.98	-1.26	3.03
SCAPPLY	17505	-0.04	1.00	-2.46	2.63

Estadísticos descriptivos N2

Variable	N	M	D.T.	Mínimo	Máximo
XESCS	610	-0.21	0.55	-1.59	1.42
XRURAL	610	0.36	0.48	0.00	1.00
XCITY	610	0.39	0.49	0.00	1.00
PRIV	610	0.44	0.50	0.00	1.00
GOVFUND	610	86.32	19.71	0.00	100.00
SCHSIZE	610	6.46	3.94	0.34	25.39
SCHSIZE2	610	57.32	77.04	0.12	644.65
SCHSIZE	610	0,34	0.77	0.13	9,40
SCHSIZE2	610	0,79	7.10	0.02	88.36
LOSELE	610	0.83	0.37	0.00	1.00
HISELE	610	0.00	0.06	0.00	1.00
ABGROUP	610	0.13	0.34	0.00	1.00
PRESSPA	610	0.38	0.49	0.00	1.00
FACS	610	0.08	1.01	-1.26	1.48
FACB	610	0.02	1.01	-4.85	0.52
FACC	610	0.14	1.01	-3.45	1.74
STRATIO	610	11.73	4.36	1.19	29.60
TCSHORT	610	-0.66	0.72	-1.06	3.62
IRATCOM	610	0.11	0.08	0.02	0.72
SCMATED	610	0.08	0.94	-3.43	2.14
TIMELEARN	610	5.46	1.29	2.09	11.00
ANYSCIE	610	90.57	10.56	26.12	100.00
ENVLEARN	610	2.34	0.96	0.00	4.00
SCIPROM	610	0.09	0.85	-2.27	1.64
ACC1	610	0.42	0.49	0.00	1.00
ACC2	610	0.10	0.31	0.00	1.00
ACC3	610	0.09	0.29	0.00	1.00
ACC4	610	0.08	0.27	0.00	1.00
ACC5	610	0.10	0.29	0.00	1.00
ACC6	610	0.39	0.49	0.00	1.00
ACC7	610	0.42	0.49	0.00	1.00
ACC8	610	0.67	0.47	0.00	1.00
XPRESSPA	610	0.38	0.49	0.00	1.00
MPRESSP	610	0.01	0.09	0.00	1.00